AD-A270 508

DOT/FAA/AM-93/15

Office of Aviation Medicine
Washington, D.C. 20591

Human Factors in Aviation Maintenance – Phase Three, Volume 1 Progress Report

Galaxy Scientific Corporation
Pleasantville, NJ 08323

Office of Aviation Medicine
Federal Aviation Administration
Washington, DC 20591

August 1993

Final Report

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

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-93/15		

4. Title and Subtitle	5. Report Date
Human Factors in Aviation Maintenance - Phase Three, Volume I Progress Report	August 1993
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Galaxy Scientific Corporation	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Galaxy Scientific Corporation 2500 English Creek Avenue, Suite 1100 Pleasantville, NJ 08323	11. Contract or Grant No. DTFA01-92-Y-01005
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, DC 20591	Interim February 1992- December 1992
	14. Sponsoring Agency Code

15. Supplemental Notes

16. Abstract

The third phase of research on Human Factors in Aviation Maintenance continued to look at the human's role in the aviation maintenance system via investigations, demonstrations, and evaluations of the research program outputs. This report describes an evaluation of a computer-based training simulation for troubleshooting an airliner environmental control system and an evaluation of the aircraft maintenance visual environment. A job aid for Aviation Safety Inspectors is also reported on, along with an evaluation of pen computers considered for the job aid. This progress report also describes research on ergonomic factors related to posture and fatigue; identification of characteristics of personnel best suited for inspection-oriented jobs; redesign of work control cards; and visual inspection and training alternatives in aviation maintenance.

17. Key Words	18. Distribution Statement
human factors, maintenance, computer-based instruction, simulation visual environment, job aid, posture, fatigue, workcards, training, NDI performance	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	167	

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

Acknowledgments

The Aviation Maintenance Human Factors research team is directed by Dr. William T. Shepherd and Ms. Jean Watson of the Office of Aviation Medicine.

Galaxy Scientific Corporation is the prime contractor. Galaxy Scientific is technically responsible for Chapters One through Three. Dr. Richard Thackray prepared Chapter Four. Chapters Five through Eight were prepared by the Department of Industrial Engineering at the State University of New York at Buffalo. The total report was edited by Ms. Suzanne Morgan of Galaxy Scientific Corporation.

The pragmatic nature of this research is attributable to many FAA, airline, manufacturer, and vendor personnel who provided advice to the research team.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AANC .... Aging Aircraft NDI Development and Demonstration Center
ANOVA .... ANalysis Of VAriance
ASIs .... Aviation Safety Inspectors

BITE .... built-in test equipment
BPD .... Body Part Discomfort Chart

CAI .... computer-aided instruction
CAMI .... Civil Aeromedical Institute
CBT .... computer-based training
CD-ROM .... compact disc, read only memory
CPU .... central processing unit

ECS .... Environmental Control Simulation
EFT .... Embedded Figures Test
EPRI .... Electric Power Research Institute

FAA .... Federal Aviation Administration
f-c .... foot-candles
FIM .... Fault Isolation Manual
f-l .... foot-lamberts
FOV .... field of view
FTC .... Feeling Tone Checklist

HIT .... Harris Inspection Test
HTA .... hierarchical task analysis

IES .... Illuminating Engineering Society
IIMS .... Integrated Information Management System
ITS .... intelligent tutoring system

KSA .... knowledge, skills, and abilities

LC .... Locus of Control
LRU .... line replaceable unit

MFFT .... Matching Familiar Figures Test
MSACL .... Modified Stress-Arousal Checklist

NEC .... National Electrical Code
NDI .... nondestructive inspection
NDT .... nondestructive test
NiCad .... nickel-cadmium

OJT .... On-the-Job Training
OSHA .... Occupational Safety & Health Administration

PCMCIA .... Personal Computer Memory Card International Association
PENS .... Pens ENhancement System
POD .... probability of detection

R.I.I. .... Required Inspection Item
ROC .... receiver operating characteristic
ROM .... read only memory

SATO .... speed/accuracy tradeoff
SDT .... signal detection theory
SUNY .... State University of New York
SwRI .... Southwest Research Institute

TLX .... Task Load Index

USAF .... United States Air Force

WAIS .... Weschler Adult Intelligence Scale

Chapter One
Phase III Overview

1.1 SUMMARY

This is Volume I (of II) of the Phase Three report of the Office of Aviation Medicine research program on Human Factors in Aviation Maintenance and Inspection. The research program has matured since it began in 1989. **Figure 1.1** shows that the research program has fully transitioned to the final stages of Implementation/Evaluation. The government and the aviation industry have begun to embrace and adopt the products of the research program. These products and research results are described herein.

The success of this research and development program is founded on the principle that "good science" must be the basis for "good practice." However, basic scientific research must not be confined to the laboratory - end users must be involved in all stages of the research. This research attributes its success to the active participation of end users. Those participants include the FAA Flight Standards Service; industry consortiums like the U.S. and International Air Transport Association and the Aviation Technician Education Council; individual airlines like Delta, Continental, USAir, Northwest, United, and others; and labor representatives like the International Association of Machinists. Under such guidance the various members of the research team have been able to develop, implement, and evaluate human-centered maintenance performance enhancements.

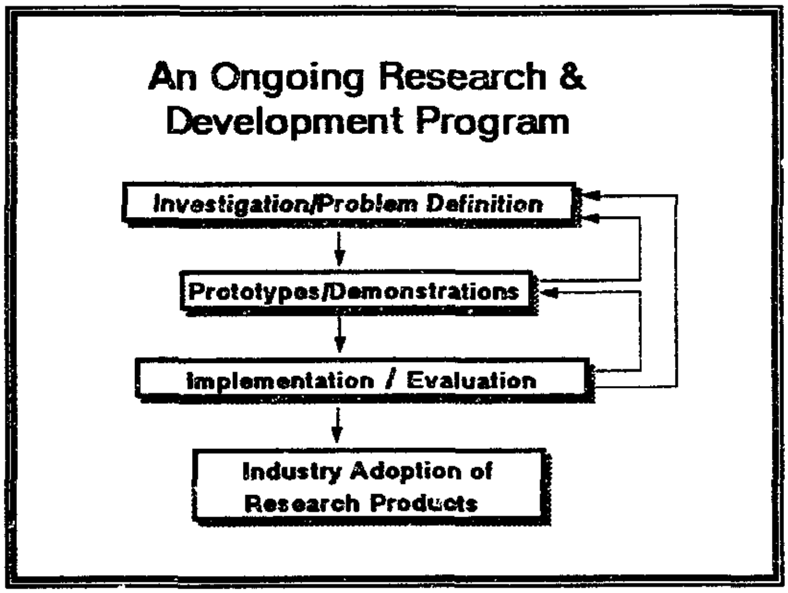


Figure 1.1 The Research Program

Each chapter contained in this volume addresses an aspect of performance enhancement. The research program recognizes that outputs must have a focus on safety and on cost control. Our current air transportation system is safe and all trends show increasing reliability and safety. The safety must continue in concert with cost control. Cost control means working smarter, reducing errors, reducing flight delays or cancellations, and generally improving the overall efficiency and effectiveness of the human in the total aviation maintenance system.

1.2 AN AIRLINE EVALUATION OF ADVANCED TECHNOLOGY TRAINING (Chapter Two)

During Phase I and II of the Aviation Medicine Human Factors in Aviation Maintenance and Inspection research program, training was a key research topic. A key product of the initial phases was advanced technology computer-based training for technicians. The focus of training is the Boeing 767-300 environmental control system (ECS). The training system includes an operational simulation of the ECS and a robust method of tracking student performance and providing advice and feedback.

This third phase of the research had the goal of evaluating the instructional effectiveness in an airline training setting. A formal evaluation was conducted comparing individualized student computer-based training (CBT) versus instructor-lead CBT. Results of the experiment are reported in this section. The results should be helpful to training personnel as they make decisions regarding the best application of instructional technology.

1.3 IMPROVING PERFORMANCE WITH BETTER INFORMATION CONTROL (Chapter Three)

The Performance ENhancement System (PENS) is being designed for the 2600 Aviation Safety Inspectors (ASIs) of the FAA Flight Standard Service. PENS capitalizes on advanced technology software and hardware to improve the collection, storage, analysis, and distribution of field data. Chapter Three describes the requirements analysis, early design and evaluation of a variety of hand held portable computer systems.

1.4 PERFORMANCE IN INSPECTION TASKS (Chapter Four)

Aviation maintenance requires the highest quality assurance. Thus, continuing inspection is a critical component of quality. This chapter describes an example of the laboratory research underway to identify characteristics of personnel best suited to inspection-oriented jobs.

The chapter reviews over twenty-five years of Department of Defense literature related to personnel and inspection.

1.5 ERGONOMIC FACTORS AFFECTING POSTURE AND FATIGUE (Chapter Five)

Aviation maintenance tasks require the technician to bend over, squat, and perform a variety of anatomical contortions. Such forced changes in posture lead to fatigue, back and limb soreness, and perhaps, error. This chapter, first, reviews the research literature related to such topics as the following: restrictive space factors, stress, and fatigue. Second, the chapter presents a plan to identify aviation maintenance tasks that are likely to force unnatural posture and, thus, increase the likelihood of fatigue and resultant error. Ultimately this research will prescribe a program to identify maintenance tasks where posture demands are beyond that at which a human can perform safe and reliable work performance. The research will offer ways to improve human performance under such conditions.

1.6 AN EVALUATION OF THE AIRCRAFT MAINTENANCE VISUAL ENVIRONMENT (Chapter Six)

Visual inspection accounts for 90% of all inspection activity. Therefore there is a high value in research that will improve the inspection visual environment. This study used a visual environment evaluation at the facilities of an airline partner to develop a general methodology for recommending the correct equipment. Ambient illumination must be supplemented by both portable area lighting and personal light sources to achieve the necessary illumination levels. The importance of a glare-free visual environment that makes use of surface reflectance is stressed. The developed methodology used task analysis data, lighting evaluations, input from inspectors and the evaluation of light sources to specify better equipment and visual surroundings.

1.7 A REDESIGN OF MAINTENANCE WORK CONTROL CARDS (Chapter Seven)

Aircraft maintenance and inspection is often driven by workcards. They present a detailed and organized ordering of the subtasks necessary to complete a job. This chapter describes an effort to improve the method in which workcards are designed and presented to the aviation maintenance technician.

As part of this project new workcards were designed for A-check and C-check on DC-9 aircraft. The results of an airline evaluation are reported. Not only does this chapter propose specific design solutions, but it also provides a highly generic methodology for design of quality technical documentation, both written and digital.

1.8 TRAINING FOR VISUAL INSPECTION (Chapter Eight)

During previous phases of the research a computer-based simulation was built for laboratory research on training for visual inspection. This chapter summarizes the results of laboratory experiments and offers concrete examples of the necessary components of a training program for visual inspection.

This chapter reports on the status of visual inspection in airlines, aircraft manufacturers, and other non-aviation maintenance environments. The chapter describes training alternatives such as part-task, whole-task, adaptive, active, on-the-job, and computer-based training for visual inspection. The chapter also describes training and inspection feedback that is likely to improve technician performance in visual inspection.

1.9 CONTINUING RESEARCH

Phase III, Volume II will be published about four months after this volume. The next volume will place increasing importance on the measurable impact of the research on human performance enhancement. The aviation industry is struggling through increasingly difficult financial hardships. Research programs must continue to improve the "bottom line" by providing procedures and products that improve maintenance efficiency. That will remain a highest priority of this program.

Chapter Two
Results of the Environmental Control System Tutor Experiment

2.0 INTRODUCTION

This study investigates the effect of presentation methods on computer-based training effectiveness. The experiment was conducted at the technical operations training center of a major airline in the Fall of 1992. Subjects used the Environmental Control System Tutor with both instructor-led and individual-use teaching methods. The experiment found no significant difference in overall performance between the two groups, although the instructor-led group did perform slightly better on the part identification section of the examination. Also, the experiment found no significant difference in the preference of the presentation method between the two groups. This report also covers shortcomings in the design of the experiment.

2.1 PURPOSES OF THE EXPERIMENT

The study had two motivations: verification of the effectiveness of the tutor and comparison of computer-based training (CBT) methods. First, we wanted to ensure that the use of the Environmental Control System (ECS) Tutor will improve the students' performance in diagnostic tasks. We have already conducted several informal usability studies that looked at the compatibility and understandability of the tutor as described in Pearce (1992), and felt it necessary to perform a formal effectiveness study of the tutor for final evaluation. By comparing the performance of technicians who have used the tutor with the performance of subjects who have been taught with traditional methods, we were able to get a better idea of the strengths and weaknesses of the ECS Tutor. In addition to testing, we also collected data on the opinions of the subject concerning the tutor, to see if there are problems with the design or implementation of the tutor.

Second, we wanted to compare the effectiveness of presentation methods for CBT systems. The two top-level classifications for presenting CBT systems are the instructor-led and individual-use methods. The instructor-led method is the traditional mode of teaching, in which the teacher controls the presentation of material. In the individual-use method, each student controls his or her own learning process. Several studies have compared the efficiency of these two methods for general instruction (Charney and Reder, 1986; Czaja et al., 1986), but no studies could be found that have compared these methods for teaching troubleshooting. The information obtained from these results will help to determine specific components of CBT systems that improve student performance. The data from this study will also be useful in the evaluation of the cost effectiveness of computer-based training systems.

2.2 ENVIRONMENTAL CONTROL SYSTEM OVERVIEW

The ECS, found in all modern airliners, controls the pressure and temperature of air in the airplane. The ECS of the airliner that the tutor simulates consists of three control and display panels in the cockpit, several electronics modules in the avionics bay, the distribution system, and the two cooling packs located in the fuselage. The ECS is a very complex system and consists of electrical, mechanical, and air flow subsystems that interact to provide the cool, pressurized air. It was chosen for the training domain of the tutor because the ECS is fairly common across airliner types, and therefore the training could be generalized across airliner types. Built-in test equipment (BITE) of modern airliners makes the job of the technician easier, since it tests some components with the push of a button. Not all components are tested by the BITE, so the technician must know when and how to use external test equipment to isolate malfunctions.

2.2.1 The Flightline Technician

The flightline technician must quickly diagnose and repair malfunctions on the aircraft on which they are certified to work. Technicians must know about the systems of several different types and models of aircraft. Their task is time constrained, since most flights have about 40 minutes on the ground between landing and takeoff. Also, some repairs take more than 40 minutes, and the technician must find these faults quickly to minimize delays in the flight schedules.

It is standard procedure for the flightline mechanic to use the Fault Isolation Manual (FIM), which is a logic tree used to diagnose malfunctions. The technician follows the branches of the FIM based on the outcomes of tests and inspections. The FIM specifies a "minimal path" of actions to repair a failure, from the high-level description of the malfunction to the malfunctioning component. In some cases, it is possible to diagnose malfunctions with a single test (for example, by looking for abnormal temperatures in the airflow path), so in practice the FIM is not always used.

2.2.2 Overview of ECS Tutor

The ECS Tutor is a computer-based training (CBT) system that trains aircraft technicians to diagnose and repair malfunctions of the ECS of the Boeing 767 (**Figure 2.1**). The tutor contains a deep simulation model of the ECS, which allows the user to see the consequences of his actions on the system down to the sensor level. The user can change switch settings to observe the values of various system parameters. The tutor is also highly graphical, allowing for direct manipulation of ECS components, and contains realistic pictures and animations of system components and schematics.

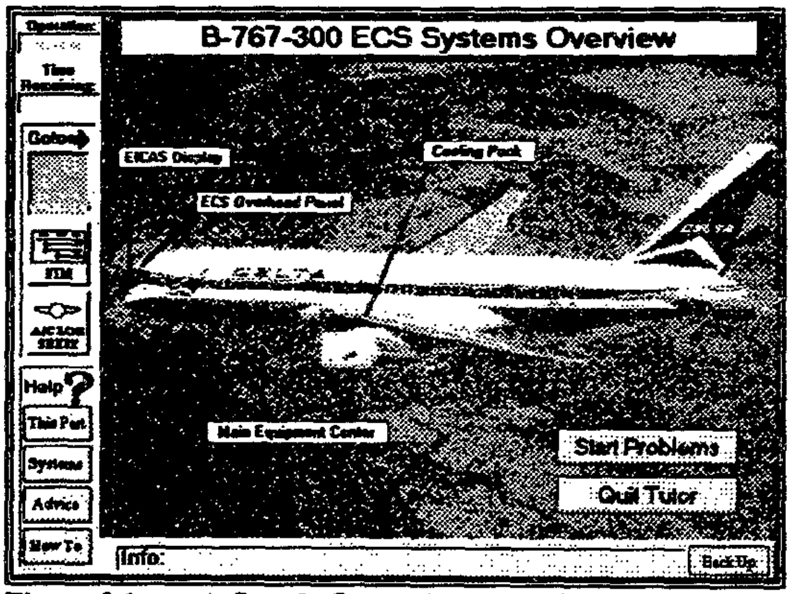


Figure 2.1 A Sample Screen from the ECS Tutor

The tutor allows four types of actions on the components of the ECS: operate, inspect, test, and replace. The first is to *operate* the ECS equipment. For example, the student can change switch settings of the cockpit control panels. The second action is to *inspect* a component; this action includes reading of display values on control equipment or looking for visible failures in pack components. The next action, *test*, differs from inspection in that the technician has to perform some type of action, usually operating some internal or external test equipment, rather than just observing a component. One example of this activity is when the technician tests the pack controller by operating the BITE. The last action, *replace*, allows the user to swap out line replaceable units (LRU's) with working components.

2.3 METHOD

The experiment was designed to measure differences in performance between students taught to troubleshoot using a "traditional" instructor-led training method and an individual-use training method. Because the participating airline does not give a formal course that explicitly teaches troubleshooting skills, we had to design a short instructor-led CBT session based on traditional teaching methods. To standardize the information that was being presented in both groups, the instructor for the instructor-led group presented the same version of the tutor that was used by the subjects in the individual-use group. The only difference between the two groups was in the method of presentation; the individual-use group interacted directly with the tutor, while the instructor-led group had this information presented by an instructor. Thus any differences in performance could be attributed to the method of presentation, rather than any differences in content.

2.3.1 Subjects

The subjects participating in the experiment consisted of 10 ground training instructors and 10 flightline technicians. All of the subjects had some level of general knowledge of the operation of ECS, from either troubleshooting experience or courses on the ECS's of other aircraft. None of the subjects had worked on the Boeing 767 ECS. Instructor experience ranged from two to 19 years as instructors, while all of the flightline technicians had less than two years experience as aircraft technicians. Subjects were randomly assigned to one of the experimental groups, with half the instructors and half the technicians going to each experimental group.

2.3.2 Procedure

The experiment was divided into three phases: introductory lesson, tutor usage, and testing (shown in **Figure 2.2**). All of the subjects participated in an introductory lesson on the basic operating principles of the B-767 ECS. This course, developed by an instructor of the technical operations training department of the participating airline, covered the general operation of the B-767 ECS, modes of operation of the ECS, and the functions of the sensors, valves, and electronics that control ECS operation. All subjects went through this one-hour course before participating in the tutor usage portion of the experiments.

After this one-hour course came the tutor usage phase, in which the students were split into the two groups ("instructor-led" and "individual-use") for the 2 1/2 hours troubleshooting training course. Each member of the individual-use group used the ECS Tutor individually to solve as many problems in the tutor as possible in the allotted time. The students used the tutor on the participating airline's training computers. The instructor-led group was given a stand-up lecture on ECS troubleshooting, with the instructor using the ECS Tutor as an instructional aid. This tutor usage phase was stopped for both groups when the instructor finished all ten problems in the tutor. Thus the time of instruction was the same for both groups, except for the three subjects in the individual-use group who finished early.

2.3.3 Data

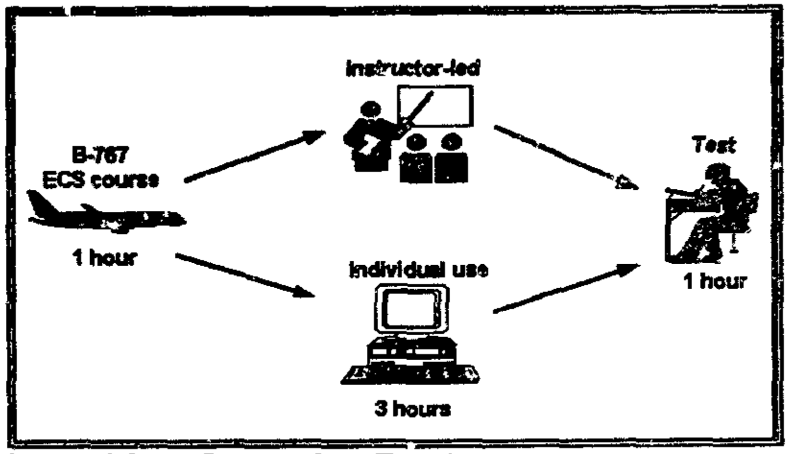


Figure 2.2 Design of the Experiment

After the instructor-led group had finished all 10 troubleshooting problems in the tutor, both groups were given a short examination that measured troubleshooting skills. This one-hour examination, developed by an instructor at the participating airline, was designed to measure a variety of skills. Most of the questions were multiple choice, with some fill-in-the-blank questions. Questions were divided into four sections, and data was collected on completion times for each of these sections (the questions are described in the "Results" section). During the examination the subjects could use a diagram of the ECS and the fault isolation manual to help them solve all the problems.

The exam also contained a poll with questions about the user's satisfaction level with the tutor. We also administered a background poll to determine the distribution of skill levels for computer use and ECS maintenance. After the subjects finished the examination and polls, they were asked to write about any impressions or observations concerning the tutor.

2.4 RESULTS

This section is divided into an *Examination results* section, covering the analysis of the data from the tutor examination, the *Examination comments* section, which describes the results of the poll, and the *Written comments* section, covering the written comments concerning the tutor.

2.4.1 Examination Results

The examination contained 23 questions, divided into four question types: components, procedures, systems, and troubleshooting. Component questions measured knowledge of the parts of the ECS. The procedure questions measured knowledge of the procedures necessary to diagnose the ECS. Systems questions addressed the various control systems and their relationships. The largest part of the examination was the malfunction section, which tested knowledge of troubleshooting performance on the B-767 ECS. Scores on the questions were weighted by difficulty; for example the troubleshooting questions counted about twice as much as the component questions, because they were more difficult and time consuming.

The only significant difference between the group's scores was in the component section. These questions dealt with the ECS on a component level; for example these questions concerned the function of the parts, connections to other parts, or behavior of a specific part. There were two main reasons for the superior performance of the instructor-led group in the component-related questions. First, since the ECS Tutor does not explicitly teach the user about the ECS at the component level, the individual-use subjects were disadvantaged when it came to learning about the parts of the system. Second, students in the instructor-led group could ask the instructor to explain the finer points of the operation of the ECS, and the instructor would often answer questions by describing the behavior of one of the components. On the other hand, students in the individual-use group could ask the tutor about a component by clicking on one of the "Help" buttons, but dialog with a computer is not always as robust or meaningful as that with an instructor. This result, along with several of the written comments, points out the importance of giving adequate background information before attempting to teach troubleshooting. This background information could be taught with the computer, although in many cases it may be more effective and efficient for an instructor to teach this material.

The examination data showed no significant difference between the two groups in the time to complete the examination or in the weighted overall examination scores, as shown in **Table 2.1**. This is shown graphically in **Figure 2.3**, which is a "box and whiskers" plot of the median (line in the box), the first standard deviation (the box), and the second standard deviation (ends of the whiskers) for the overall scores of the two groups. Also, there was no significant difference in performance on the procedures, systems, and troubleshooting sections of the examination. There was also no significant difference in the average time that it took to complete the examination. **Figure 2.5** in the Appendix A shows the score distributions of the two groups for the four examination sections.

Table 2.1 Summary of Test Statistics

	Instructor-led	Individual-use
Time to complete test	31 min.	32 min.
Average component (A) score	80.0%	46.7%
Average procedure (B) score	73.3%	63.3%
Average system (C) score	63.4%	54.8%
Average troubleshooting (D) score	71.3%	76.6%
Average overall score	70.3%	65.1%

2.4.2 Examination Results

The examination contained 17 questions about various aspects of the tutor. Questions were of two types: general questions dealing with the usability and general behavior of the tutor, and questions about several features of the tutor. Subjects were asked to rate their agreement with each statement, using the scale "agree strongly," "agree," "no opinion," "disagree," and "disagree strongly." The questions were equally mixed between positively and negatively phrased sentences. **Figure 2.6** in the Appendix A shows the distribution of responses for the subjects in the individual-use group.

Overall, satisfaction with the tutor was high. Some subjects in the individual-use group thought that the tutor behaved in unexpected ways. Ongoing usability studies are being used to locate the source of this problem. Also, fine-tuning of the context-sensitive help will improve the system's ability to offer guidance to the user at the appropriate time. There were no strong negative responses to the overall design of the tutor or to specific features, although the responses to several questions indicated that the wording of error messages could be improved.

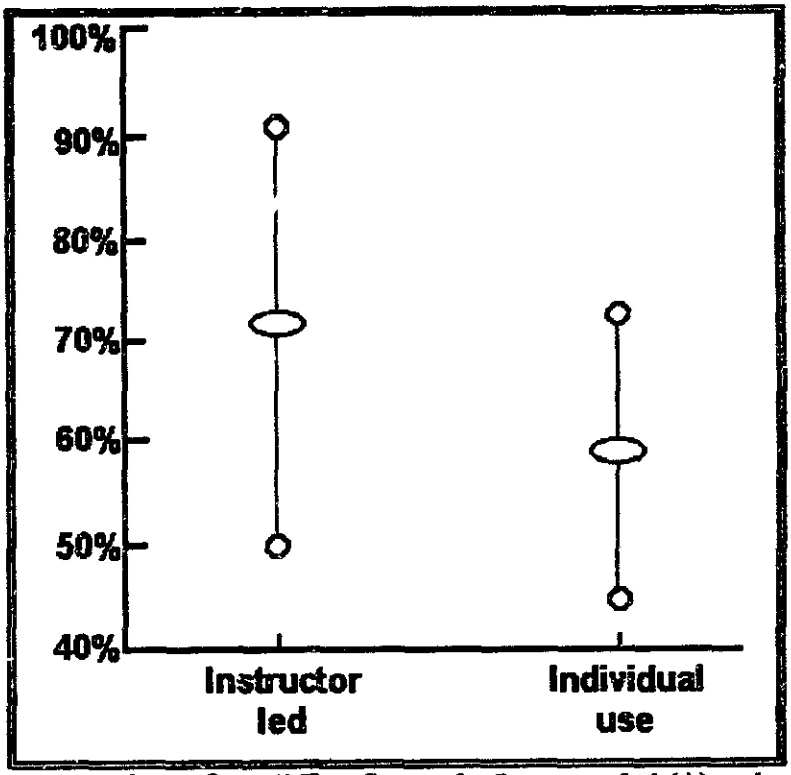


Figure 2.3 Overall Test Scores for Instructor-led (A) and Individual-use (B) Groups

A comparison between the two groups showed no significant difference in satisfaction with the tutor. This comparison was done by removing all examination questions that were not relevant to satisfaction or were not relevant to both user groups. The negatively-phrased questions were then inverted, and then the responses for each question were sorted by value, with "one" being the highest level of agreement and satisfaction, and "five" being the lowest. The distributions for the two groups are shown in **Figure 2.4.** The two distributions are almost identical, indicating that preference for the tutor was independent of the method of presentation. This seems to indicate that the positive features of instructor-led instruction (for example, natural instructor-student dialog) were balanced by the positive features of individual-use instruction (for example, full control over instructional rate).

2.4.3 Written Comments

The examination asked the subjects to write down any specific problems that they had with the tutor, and about their general opinion of the tutor. Only four subjects (of twenty total) responded to this section, all of whom were in the individual-use group. **Table 2.2** shows all of the written comments from the examination.

2.4.3.1 Hardware and Software Bugs

The written responses indicated several different problems that subjects had with the tutor and the training program. On the first level were problems with the computer hardware. Since there was not enough time to allow the students to calibrate their own touchscreens, the touchscreens were calibrated by one of the evaluators. But individual differences in height, handedness, and hand-eye coordination were significant enough to cause problems for some subjects. Since not all user's actions are confirmed, some errors in screen touching were attributed to logic errors by the tutor. There was no way to correct these problems during the experiment, and several users became irritated at being "falsely accused" of making mistakes. Similarly, the tutor would unpredictably crash on two of the training computers, and the subjects would have to start again from the beginning. Thus some individual-use subjects paid more attention to the mechanics of using the tutor, rather than focusing on the content of the tutor. This software problem was found after the experiment and has been fixed.

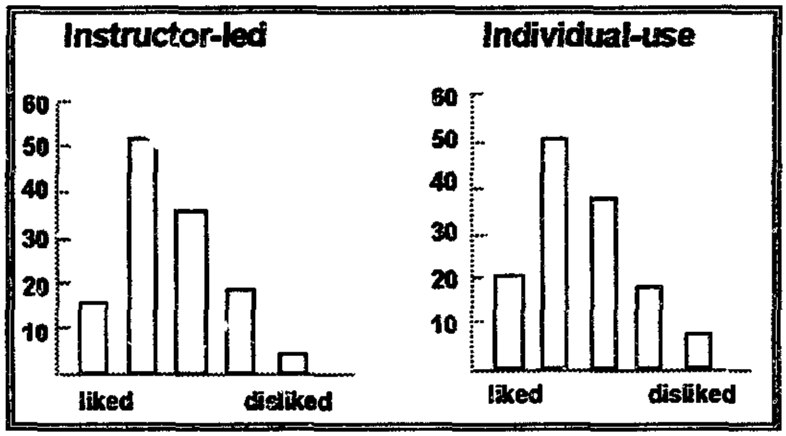


Figure 2.4 Tutor Preferences of the Two Groups, Measured by Poll Responses

Table 2.2 Written Responses to Poll

Subject B-49: This training without an instructor is close to useless; it should only be used to enhance the class.
Subject B-50: ECS Tutor is very effective, but takes time to get relaxed with. For someone with a lot of computer background the ECS Tutor would help a great deal.
Subject B-55: 1) Arrow was not always accurate; touch screen was not aligned correctly 2) No chance to correct yourself if computer made wrong selection, because of poor alignment! 3) Twice system locked up - may be Delta's computer not program - and made student go to beginning of lesson and the time and mistakes shown at end of lesson were unrealistic Was very interesting thanks.
Subject B-56: I found it frustrating when I tried to get back to a problem to see how it was written. The backup selection would only take me one screen back but not any further. I did not like the CBT simply because it did not answer my questions, only gave me information that the program thought I needed.

2.4.3.2 Response Time

Another problem some subjects had with the tutor was the long lag between the time that the students performed an action and the time that the tutor responded to that action. The computers used for development were two to three times as fast as those used for the experiment, and the lag was not significant on the development computers. The lag was much more noticeable on the computers used in the experiment. The subjects would sometimes repeat an action several times if there was not feedback showing that the computer was processing their input. As a result, user actions were applied to subsequent displays, which led to unwanted behaviors in the tutor.

2.4.3.3 Navigation

Some subjects were also uncomfortable with the behavior of navigation between screens. For example, the version of the tutor that was used in the experiment had a "back" button that allowed the user to toggle back and forth between the previous two screens. In designing the tutor we thought that this would be as much as the user would need, and went with this simple design. But the users expected that the button would cause the tutor to keep going back to previous screens until they arrived at the starting screen. There were several times when it would be faster to have the button work in the expected way, and some users became frustrated when it did not work the way they wanted. Earlier testing did not indicate that this was a problem; only in a classroom setting did it tend to confuse the users.

2.4.3.4 Time Constraints

Probably the largest source of frustration was the limited information that the subjects had about the ECS. In the standard two-week training course that technicians must take before being certified on a plane type, about five hours are spent on the ECS. But because of time limits, only one hour was spent on this background material. Thus there were many uncertainties about the ECS, which led to frustration when trying to use the tutor. In retrospect, more time should have been spent on the introductory part of the experiment.

These results show that it is important that the subjects have all the necessary background information. Students that lack the prerequisite knowledge will have trouble in learning complex reasoning tasks such as troubleshooting. Also, having robust, usable tutoring software that has been thoroughly tested on the target platform is important. Software that is frustrating to use will not have the full impact of usable software.

2.4.4 Caveats

Although we did our best to ensure that the ECS Tutor experiment was a valid experiment, there were several problems with the design of the study. These problems are quite common in the design and execution of training evaluations (Goldstein 1987). The problems that we experienced were not a result of poor planning, but had more to do with constrained resources. These problems need to be considered when drawing conclusions from this study. The problems have to do with sample size, experience level of the subjects, and the testing of troubleshooting skills, and are described below.

2.4.4.1 Subject Group Sample Size

It is difficult to draw hard conclusions from subject groups that are smaller than 40 persons. Individual differences play an important role when population sizes are small, as they were in this experiment. Most of the results were not statistically significant, but this could be from individual differences between the two groups, and not because of inherent differences in the efficiency of the teaching methods. Future studies should use larger subject groups.

2.4.4.2 Experience Level of Subjects

Another problem had to do with the experience levels of the subjects. Of the twenty subjects, ten were experienced instructors with between two and 19 years experience, and the other ter participants were flightline technicians with less than two years experience. This should be compared with the two goals of the ECS Tutor: to teach general troubleshooting skills, and to teach specific knowledge to troubleshoot the B-767 ECS. The first goal, teaching general skills, is the most important of the two, as general skills can transfer across instances of airliner troubleshooting, and across airliner types. Clearly, the participants in the experiment already had strong troubleshooting skills, although they did not have specific knowledge of the B-767 ECS. Because of the mismatch between the intended users and the actual users of the tutor in this study, it would be very unlikely for any teaching method to cause a measurable difference in performance, since most of the subjects were already fairly knowledgeable about troubleshooting.

2.4.4.3 Testing Troubleshooting Skills

The last problem has to do with the difficulty of testing troubleshooting skills. Because troubleshooting requires a combination of several types of skills and knowledge, tests that attempt to measure troubleshooting skills should also measure a variety of skills. It is also difficult to design an examination to measure this knowledge "out of context," because performance on the job may be very different from performance in the classroom. In an ideal experiment, the subjects would have been tested on real equipment and evaluated by a panel of experts, but this was not possible. Future examinations should attempt a more realistic task for measuring troubleshooting skills, even if the testing must be done on paper.

2.5 CONCLUSION

The experiment found no significant difference between the troubleshooting performance of subjects that used the tutor in instructor-led and individual-use modes. An analysis of the performance on the four question types (components, procedures, systems, and malfunctions) also found no statistically significant difference. An analysis of user satisfaction with the tutor found no difference between the two groups. Possible problems with the significance of these results are small population size, experience levels of the subjects, and difficulty of testing troubleshooting skills.

This experiment also led to several important points to be considered in designing CBT courses and CBT evaluations. First, the subjects should have all the necessary background information, since students who lack the prerequisite knowledge will have trouble in learning complex reasoning tasks. Also, having robust, usable tutoring software that has been thoroughly tested on the target platform is important. Difficult software that is frustrating to use will not have the full impact of usable software. Third, because individual differences play an important role when population sizes are small, studies should use larger subject groups with at least 40 subjects.

2.6 REFERENCES

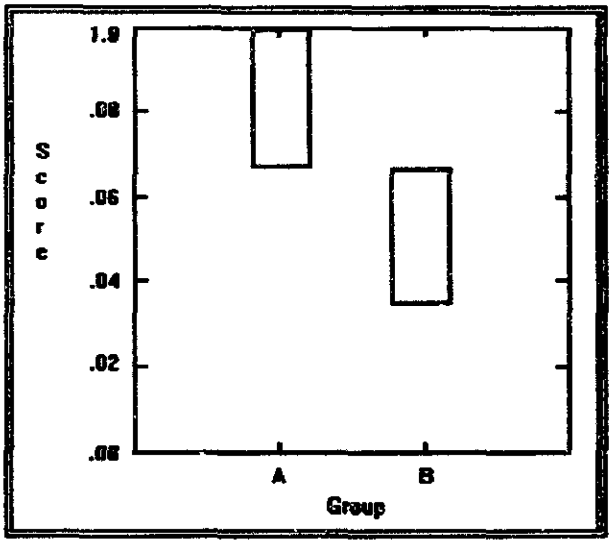
Charney, D. H., and Reder, L. M. (1986). Designing interactive tutorials for computer users. *Human-Computer Interaction,* 2, (4), 297-317.

Czaja, S. J., Hammond, K., Blascovich, J. J., and Swede, H. (1986). Learning to use a word-processing system as a function of training strategy. *Behavior and Information Technology, 5,* (3), 203-216.

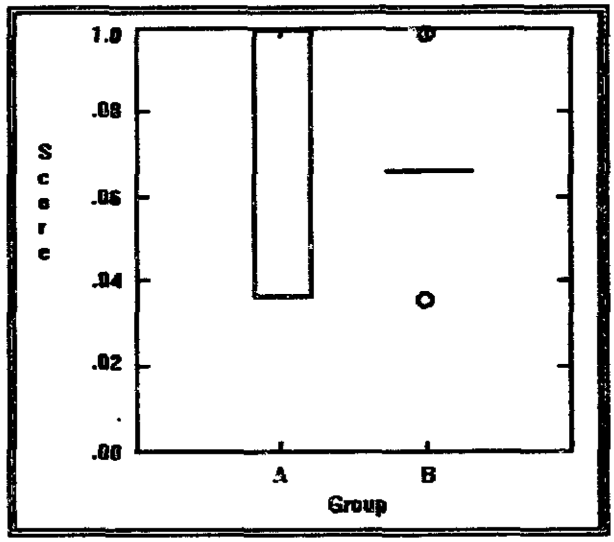
Goldstein, I. L. (1987). The relationship of training goals and training systems. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 963-975). New York: John Wiley & Sons.

Pearce, M. C. (1992). Report on the environmental control system tutor usability study at clayton state college. Unpublished manuscript. Galaxy Scientific Corporation.

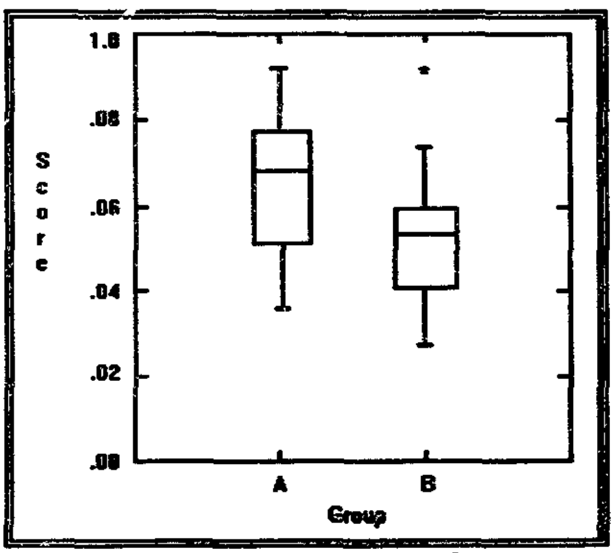
Chapter Two
Appendix A



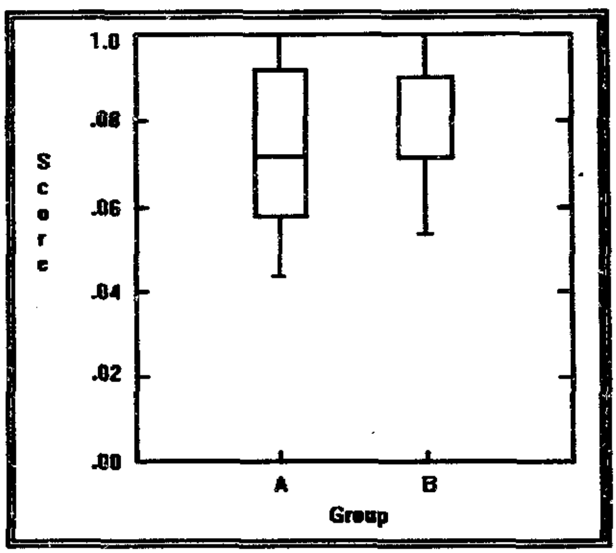
Scores for Test Section A



Scores for Test Section B



Scores for Test Section C



Scores for Test Section D

Figure 2.5 Distribution of scores for examination sections A, B, C, and D. These "box and whiskers" show the median (line in the box), the first standard deviation (the box), and the second standard deviation (ends of the whiskers) for the section scores of the two groups. Group A was instructor-led and Group B was individual-use.

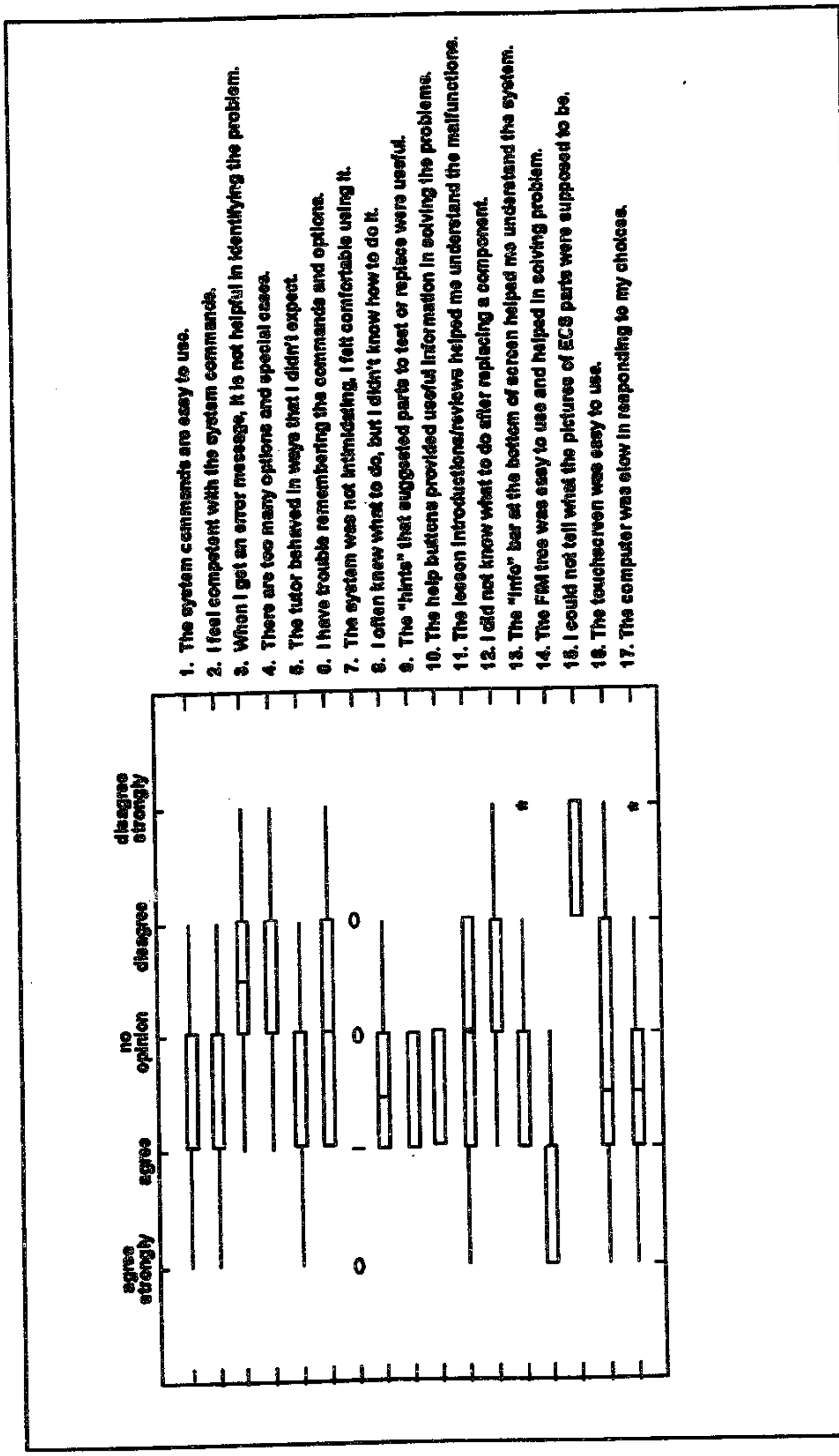


Figure 2.6 Examination Response Distribution for the Individual-use Group

Chapter Two
Appendix B

Evaluation # ________

ECS Tutor Evaluation

This evaluation will measure your knowledge of the B-767 ECS and of troubleshooting procedures. The evaluation is to be given after you have seen the ECS Tutor, and is divided into background and troubleshooting questions. The background questions will measure your knowledge of the various component, systems, and procedures related to the ECS. The troubleshooting questions measure your ability to diagnose malfunctions of the ECS.

Most of the questions are multiple choice; simply circle the letter in front of that answer you think is correct. If you need more room for the "short answer" questions, continue on the back of that page, but the answers should not be very long.

You may use the materials from the introductory section of the evaluation (the pack diagram, equipment drawings, and FIM), including any notes you took.

For each of the sections, you should write the times that you begin and end that section; an area is provided. You should not rush to finish the evaluation; correctness is more important than time.

The last two pages of this package are forms that ask your opinion of the tutor and about your experience with ECSs. Please take the time to fill these out, as they will influence the changes that are made to the tutor.

NOTE: This is not a test! This results of this evaluation are confidential and will not be used to evaluate you. You should not sign your name to this evaluation. The study is designed to evaluate the ECS Tutor, and not the people taking this evaluation.

A. Components

These questions will evaluate your knowledge of the parts of the B-767 ECS. These questions are all multiple choice; just circle the correct answer.

Start Time: ______

1. The function of the Primary Heat Exchanger is to:
   A. limit flow of bleed air into the pack.
   B. cool the air before it enters the compressor.
   C. cool bleed air coming out of the compressor before it enters the condenser.
   D. none of the above.

2. The low limit valve:
   A. receives differential pressure to command the valve.
   B. can receive an electrical command from the standby pack controller to command the valve.
   C. can receive an electrical command from the pack temperature controller to command the valve.
   D. all of the above.

3. When performing a pack temperature controller BITE, the pack sensor is faulted. Where is the sensor located?
   A. In the ceiling of the cabin.
   B. Between the primary water extractor and the turbine.
   C. Near the flow control valve.
   D. Downstream of the turbine.

Finish Time: ______

B. Procedures

The questions in this section will measure your knowledge of the procedures necessary to diagnose the B-767 ECS. You will be given a task, and must describe the steps to perform the task. For the "short answer" questions, please include an explanation.

Start Time: ______

4. On the pack temperature controller, if you press the PRESS TEST switch and the GO light does not illuminate, you should:
   A. replace the components with the lights illuminated
   B. press the VERIFY switch to reset the controller.
   C. replace the controller.
   D. replace the faulty lamp.

5. A B-767 has a problem with the ECS system. Using the FIM and fault code 21-51-19, go to the appropriate chart to find the fault. Starting at block #1 answer YES. The next block would be answered NO. What is the problem?

6. A B-767 has a problem with the ECS system. Using the FIM, find the correct fault code. On the overhead panel with the pack selector switch in "AUTO," the "INOP" light on the reset switch is illuminated. When STBY was selected the light extinguished. What is the fault code for this problem?

Finish Time: ______

C. Systems

This section will measure your knowledge of the various systems and the relationships between components of the ECS. The goal is to measure what you know of the behavior and functioning of the pack.

Start Time: ______

7. What controls the pack when the selector switch is in auto?

____

8. What controls the pack when the selector switch is in STBY N?

______________________ and ______________________

9. Circle the correct answers: In STBY C the Low Limit Valve is normally (closed, open), and the Temperature Control Valve is normally (closed, open).

10. If the INOP light and the PACK OFF lights are illuminated on the reset switch in the AUTO position, the problem is with the
   A. Trim Valve failed.
   B. Pack Outlet overtemp.
   C. Gasper Fan failed.
   D. Compressor Outlet overtemp.

11. What is the purpose of the altitude switch?
   A. to increase airflow through the pack at 31,000 feet, since the air is thinner at higher altitude.
   B. to allow colder air to the turbine at 31,000 feet
   C. so the EICAS can record the altitude on the autoevent if the pack fails.
   D. all of the above.

12. What indication would you get in the cockpit if the altitude switch failed?
   A. Failure flag on the captains altimeter.
   B. Both packs trip and will not reset.
   C. No indication in the cockpit.
   D. EICAS maintenance message "PACK CONTROLLER BITE."

Finish Time: ______

D. Malfunctions

This is the most important section of the evaluation, since it measures what the ECS Tutor was designed to teach to its users. Include a short explanation for your answers.

Start Time: _______

13. The left pack fails. In the cockpit the left PACK OFF light on the RESET switch is illuminated and will not reset when the pack selector switch is in any position. There is also an EICAS message L PACK OFF. What would you suspect the problem to be?

14. The pilot tells you the INOP light on the pack reset switch for the right pack illuminated and will not extinguish. You go up to the cockpit and move the selector switch to the STBY positions and try to reset the switch. The INOP light stays illuminated. There is also an EICAS message L PACK TEMP. You go down to the pack bay and physically check the heat exchangers. Heat exchangers check OK. You also do a BITE check on the Pack Temperature Controller; the BITE test is OK. You check the Flow Control Valve. It checks out OK. You do a BITE check on the Standby pack controller; checks OK. Using the FIM manual, what is the problem?

15. When performing a BITE test on the standby pack controller, you get a NO GO light in position 3. What component is faulted?

16. When performing a pack temperature controller BITE, what position should the pack selector switch on the overhead be set to?

17. A B-767 is incoming with a pack malfunction. The indications were not radioed in. Once the plane lands, in what order would you check these items?

__ inspect heat exchangers.
__ operate pack temperature controller BITE.
__ check condition of the PACK OFF/INOP lights.
__ try to operate the pack in different modes.

18. A B-767 has a problem with the ECS system. On the ground with the packs operating, the left pack trips off. You have an "INOP" light and "PACK OFF" light illuminated. You allow the pack to cool and reset the switch, the lights extinguish. After a while the pack trips again in AUTO and STBY. You go down to the left pack bay and notice water dripping from a drain tube on the water extractor. This would indicate:
   A. normal condition.
   B. coalescent bags need to be replaced.
   C. water nozzles in the ram air duct are clogged.
   D. leak in the potable water system.

Finish Time: ______

You have completed the evaluation; please fill out the following evaluation.

Evaluation for ECS Tutor

This is an evaluation to determine how effective you think the ECS Tutor is. Please choose a number **between 1 and 5** that describes your agreement with each statement, using the definition in the scale below. Be sure to read the statements carefully. Write your choice to the left of the question.
NOTE: This is not a test! This study is confidential and will not be used to evaluate you. You should not sign your name to this evaluation. The evaluation is designed to determine which parts of the ECS Tutor need improvement.

1	2	3	4	5
strongly agree	**agree**	**neutral**	**disagree**	**strongly disagree**

General System Questions

__ 1. The system commands are easy to use.
__ 2. I feel competent with and knowledgeable about the system commands.
__ 3. When I get an error message, I find that is not helpful in identifying the source of the problem.
__ 4. There are too many options and special cases.
__ 5. The tutor behaved in ways that I didn't expect.
__ 6. I have trouble remembering the commands and options and must ask questions frequently.
__ 7. The system was not intimidating, I felt comfortable using it.
__ 8. I often knew what to do, but I didn't know how to do it.

Questions about Specific Components of the ECS Tutor

__ 9. The "hints" that suggested possible parts to test or replace were useful.
__ 10. The help buttons provided useful information in solving the problems.
__ 11. The lesson introductions and reviews helped me to understand how the malfunctions were related.
__ 12. I did not know what to do after replacing a component.
__ 13. The "Info" bar at the bottom of screen helped me understand the system.
__ 14. The FIM tree was easy to use and helped in solving problem.
__ 15. I could not tell what the pictures of ECS parts were supposed to be.
__ 16. The touchscreen was easy to use.
__ 17. The computer was slow in responding to my choices.

If you have any other comments about the ECS Tutor or about your answers to these questions, please write them on the back of this paper. Thank You.

Chapter Three
PEN COMPUTERS: EVALUATIONS, RECOMMENDATIONS, AND THE PENS PROJECT

3.0 INTRODUCTION

Pen computer technology has the potential to revolutionize the computer industry. Pen computers are compact, easy to use, and designed for field use. These factors make pen computers ideal tools for field data collection and analysis, even for individuals who do not currently use computers. Galaxy Scientific is working with the Flight Standards Service and the Office of Aviation Medicine to develop a job aiding system that is based on this exciting new technology.

The following is a discussion of the general characteristics of pen computers, a comparison of pen computers available from a variety of manufacturers, and a description of the progress of the Performance ENhancement System (PENS) for Aviation Safety Inspectors.

3.1 GENERAL CHARACTERISTICS OF PEN COMPUTERS

Pen computers are similar to personal computers in that they consist of a display, a central processing unit (CPU), and an input device. Unlike personal computers, however, pen computers put the CPU and display in one small box. Instead of a keyboard and mouse, a pen computer uses a special pen stylus for input. The pen stylus not only functions as a pointing device, it also serves as the primary means for entering data. Figure 3.1 illustrates both a typical personal computer and a typical pen computer.

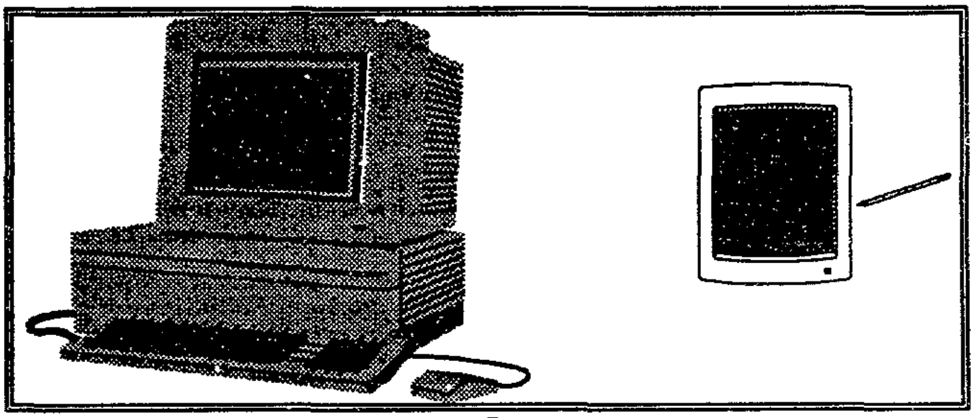


Figure 3.1. Typical Desktop and Pen Computers

Unlike a personal computer, data are written on the screen, rather than typed; a handwriting recognizer translates the printed input into "typed" characters. (Script or "cursive" recognition software is currently being developed by several companies.) Additional gestures are used for editing. Each person customizes the recognizer to her/his handwriting style for improved recognition accuracy. Pen computers also come with "virtual" keyboards (software versions of keyboards), with a connection for an actual keyboard, or with an actual keyboard located beneath the pop-up display.

Pen computers, like notebook computers or laptops, are battery powered. Extending the charge life of batteries is one of the hottest areas of portable computer research. Battery life currently ranges from one hour (with no energy conserving features turned on) to three or more hours (with all energy conserving features in use) on a single charge. While manufacturers take diverse approaches to battery charging and power management, nearly all pen computers come with user-replaceable batteries and ac adapters; several manufacturers also offer automobile cigarette lighter adapters.

Personal Computer Memory Card International Association (PCMCIA) slots also differentiate pen computers. Whereas notebook and laptop computers have seen limited use of these slots (and they're practically nonexistent on desktop computers), nearly every pen computer has at least one PCMCIA slot. These slots can be used for fax/modem cards, network cards, removable storage devices, and memory extension cards. The cards are approximately the size of a credit card and will likely see widespread use in the near future. Such devices allow quick and easy addition of peripherals or personal data. The PCMCIA slots make it easy to have a pool of computers for field workers. When someone needs to go to the field, she/he can grab a computer and a few PCMCIA cards and quickly have a customized machine.

Finally, pen computers are generally lightweight. Units range in weight from approximately three and a half pounds to around seven and a half pounds. This broad range is due to the unique features of each computer. For instance, the lightest computer uses a low voltage power system that reduces battery size (hence weight) and constrains the sizes (weights) of internal components. Two of the computers in the middle weight range come with built-in keyboards that can be stored beneath the displays. At the high end of the weight range are the ruggedized units; these units can withstand environmental extremes, such as cold temperatures and rain, and the general hazards of portable use, such as drops or collisions. The individualizing features of pen computers are discussed later in this document.

3.2 KEY BENEFIT OF PEN COMPUTERS

Pen computer technology capitalizes on the evolution of several branches of computer science and engineering. Graphical operating environments, such as Windows, allow the user to operate a computer almost entirely through pointing and "clicking" (tapping twice in rapid succession with the pointing device). The pen stylus not only supports such pointing and clicking, but when it is combined with handwriting recognition, it allows the user to enter data or issue commands. Thus, one simple device can be used as the sole means of computer operation and data collection. The result of such technological advances is that pen computers offer the promise of empowering field workers with computer technology. Even those people who don't traditionally use computers can be brought up to speed with relative ease.

3.3 POTENTIAL USES OF PEN COMPUTERS

Because pen computers are designed for field use, they have a variety of applications. Some of these application areas include sales, production, health care, census, law enforcement, delivery services, investigation, and inspection.

For example, sales people can make sales calls, assess the customer's needs, quote a price, and even sign up the customer, all on the computer. Production personnel can document production difficulties and track work in progress as they walk through the plant. Health care applications include patient forms, pharmaceutical orders, meal planning, and patient tracking and charting. Instead of using paper forms and waiting months for the data to be entered into a computer database, a census could be taken with on-line forms, thus facilitating quick compilation of a database.

Law enforcement personnel could use pen computers in a variety of ways, from mundane tasks such as writing tickets, to more involved tasks such as documenting and investigating crimes. Personnel in the National Transportation Safety Board could use them for aircraft accident investigations. Delivery services currently use custom pen computers for package tracking, delivery schedules, and recipient signatures. Any regulatory agency could use pen computers for inspections. For example, Food and Drug Administration personnel could use them when inspecting food production and sales facilities (e.g., meat packing plants, restaurants, grocery stores, etc.). Occupational Safety and Health Administration officials could use pen computers for inspections of workplace environments. Aviation Safety Inspectors could use pen computers to speed data collection, information retrieval, information distribution, and certification.

3.4 COMPARISON OF PEN COMPUTERS

Eight pen computer models from a variety of manufacturers were obtained, evaluated, and compared on the basis of CPU type and speed, hard disk capacity, display type, weight, ruggedness, cost, and a number of other factors. (Specific computers are hereafter identified as Computer #1 through Computer #8.) The following specifications, figures, and tables describe the results of that evaluation. While none of the pen computers evaluated could be considered "perfect," some were clearly better than others. While pen computers come in a variety of models and configurations, these units were selected because they are all capable of running Windows for Pen Computing. The models evaluated represent the bulk of the currently available pen computers that will run that operating environment. (Computers that use the NEC V.20 CPU are incapable of supporting Windows; therefore, computers that use this type of CPU were not evaluated.) Pertinent specifications of the evaluated units can be found in Appendix A.

(Two of the units, Computers #4 and #5, were not available in time for a "hands-on" evaluation. The specifications reported here were obtained from printed materials from the manufacturer and from published reports.)

3.4.1 Evaluated Characteristics

Central Processing Unit. Central Processing Unit (CPU) type and speed are central to the response time of a computer. An 80386 CPU should be considered the absolute minimum for portable use, particularly if running Windows. Indeed, one would be hard pressed to find a currently manufactured portable computer that uses an older generation processor. Whereas an 80386 is a minimum, it is difficult to conceive of a unit that is too powerful; many portable computer manufacturers are unveiling 80486-based models. In the future we may expect more powerful CPUs, such as an 80586.

The clock speed of the CPU affects response time nearly as much as the type of CPU. A 20 megahertz clock rate is an effective minimum for portable use, particularly when using the handwriting recognition software that comes with Windows for Pen Computing. While desktop computers now have clock rates of 33, 50, or 66 megahertz, pen computers typically have a 25 megahertz upper limit. However, the higher the clock rate, the better the response time.

Most of pen computers use an 80386 CPU with a 25 megahertz clock rate. However, Computers #2 and #3 have a 20 megahertz clock rate (and both are produced by the same manufacturer). Computer #4 uses an 80486 CPU with a 20 megahertz clock rate; given the increased power of the 80486 CPU, this unit is faster than the units that had 25 megahertz '386 CPUs. Another model from the same company, Computer #5, has an 80486 CPU with a 25 megahertz clock rate. (A 80486/33 megahertz model is to be released later this year.) One manufacturer chose to use a special low voltage 80386 CPU in its computer, Computer #7, because it allows for a smaller battery, thus reducing weight. **Figure 3.2** compares the CPU characteristics of each unit.

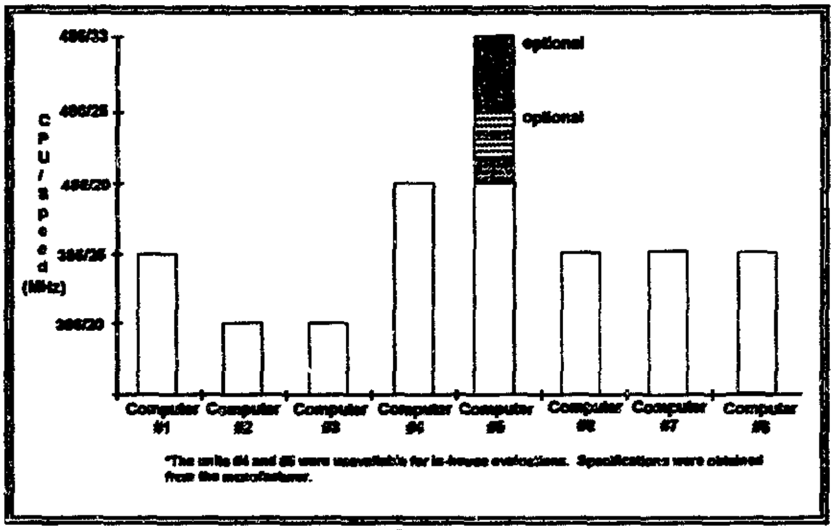


Figure 3.2 Comparison of CPU Type and Speed

Hard Disk Capacity. The hard disk capacity varied greatly across pen computer products. Capacities ranged from 40 megabytes (Computers #2 and #7) to 190 megabytes (Computer #8). Although there can be diminishing returns for large hard disks (in terms of capacity versus cost), software is becoming more space-intensive, particularly with regard to Windows programs. For example, the Windows operating software can use over seven megabytes of disk space, and a typical word processing application can use over 10 megabytes of space. Therefore, 40 megabytes is an effective lower limit on capacity, while 190 megabytes cannot be considered excessive. Currently Computer #7 is limited to a 40 megabyte hard drive because that is the only available size that runs on the low voltage system chosen by the manufacturer. **Figure 3.3** represents the distribution of disk capacities for the evaluation units.

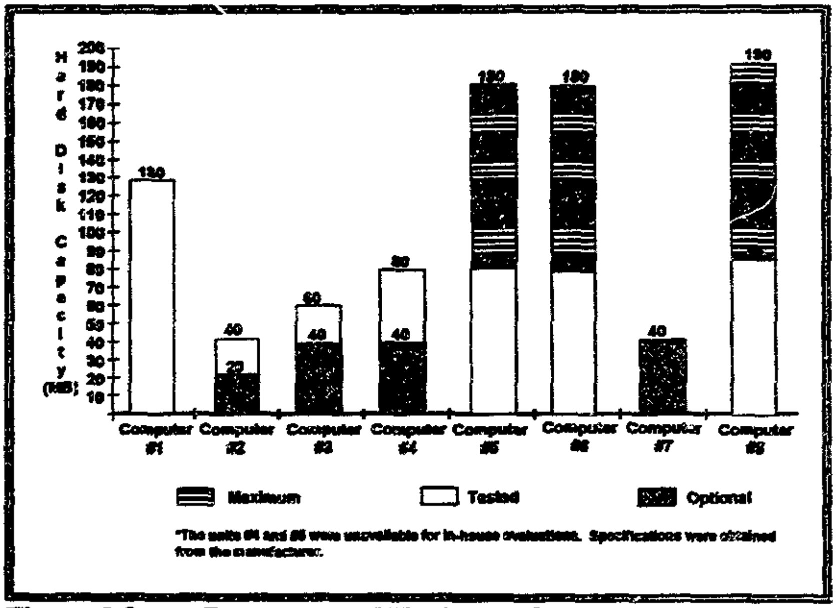


Figure 3.3 Comparison of Hard Disk Capacities

<u>Display Type</u>. Display type greatly affects the ability to read the display in a variety of lighting conditions. Transflective displays work best in bright light, and they work fine in typical indoor lighting. However, transflective displays are nearly impossible to read in the dark. Backlit displays work best in the dark, and they work fine in typical indoor lighting. However, backlit displays can be difficult to read in bright light (this problem is greatly ameliorated by separate brightness and contrast controls). (It is extremely difficult to describe exactly what transflective and backlit displays are or how they look. One really needs to see these displays to understand more than the facts that one works best in the light and the other works best in the dark.) Computers #7 and #8 are unique in that they have backlit displays that can be completely turned off, in which case the display becomes transflective. Computer #5 is the only unit available with a color display. Figure 3.4 depicts the display types of the compared units.

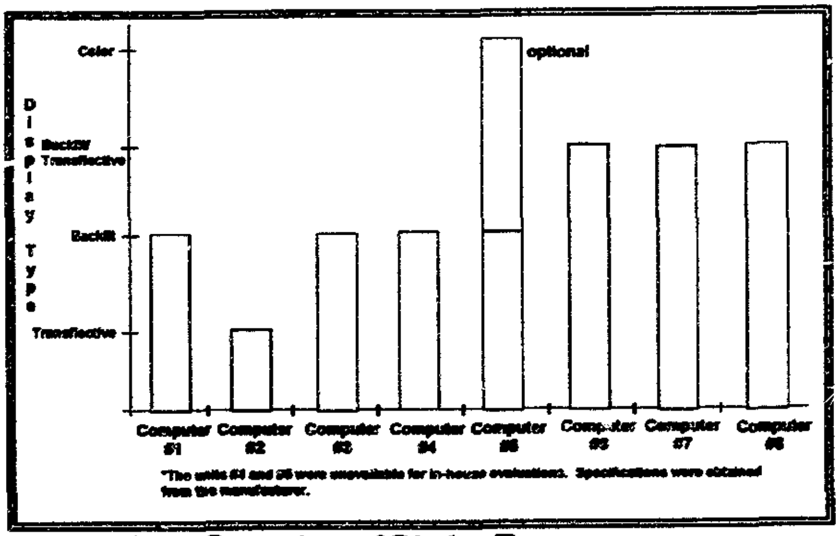


Figure 3.4 Comparison of Display Types

<u>Weight</u>. Weight is a critical factor when evaluating computers for field use. The computer must be easy to hold and carry for a significant portion of the workday. The weight of pen computers is highly correlated with ruggedness; the more rugged a machine, the heavier it is likely to be. Computer #7 was the lightest evaluated unit at 3.3 lbs. Computer #8, which is ruggedized, was the heaviest unit at 7.5 lbs. (The unit evaluated was a pre-production unit; production units are supposed to weigh 6.5 lbs.--which would still make it the heaviest pen computer.) Computer #1, because of its built-in keyboard, is toward the upper end of the weight range. Computer #5, which also has a built-in keyboard, weighs 7.0 lbs., which is also at the upper end of the weight range. However, the pen tablet can be removed from the keyboard base unit to lighten the load. Figure 3.5 shows the weights of the evaluated units.

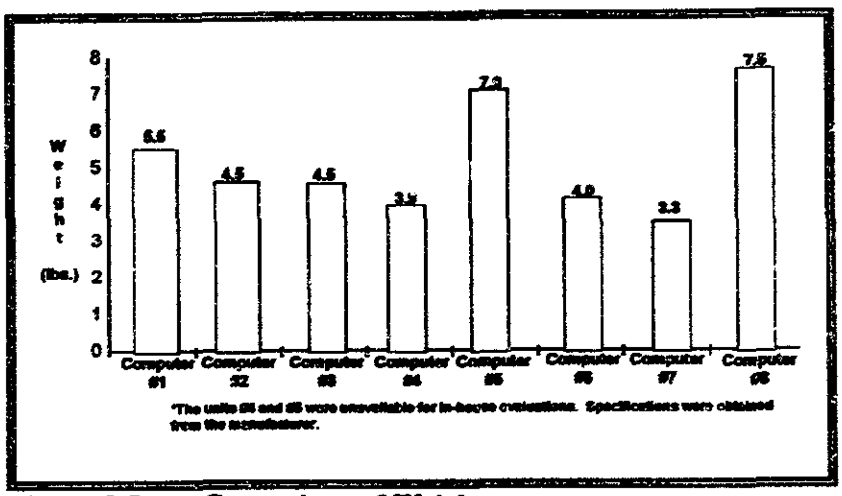


Figure 3.5 Comparison of Weights

Battery Life. Another factor that affects the utility of a portable computer is battery life; that is, how long can the computer be operated before it is necessary to change the battery or recharge it. Battery life can be extended by incorporating various energy-saving features into the computer. One of the most common energy-saving features is to shut off components that are idle. For example, if the hard disk has not been used for a period of time, say 30 seconds, it can be shut down to conserve power. Similarly, the screen and the CPU can be turned off when not in use. The amount of time that passes before a component is shut down is usually adjustable by the operator. While pen computer manufacturers tend to specify long battery lives based on all energy-saving features enabled, real use tests indicate that a standard pen computer battery will last about an hour of continuous use when none of the energy-saving features are enabled. (This assumes the display is backlit; a display that is not backlit consumes considerably less power. For example, Computer #2 will run approximately two hours on a single charge.) Computer #8 is an exception to this rule in that it will run approximately two hours on a single charge with the display on. The pen computer manufacturers are combating the problems associated with short battery life in several ways: the manufacturer of Computer #7 supplies two batteries as standard equipment with their product; Computer #8 uses a quick-charge battery that recharges in an hour; the manufacturers of Computers #7 and #8 both use non-replaceable backup batteries that allow the operator to change main batteries without shutting down the unit; Computer #1 has an optional heavy duty battery that gives two hours of continuous use. All of the manufacturers have designed their units such that a dead battery can be quickly replaced with a fully charged one. **Figure 3.6** compares the battery lives of the evaluated units, with and without energy conserving features turned on. The manufacturer's specifications should be taken with a grain of salt.

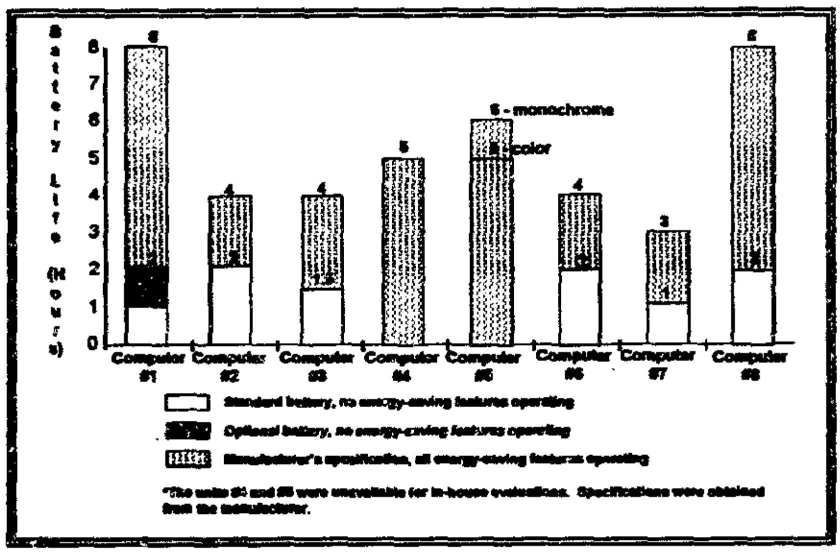


Figure 3.6 Comparison of Battery Life

Ruggedness. Ruggedness, or the immunity from damage due to drops, collisions, water, extreme temperatures, etc., can be an important criterion on which to evaluate pen computers. Most field environments are rather harsh compared to the typical office environment. Instead of sitting quietly on a piece of furniture, as would a desktop computer, a field computer will (at a minimum) be subjected to a lot of handling. In the course of such handling, it is likely that: the computer will be dropped; the operator will bump into things while operating the computer; it will rain or snow on the computer; or, the computer will be left on the dash of a locked car on an August afternoon. All of these things can take a toll on the hardware if it is not designed with such factors in mind. While Computer #8 is specifically designed to handle such environments, most of the other evaluated units were designed to be semi-rugged. Instead of making ruggedness a fixed aspect of their units (thus making them heavy all of the time), the manufacturers of these units have opted for ruggedized carrying cases. These cases improve impact and water resistance, and they have straps and handles to ease carrying the computer. **Table 3.1** summarizes the ruggedness characteristics of the evaluated units, along with a number of other factors.

Other Factors. Other factors that contribute to the desirability of a given pen computer over another include such things as a built-in keyboard, separate brightness and contrast controls, a standard internal or external floppy disk drive, the number of PCMCIA slots, a built-in fax/modem, and the size of the unit. There are tradeoffs involved with many of these factors. For example, a computer that has one PCMCIA slot and an internal fax/modem is nearly equivalent to a computer that has two PCMCIA slots but no internal fax/modem. (One of the PCMCIA slots can be used for a fax/modem. A fax/modem can be very important to remote field workers who need to communicate with others in other field locations or at a central office.) **Table 3.1** lists the factors and their presence on the units.

Price. For many people, the determining factor on purchasing a pen computer will be price. It is important to realize that although the initial cost of a pen computer is likely to be higher than an equivalently equipped notebook computer, pen computers also weigh less and are smaller. Furthermore, pen computers were designed to speed data collection and reduce reliance on data-entry personnel. In other words, many agencies who use paper forms rely on data entry clerks to read the data off those forms and transcribe them into a computer database. Data entry clerks are an intermediate step in the data collection and distribution process; such intermediate steps can reduce data integrity and slow assimilation of the data into databases. Pen computers allow data to be directly entered in the proper database format at the time of collection. This method ensures that data are entered correctly and it reduces reliance on data entry personnel. While price is highly correlated with capabilities, it is not a reliable indicator. Price should be one factor used to choose between pen computer models, but it should not be the sole factor. The prices of the pen computers evaluated are shown in **Figure 3.7**

3.4.2 Common Features of Pen Computers

All of the evaluated pen computers had an external keyboard port (with the exceptions of Computer #1 and Computer #5, which had built-in keyboards). The units also had serial and parallel ports; on computers other than Computers #2, #3 and #8, the parallel port doubled as a floppy disk drive connector. Serial ports are often used for communications (including loading software onto the pen computer) and a mouse (which, of course, is unnecessary when using a pen stylus as a pointing device). Parallel ports are also used for communications, but their primary use is for connecting printers. Computer #7 was the only model that did not have a VGA port; a VGA port is a convenient feature if the operator wants to display items on a large monitor instead of the pen display. All of the units were configured with eight megabytes of RAM; this is an effective minimum for running Windows for Pen Computing.

Table 3.1 Miscellaneous Pen Computer Characteristics

	Computer #1	Computer #2	Computer #3	Computer #4	Computer #5	Computer #6	Computer #7	Computer #8
Rugged	Optional Case	Optional Case	Optional Case	Optional Case	Optional Case		Optional Case	Yes
Built-in keyboard	Yes	No	No	No	Yes	No	No	No
Number of Brightness and Contrast Controls	2	1	1	2	2	2	2	2
External floppy disk drive standard	Yes	No	No	Yes	Yes*	No	Yes	No
Number of PCMCIA slots	1	1	1	2	2	1	2	0
Built-in fax/modem	Yes	No	No	No	No	No	No	Yes
Dimensions	1.6"H x 11.5"W x 9.3"D	1.2"H x 11.7"W x 9.4"D	1.2"H x 11.7"W x 9.4"D	1.2"H x 10.9"W x 9.8"D	2.1"H x 11.7"W x 9.3"D	1.3"H x 11"W x 11"D	1.5"H x 10.6"W x 8.3"D	2.0"H x 12.5"W x 10.1"D

*internal

3.4.3 Common Limitations of Current Pen Computers

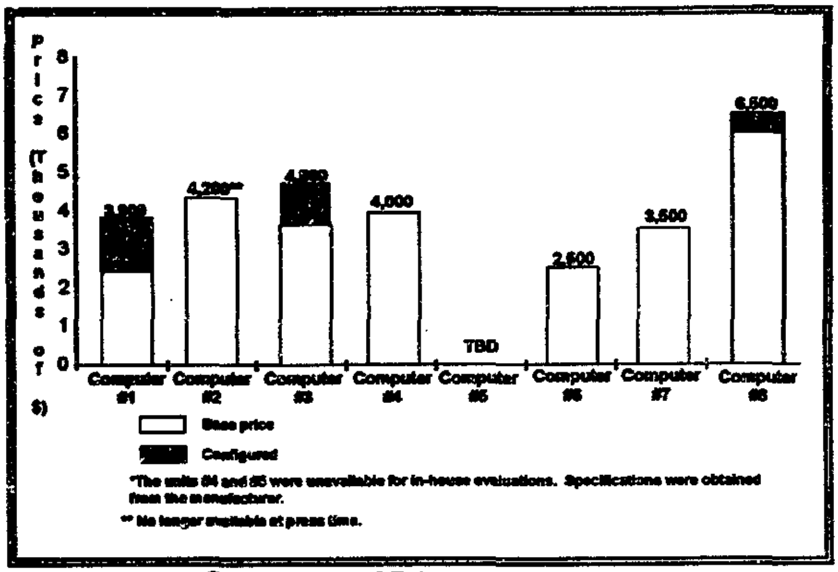


Figure 3.7 Comparison of Prices

As mentioned above, a current common limitation to pen computers is the short battery life. Batteries also make significant contributions to the weight of a unit. Several manufacturers are working to improve battery life, while maintaining lightweight batteries. Many current pen computers use nickel-cadmium (NiCad) batteries, which appear to have an upper limit on battery life of one hour (given a backlit display and current technology). Nickel-metal hydride batteries are now beginning to appear; these batteries appear to have longer lives than do NiCad batteries. It is only a matter of time before compact, lightweight, long-life batteries are available.

On the subject of batteries, a limitation of all current pen computers, except Computers #2 and #3, is that the pen stylus requires batteries. Unless there is an accompanying keyboard, there is no way to operate the computer if the pen stylus batteries die. The pen stylus batteries are typically readily available (and inexpensive) calculator or hearing aid batteries. It is probably a good idea to carry a spare set of pen stylus batteries in the field. The pen styli for Computers #2 and #3 use internal inductors that are sensed by the radio frequency grid on the computer. Other manufacturers are probably moving in the inductor direction.

The biggest limitation of pen computers is the handwriting recognition software. Recognition accuracy is unacceptably low when using the recognizer that comes with Windows for Pen Computing. The standard recognizers that run under other operating environments, such as PenPoint and PenRight, do not perform any better. However, there is a third-party recognizer, Nestor Writer, that has very high recognition accuracy. According to a recent publication (Bachus & Weston, 1993), a novice user can get initial accuracy rates above 80%. The authors reported rates above 90% after one hour of training. This software is currently available for PenPoint and should be available for Windows within six months. While all currently available recognizers require printed text input, several companies are working on cursive recognizers.

3.4.4 Tradeoffs Between Pen Computers

It should be apparent that the perfect pen computer has not yet been invented. As mentioned above, there are several tradeoffs between the features of pen computer models and each model has features that make it unique. Computer #1 is modestly priced and has a built-in keyboard, but the keyboard makes the unit heavy. Computers #2 and #3 have the benefit of not requiring batteries for the pens, but they are more expensive than similarly configured units. Computer #4 has an 80486 processor and is lightweight. Computer #5 also has an 80486 processor, a color display, and a built-in keyboard. However, these units are rather expensive. Computer #6 has a handle that doubles as a stand and it has a "hot dock." The hot dock allows the unit to be quickly placed in a docking station that supplies power and connects the unit to a network. Computer #6 falls in the middle of the group in terms of overall performance. Computer #7 has two PCMCIA slots, is the lightest unit evaluated, and has a backlit/transflective display. Unfortunately, it has a small hard disk and the evaluated pre-production units had very short life main and pen batteries. Computer #8 is very rugged and has quick-charge batteries, but it is heavy and expensive.

Such tradeoffs make it difficult, if not impossible, to dictate which unit to purchase for a given application. When implementing any application, the designer must perform an extensive field evaluation to fully understand which features are most important to the people who will actually be using the equipment. Just such an evaluation is proposed in the next section.

3.5 PENS: A PERFORMANCE ENHANCEMENT SYSTEM

The Performance ENhancement System, PENS, is a tool to aid Aviation Safety Inspectors in performing their tasks. Aviation Safety Inspectors (ASIs) make up the inspection team for the FAA. Aviation Safety Inspectors perform a variety of tasks, in both commercial and general aviation areas, including: inspecting aircraft and equipment, reviewing manuals and records, certificating pilots, and evaluating training programs.

There are 2600 ASIs in the nine regions of the FAA. The initial target of PENS is an ASI performing an airworthiness (safety) ramp inspection. (A ramp inspection consists of inspecting an aircraft, while it is at the gate, before a scheduled departure.) PENS is an electronic performance support system (Gery, 1991) that consists of two components: a "smart" forms application and an on-line documentation system. PENS capitalizes on the recent advances in pen computer technology outlined above.

3.5.1 Improved Forms

As is typical with regulatory agencies, there are several forms that must be completed while performing an ASI task. Currently, these forms are on paper and require that redundant information be recorded on each form. After completing the forms, the ASI either types the data into a local computer database or he/she submits the forms to a data entry clerk. There are several drawbacks to such an approach. First, redundant recording of data on multiple forms takes time that could be devoted to more productive activities. Second, the two-step process of recording data on paper and then entering the data into a computer is inefficient. Third, one is either paying an inspector to do a task for which he/she is over-qualified, or one is paying for a staff of data entry clerks. Fourth, a data-entry clerk may make transcription errors (due to misreading the inspector's handwriting) or errors due to incomplete knowledge and understanding of the inspector's activities. Such errors mean that the database is an unreliable source of information. Finally, the current process takes considerable time, which means there is a delay in getting safety data into the national database where it can be accessed by other members of the FAA.

Pen computer technology can be easily applied to such tasks to minimize the number of steps required to collect data and assimilate it into the database. Forms will be linked together so that an entry in one form propagates to the other forms, thus eliminating redundant data entries. Furthermore, the data will be collected so that they are ready for direct downloading into the database. This method of collecting data reduces the need for data entry clerks and it reduces data transcription errors. At the end of the work day, the inspector will return to the office, connect the pen computer to the network, and initiate a downloading procedure that will be carried out overnight.

3.5.2 On-line Documentation

The second major contribution of PENS is an on-line documentation system. Whereas ASIs currently must carry two briefcases full of books (including Federal Aviation Regulations, ASI Handbooks, and other regulatory documents), the necessary data will be stored on the hard disk of the pen computer or on a CD-ROM (compact disc, read-only memory). Not only is the computer media more lightweight and compact, it also facilitates quick retrieval of specific information. For instance, an ASI will be able to search the regulations for the word "corrosion" to answer a question on reporting defects. PENS would then indicate all of the instances of the word corrosion. The ASI could then ask PENS to retrieve the relevant documents and display the pages that discuss the term.

Besides the bulk and inefficiency of the books, inspectors must deal with problems of information currency. One complaint made by inspectors is that they will tell an operator that it is not in compliance with the regulations, only to be shown a more recent edition of those regulations. That is, sometimes the operators get the most recent editions of the regulations before the inspectors do. This problem could be dealt with by distributing updated documents to the pen computers when they are connected to the database computer network. Thus, a new edition of a document could literally be published one day and in the inspector's hands the next.

3.5.3 Additional Benefits

A side benefit of using a computer to support inspection activities is that it opens the door to other types of activities and methods for documenting an inspection. For example, an inspector could follow an on-line checklist for an inspection. The checklist would then become the focus of interaction with the computer; by completing the checklist, all of the necessary forms would be automatically completed. We could even develop a scheduling component that would remind the inspector to follow up on an inspection. When documenting an inspection, ASIs currently must record their findings verbally. However, because the bulk of a ramp inspection is conducted by visually inspecting an aircraft, sketching is a more natural method for recording the results of such an inspection. Thus, if an inspector found a leaking seal on the wing of an aircraft, the inspector could annotate a line art drawing of that aircraft on the computer. This graphic could then be stored along with the completed form.

Another important benefit of giving ASIs computer-based inspection tools is that it would greatly ease inspection of air carrier records. Nearly all air carriers keep their records in computer files, as well as paper files. (At least one airline has only computer records.) Whereas searching paper files for specific data can be tedious and cumbersome, computer databases were designed for just such activity. Indeed, some industry officials are promoting the notion of allowing the FAA to inspect their records:

> "We're not taking advantage of the data systems airlines have in place," [a] senior vice president of technical operations at [an airline], said. Those systems could be the foundation of a new surveillance system "that penalizes bad behavior and rewards good behavior." (McKenna, 1992)

The proposed concept would consist of reducing the frequency of inspections for operators who consistently meet airworthiness standards, while increasing the frequency of inspections for those operators who do not meet those standards. (A similar concept is already applied to other types of activities.) This approach should benefit the airlines by streamlining maintenance (thus reducing costs due to out of service aircraft) and reducing the amount of company time spent on inspections.

3.6 EVALUATION AND IMPLEMENTATION

There are a number of issues that can affect the success of introducing new technology into the ASI work environment. Many inspectors do not have experience using computers. Of those inspectors, some are willing to try the new tools based on promised increased productivity, while others are hesitant to embrace a new method for performing their work. Some inspectors are even concerned with how they will be perceived by the operators when they are carrying a pen computer.

Perhaps the most significant hurdle to widespread implementation of PENS, however, is the adequacy of the handwriting recognition software. The difficulties involved with handwriting recognition (writer independence, print vs. cursive writing, intraindividual variations in writing style) are directly analogous to difficulties with speech recognition; however, handwriting recognition is five to ten years behind speech recognition. Although much research and development is going in to new methods for handwriting recognition, we cannot wait for such advancements before fielding a system. Therefore, we are capitalizing on constraints built into the forms and data to reduce dependence on handwriting recognition. For instance, because many fields on the forms require one item out of a finite set of possible entries, one can display that set and select an item from it. This approach has the added benefits of reducing memory demands on the inspectors and of increasing data reliability.

Pen computer configurations and durabilities must also be considered, as there are significant tradeoffs in these areas. Questions that should be asked include: Is it better to have a lightweight unit without a keyboard, or a slightly heavier unit with a keyboard? Which is more important to inspectors, weight or ruggedness? Is battery life sufficient to even consider using such a device? Appendix B lists these questions and others, along with our recommendations. These recommendations are based on very informal evaluations, however, and should be considered only as preliminary guidelines.

3.6.1 Evaluation Plan

Given the above concerns, the following evaluation is proposed as a means to assess the utilities of various hardware configurations and the effectiveness of the software. We expect to modify the software based on inspector feedback, but the field evaluations will largely determine which models of pen computer hardware will be put into actual use. Although we expect the hardware to withstand most environmental conditions, it is possible that some extreme conditions will preclude the use of computer hardware. The following experimental plan will provide inspectors with experience with a range of models and it will subject the hardware to a range of operating environments.

3.6.2 FAA Regions

We will field units in six to nine Regions, in a variety of locations. This will give the project broad exposure to field inspectors and it will subject the hardware to a range of environmental conditions. The six Regions identified below are suggested based on the worst-case environmental conditions present in those regions.

Suggested location:	Reasons:
Alaska	Cold, snow
Northwest Mountain	High humidity, low temp., rain
Central	Average temp., average humidity
New England/Eastern/Great Lakes	Cold, snow
Southwest	Low humidity, heat
Southern/Western Pacific	High humidity, rain, heat

3.6.3 Pen Computer Models

We will field four different models, each from a different manufacturer; this will reduce reliance on one manufacturer and it will help identify design factors important to the inspector population. From this evaluation, it is likely that two of the models (or subsequent versions of them) would be chosen for final implementation. However, all of the purchased units would remain in service after the evaluation.

Each unit would have nearly identical hardware configurations, so as not to bias the results.

Seventy-two computers (and peripherals) will be purchased; this will provide each Region with two units from each manufacturer.

Suggested Computer:	Reasons:
Computer #1	Built-in keyboard
Computer #4	80486, medium weight
Computer #7	Lightest
Computer #8	Rugged

(Note: Computer #5 is also an 80486 unit with a built-in keyboard; this may be used instead of Computer #4.)

3.6.4 Experimental Design

A team of eight inspectors in each Region will evaluate these units. These inspectors will represent a cross-section of the inspector population in terms of age, sex, work experience, and computer experience. Each inspector will use one of the computers for a week and then switch to a different model. The rotation would be counterbalanced to eliminate order effects. This rotation will continue until each inspector has had an opportunity to use each model. At the end of the rotation, each inspector will complete an evaluation form (sample attached) that requests him/her to rate each unit and answer some general questions. The inspectors should still have access to the units at this time to refresh their memories of the specifics of each unit. From these data, we will recommend two of the models (or their subsequent versions) for final implementation.

3.7 SUMMARY AND CONCLUSIONS

As discussed above, pen computers use handwriting recognition software and a pen stylus for input, rather than a keyboard. The operator writes on the screen and the handwriting recognition software translates the written characters to typed characters. The pen stylus also acts as a pointing device, much like a mouse. When combined with graphical user interfaces, such as Windows for Pen Computing or PenPoint, the pen stylus and handwriting recognition software hold the promise of making computers easier to use than traditional desktop computers. Many pen computer models from a variety of manufacturers have undergone preliminary in-house evaluations. These evaluations have identified several differences in the design of such devices and have identified some tradeoffs involved in these design choices. While such evaluations are valuable, they should be seen as only a first step in selecting equipment; final selections must be made based on field evaluations by the actual user population.

As with the introduction of any new tool into an existing system, the effects are widespread (Chapanis, 1982; Helmreich, 1987; and, London, 1976). The potential for enhancing the productivity and job satisfaction of Aviation Safety Inspectors is great. However, with that potential comes the possibility of either having no effect (because of rejection of the tool) or, worse yet, actually decreasing performance. The PENS project is taking a cautious, iterative approach to design and introduction of the tools. Only through careful cognitive task analysis, rapid design and prototyping, and empirical evaluation will PENS be seen in the eyes of the inspectors as a beneficial cognitive tool, rather than another doorstop or paperweight.

3.8 REFERENCES

Bachus, K. & Weston, R. (1993, January). Pen power. *Corporate Computing*, 68-95.

Chapanis, A. (1982). Computers and the common man. In R. A. Kasschau, R. Lachman, & K. R. Laughery (Eds.), *Information technology and psychology: Prospects for the future* (pp. 106-132). Praeger.

Gery, G. J. (1991). *Electronic performance support systems* (2nd ed.). Boston: Weingarten.

Helmreich, R. (1987). Changing potential users to actual users: An evolutionary approach to office system acceptance. In M. Frese, E. Ulich, & W. Dzida (Eds.), *Human-Computer Interaction in the Work Place* (pp. 81-94). Amsterdam: North-Holland.

London, K. R. (1976). *The people side of systems.* London: McGraw-Hill.

McKenna, J. T. (1992, December 7). Airlines seek better sharing of safety data with FAA. *Aviation Week & Space Technology*, p. 43.

Appendix A, Computer Specifications

Computer #1

Features
Dimensions--1.6" H x 11.5" W x 9.3" D
80386/25 MHz CPU
8 MB RAM (2 MB Std., 6 MB upgrade)
130 MB Hard Drive
Built-in keyboard
External floppy drive (parallel port) standard
Sidelit (backlit) 9.5" 64 shade VGA LCD display (blue) with brightness and contrast controls
Optional built-in FAX/Modem (2400 baud modem/9600 baud fax or 14.4 kbaud modem/9600 baud fax)
Serial port
Parallel port/floppy disk drive port (requires adaptor for parallel port)
Monitor port
1 PCMCIA slot
Battery-operated pen
Computer battery is replaceable
Operating temperature range 41 to 104 degrees F
Storage temperature range -4 to 140 degrees F
Operating relative humidity 10% to 80% noncondensing
Storage relative humidity 5% to 80% noncondensing
Shock tolerance--operating 5g, nonoperating 80g
Vibration tolerance--3-200-3 Hz at 0.4g (operating); 3-200-3 Hz at 1.5g (nonoperating)
Altitude--operating 10,000 ft; nonoperating 40,000 ft
Electrostatic discharge 15kV
Die-cast magnesium and injection-molded thermoplastic case
5 1/2 lbs.
$2796 base price; $4070 configured

Drawbacks
1 hour battery life with standard battery and energy conservation features disabled.
Pen ink is "noisy"; abated somewhat by supplied filtering software.
Keyboard only usable in landscape display rotation.

Other Factors
Pen "feel" simulates a felt tip pen on paper.
Screen is good indoors and outdoors (but has some glare).
Optional heavy duty battery.
Optional cigarette lighter power adaptor.

Opinion

One of the favorites until handwriting recognition produces near 100% accuracy. Even then, it will be a good unit because the keyboard gives the unit more flexibility and because many people can type faster than they can write.

Computer #2 and Computer #3

Features
Dimensions--1.2"H x 11.7"W x 9.4"D
80386/20 MHz CPU
8 MB RAM (4 M Std., 4 M upgrade?)
40 or 60 MB Hard Drive
Transflective 16 shade VGA LCD display (green)--Computer #2
Backlit 16 shade VGA LCD display (blue) with single brightness/contrast control--Computer #3
Serial port through docking strip
Parallel port through docking strip
Monitor port through docking strip
Operating temperature range 41 to 104 degrees F
Storage temperature range -4 to 122 degrees F
Operating relative humidity 5% to 95%
Altitude--operating 9800 ft; nonoperating 40,000 ft
No battery in pen
Rechargeable computer battery is replaceable
Optional FAX/Modem.
4 1/2 lbs.
$5350 configured

Drawbacks
Choice between Flash Disk and hard drive.
Pen "feel" is very slick, uncomfortable.
Parallel port does not support floppy disk drive; one must choose between a fax/modem and a floppy drive controller.

Other Factors
A favorite pen stylus because of slim design, out-of-the-way button.
Battery life of Computer #2 is longest of tested units because display is not backlit.
Terrible sales support.

Opinion
Computer #2 was discontinued while this report was being written. Computer #3 is probably in the bottom third of 386-based machines in terms of features, performance, and price.

Computer #4

Features
Dimensions--1.2" H x 10.9" W x 9.8" D
80486/20 MHz CPU
8 MB RAM (4 MB Std., 4 MB upgrade)
40 or 80 MB Hard Drive
Optional external keyboard
External floppy drive standard
Backlit 9.4" 64 shade VGA LCD display (blue) with brightness and contrast controls
Optional PCMCIA FAX/Modem (2400 baud modem/9600 baud fax or 14.4 kbaud modem/fax)
Serial port
Parallel port/floppy disk drive port (requires adaptor for parallel port)
Monitor port
2 PCMCIA slots
Battery-operated pen
Computer battery is replaceable
Operating temperature range 0 to 45 degrees C
Storage temperature range -20 to 60 degrees C
Operating relative humidity 0% to 85%
Storage relative humidity 0% to 95% noncondensing
3.9 lbs.
$3999 base price; $5770 fully configured MSR

Drawbacks
TBD

Other Factors
TBD

Opinion
TBD

Computer #5

Features
Dimensions--2.1" H x 11.7" W x 9.3" D
80486/20 MHz or 80486/25 MHz CPU
8 MB RAM (4 MB Std., 4 MB upgrade)
80, 120, or 180 MB Hard Drive
Built-in keyboard
Internal floppy drive standard
Backlit 9.4" 64 shade VGA LCD display or 9.4" 256 color Super VGA active matrix display
Optional PCMCIA FAX/Modem (2400 baud modem/9600 baud fax or 9600 baud modem/14.4 kbaud fax)
Serial port
Parallel port
Monitor port
2 Type II or 1 Type III PCMCIA slots
Battery-operated pen
Computer battery is replaceable
Optional second battery
Optional docking station
7.0 lbs.; but display can be removed to function as pen-only tablet
Price TBD

Drawbacks
Keyboard only useable in landscape display rotation.

Other Factors
TBD

Opinion
TBD

Computer #5

Features
Dimensions--2.1" H x 11.7" W x 9.3" D
80486/20 MHz or 80486/25 MHz CPU
8 MB RAM (4 MB Std., 4 MB upgrade)
80, 120, or 180 MB Hard Drive
Built-in keyboard
Internal floppy drive standard
Backlit 9.4" 64 shade VGA LCD display or 9.4" 256 color Super VGA active matrix display
Optional PCMCIA FAX/Modem (2400 baud modem/9600 baud fax or 9600 baud modem/14.4 kbaud fax)
Serial port
Parallel port
Monitor port
2 Type II or 1 Type III PCMCIA slots
Battery-operated pen
Computer battery is replaceable
Optional second battery
Optional docking station
7.0 lbs.; but display can be removed to function as pen-only tablet
·Price TBD

Drawbacks
Keyboard only useable in landscape display rotation.

Other Factors
TBD

Opinion
TBD

Computer #6

Features

Dimensions--1.3"H x 11"W x 11"D
80386/25 MHz CPU
8 MB RAM Std. (up to 20 MB)
80 MB Hard Drive (up to 120 MB)
Backlit/Transflective (backlighting can be turned off) 64 shade VGA LCD display (blue) with
brightness and contrast controls
Serial port
Parallel port
Floppy disk drive port
Monitor port
Built-in 9600 baud FAX/2400 baud Modem
"Hot Dock" docking port for power and/or other connections.
Battery-operated pen
Rechargeable computer battery is replaceable; 2 batteries std.
4 lbs.
$2500 base price to VAR; $3500 MSR

Drawbacks

Screen needs constant adjustment; screen is not evenly lit. Pre-production problem?
Pen scratches screen coating. Pre-production problem?
Pen is difficult to remove from holder
Slow hard disk drive.

Other Factors

No button on pen (pen buttons hinder more than they help).
Nice built-in, adjustable handle.
Units manufactured by under contract by large computer manufacturer.
Designed for landscape display rotation, but usable in portrait rotation.
CPU battery life is 2 hours hard use, 4 hours normal use. Well above average.

Opinion

Unit is relatively well-designed, but unspectacular.

Computer #6

Features
Dimensions--1.3"H x 11"W x 11"D
80386/25 MHz CPU
8 MB RAM Std. (up to 20 MB)
80 MB Hard Drive (up to 180 MB)
Backlit/Transflective (backlighting can be turned off) 64 shade VGA LCD display (blue) with brightness and contrast controls
Serial port
Parallel port
Floppy disk drive port
Monitor port
Built-in 9600 baud FAX/2400 baud Modem
"Hot Dock" docking port for power and/or other connections.
Battery-operated pen
Rechargeable computer battery is replaceable; 2 batteries std.
4 lbs.
$2500 base price to VAR; $3500 MSR

Drawbacks
Screen needs constant adjustment; screen is not evenly lit. Pre-production problem?
Pen scratches screen coating. Pre-production problem?
Pen is difficult to retrieve from holder.
Slow hard disk drive.

Other Factors
No button on pen (pen buttons hinder more than they help).
Nice built-in, adjustable handle.
Units manufactured by under contract by large computer manufacturer.
Designed for landscape display rotation, but usable in portrait rotation.
CPU battery life is 2 hours hard use, 4 hours normal use. Well above average.

Opinion
Unit is relatively well-designed, but unspectacular.

Computer #7

Features
Dimensions--1.5" H x 10.6" W x 8.3" D
80386/25 MHz CPU
8 MB RAM (4 MB Std. vs 4 MB upgrade)
40 MB Hard Drive
3.3 Volt "Low Power System"
External floppy drive (parallel port?) standard
Backlit/Transflective (backlighting can be turned off) VGA LCD display (gray-brown) with brightness and contrast controls
Serial port
Parallel port/floppy disk drive port (requires adaptor for parallel port)
Keyboard port
Monitor port
2 PCMCIA slots
Battery-operated pen
Computer battery is replaceable
3.3 lbs.
$3499 base price; $???? configured

Drawbacks
Pen batteries need frequent replacement (eg. weekly).
Hard disk is largest available for given dimensions, but is still too small.

Other Factors
Screen is good indoors and outdoors.
The 2 PCMCIA slots will allow a data card and a fax/modem card simultaneously.
Not rugged, but 3rd party is designing a "wetsuit", rubber, ruggedized case.
Tested unit was early production.

Opinion
Assuming they can solve the pen battery problems, this will be a nice, small, lightweight, pen-only unit.

Computer #8

Features
Ruggedized
Dimensions--2.0"H x 12.5"W x 10.1"D
80386/25 MHz CPU
8 MB RAM Std.
85 MB or 190 MB Hard Drive
Soft keys built into bezel of display
Backlit 64 shade VGA LCD 10" display (green) with brightness and contrast controls; optional 11.6" SVGA 64 shade display
Built-in 9600 baud FAX/2400 baud Modem Std.; optional 9600/9600 FAX/Modem
Serial port
Parallel port
Keyboard port
Monitor port
Optional docking station
Ballistic-composite main housing with aircraft aluminum and stainless steel fittings
Battery-operated pen
Quick-charge 3 hour computer battery with backup battery.
6.5 lbs.
$5995 base price; $6495 configured w/o docking station; $6990 w/docking station

Drawbacks
Weight.
Although battery monitor indicates over 2 hours of charge, it is closer to 1-1 1/2 hours.

Other Factors
Start-up company; company's only product; difficulty in bringing it to the market.
Optional keyboard unit has floppy drive, ports.
Getting a demo unit was extremely difficult; units received were pre-production. Company started production in 11/92.

Opinion
Ruggedness is continually mentioned as a key criterion for field units; this is a good example.

Computer #8

Features
Ruggedized
Dimensions--2.0"H x 12.5"W x 10.1"D
80386/25 MHz CPU
8 MB RAM Std.
85 MB or 190 MB Hard Drive
Soft keys built into bezel of display
Backlit 64 shade VGA LCD 10" display (green) with brightness and contrast controls; optional 11.6" SVGA 64 shade display
Built-in 9600 baud FAX/2400 baud Modem Std.; optional 9600/9600 FAX/Modem
Serial port
Parallel port
Keyboard port
Monitor port
Optional docking station
Ballistic-composite main housing with aircraft aluminum and stainless steel fittings
Battery-operated pen
Quick-charge 3 hour computer battery with backup battery.
6.5 lbs.
$5995 base price; $6495 configured w/o docking station; $6990 w/docking station

Drawbacks
Weight.
Although battery monitor indicates over 2 hours of charge, it is closer to 1-1 1/2 hours.

Other Factors
Start-up company; company's only product; difficulty in bringing it to the market.
Optional keyboard unit has floppy drive, ports.
Getting a demo unit was extremely difficult; units received were pre-production. Company started production in 11/92.

Opinion
Ruggedness is continually mentioned as a key criterion for field units; this is a good example.

Appendix B, Evaluation/Implementation Questions

The following questions need to be addressed when specifying a pen computer for the Flight Standards Service:

Environmental Immunity--How resistant is the unit to temperature extremes (eg. Anchorage in the winter to Puerto Rico in the summer), humidity, rain, etc.?

Ruggedness--Can the unit be dropped? How susceptible is the screen to damage from collision? Will the paint chip from minor collisions? Is a ruggedized case necessary and available?

Harness--Is there a harness or strap to alleviate carrying the unit and preventing damage if dropped?

Weight--Are the units light enough to be carried for an entire work day?

Lighting Conditions--Will current units work in lighting conditions ranging from bright sunlight to absolute darkness?

Display--Is the display monochrome, grey scale, or color? How many shades or colors can be displayed simultaneously? What is the resolution?

Pen--Is there a provision for tethering the pen so that it won't get lost? Will the pens allow user replacement of the batteries, rather than buying a new pen? Is the location of the button on the pen such that it is not accidentally depressed while writing?

Pen Feel--Does the feel of "writing" with the pen on the computer simulate a pen on paper?

Storage Capacity--What is the capacity of the hard disk, in Megabytes?

Speed/Computing Power--What are the fastest and most powerful CPUs available? 80386? 80486? 80586?

RAM--What are the available RAM capacities and speeds? At least 8 Megabytes are required; can more be put in?

PCMCIA card slots--PCMCIA (Personal Computer Memory Card International Association) cards allow peripherals, such as FAX/Modems, CD ROM controllers, and ROM (read only memory), to be easily added to and removed from the unit. These cards are revolutionizing the portable computer industry. How many PCMCIA slots are available?

Keyboard--Is there a keyboard built into the units? Is there a lightweight standard keyboard as an accessory to the units?

CD-ROM Players--Do the units currently support CD-ROM (compact disc, read only memory) players?

FAX/Modem--These allow for communication with computers and other parties over the phone lines. Are they available for the pen computer?

Connectivity--Is the unit capable of supporting a network connection, either through a serial port or a dedicated port? What about wireless connections?

Upgrades--Are there user-replaceable CPUs or other upgrade paths?

Based on in-house evaluations, we recommend field evaluations. However, we can make the following conservative recommendations:

Environmental Immunity and Ruggedness--A ruggedized case that allows one to use the computer while it resides in the case will improve ruggedness and environmental immunity. Most currently available products will function in mist to very light rain. Current pen computers will operate in temperatures ranging from about 20 degrees to 110 degrees Fahrenheit. Because pen computers tend to be very susceptible to damage from dropping, the best approach may be to choose one that is itself semi-rugged, but which has a ruggedized case. A ruggedized case will also allow an inspector to use the unit in the rain, snow, etc. Only one company currently manufactures a unit that is already ruggedized.

Harness--Either the unit itself or a ruggedized case should be equipped with a carrying strap or harness. The currently available ruggedized unit comes with a carrying case that has a strap.

Weight--This is the primary drawback to ruggedized units and becomes a problem when adding a ruggedized case to other units. We think that a unit that does not allow one to remove the ruggedized case will be too heavy for general acceptance. A removable case will allow a minimum weight configuration for most uses, with the flexibility to add the case (and, hence, weight) when required. Current weights range from approximately 3.5 lbs. to approximately 7 lbs. (for the ruggedized unit). A weight of approximately 5 lbs. is probably acceptable, initially.

Lighting Conditions/Displays--Most currently available 80386 based pen computers come with backlit, grey scale, VGA (16 to 64 simultaneous shades of grey, 640 x 480 pixel resolution) displays. Such units allow one to use them in the dark and in bright sunlight. The best such units allow independent control of contrast and brightness. A monochrome display is unacceptable (regardless of resolution); a grey scale, VGA display is acceptable; while color, Super VGA (greater resolution, 256 colors) displays are not currently available, they would be preferred.

Pen--A tethered pen is a little more difficult to use, but it is much more difficult to lose. Given the cost of replacement pens ($75-$100), we would recommend tethering the pen. Again, given pen cost, we recommend that the pen allow user replacement of the batteries. A pen that has a button that prevents one from accidentally depressing it while writing is preferable, but not mandatory (because the pens can usually be rotated such that the button is out of the way).

Pen feel--Ideally, writing on the pen computer would simulate writing with a pen on paper. Some of the available products are better than others in this regard.

Storage Capacity--Current hard disk storage capacities range from 40 Megabytes to 120 Megabytes. While 40 Megabytes could be considered the absolute minimum, the more capacity one can get, the more software programs/tools and data can be stored and used. We would recommend a 120 Megabyte hard drive for now, while keeping in mind that larger capacity hard disks will be available in the future.

Speed/Computing Power--Currently, 80386, 25 Megahertz CPUs are used in most pen computers. However, 80486 and 80586 CPUs with faster clock rates should be available in the near future.

RAM--Because it is likely that the pen computers will be using Windows for Pen Computing as their operating environment, 8 Megabytes of RAM should be considered the absolute minimum requirement. Some manufacturers allow 16 Megabytes or more of RAM. The availability of RAM greatly affects processing speed and response times.

PCMCIA card slots--Most manufacturers offer one PCMCIA slot. Because many desired features of the pen computer could be addressed through the use of PCMCIA cards, two or more slots would be better, although not mandatory.

Keyboard--We recommend that a concealable keyboard be built in, similar to standard notebook or laptop computers; however, the pen computer must be fully functional when the keyboard is concealed. That is, the screen must be visible and allow pen input, even when the keyboard is concealed. The built-in keyboard would allow one to readily enter large amounts of text. Pen computers with built-in keyboards are compact and convenient, whereas detachable keyboards tend to be inconvenient and cumbersome.

CD-ROM--It is becoming increasingly clear that the pen computer will need to support a portable CD-ROM player. For the foreseeable future, a PCMCIA card or a parallel port will support this function. Ideally, a PCMCIA SCSI interface card will be used to drive the CD-ROM because they are faster than parallel port devices. Portable CD-ROM players that use parallel ports are currently available.

FAX/Modem--Most manufacturers offer FAX/Modems either as standard equipment or through a PCMCIA slot.

Connectivity--All pen computers currently supply a parallel port, which would allow connection to a network. We do not recommend that a wireless network connection provide the sole access to networks; a wired connection should be available.

Upgrades--No manufacturer currently supports upgradeable CPUs, but this will likely change. Such upgrades would allow Flight Standards to take advantage of the most recent technology without scrapping the computer itself.

Appendix C, Example Evaluation Form

Please rate the following on a *relative* 1-5 scale, where *1 is worst* and *5 is best*:

	Computer #1	Computer #2	Computer #3	Computer #4
Weight	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Size	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Speed	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Display--inside	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Display--outside	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Pen Responsiveness	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Pen Feel	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Handwriting	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5
Comfort	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5

Which product do you prefer?

Computer #1 Computer #2 Computer #3 Computer #4

Which orientation do you prefer? Horizontal Vertical

Do you prefer to have the pen tethered to the unit? Yes No

Do you think you could carry any of these units for a normal work day? Yes No

If a neck, shoulder, or waist strap were available, would you use it? Yes No

Which would you prefer? Neck Shoulder Waist

Would you prefer a rugged unit to one that is less rugged, even though it weighs more? Yes No

What are the three largest drawbacks to all of these products? 1. ____________________

2. ____________________

3. ____________________

Would you prefer using a smaller, more lightweight product (eg. less than 4 lbs.) with fewer software tools (eg. no word processors, spreadsheets, etc.) available or a heaver product with more tools?

Lightweight, Few tools　　　　Heavy, More tools

Would you prefer using a smaller, more lightweight product (eg. less than 4 lbs.) if software tools available on the pen computer (eg. word processors, etc.) were very different from those used in the office or a heavier product that used the same software as used in the office?

Lightweight, Different tools　　　　Heavy, Same tools

Would you prefer using a smaller, more lightweight product (eg. less than 4 lbs.) or a heavier product with a built in keyboard?

Lightweight, No keyboard　　　　Heavy, Built in keyboard

Would you prefer using a smaller, more lightweight product (eg. less than 4 lbs.) with an operating environment (eg. DOS) that was different from the office computer operating environment (eg. Windows) or a heavier unit with the same operating system?

Lightweight, Different environment　　　　Heavy, Same environment

Would you prefer using a product that was light, but not rugged, or a heavy, rugged product?

Lightweight, Fragile　　　　Heavy, Rugged

Would you prefer a standard laptop computer, without a pen (ie. no handwriting input), to the pen computers?

Yes　　　　No

Appendix D, Summary Minimum Hardware Specifications

Currently, two different pen computer specifications are appropriate; one specification describes what is available today, while the other describes what may be cost effective in the future.

Specification for available equipment:

Pointing device (eg. pen, mouse, trackball)
Keyboard input device (either affixed, detachable, or "virtual")
Storage device (eg. 40 Megabyte or greater capacity hard disk, PCMCIA card)
20-25 Megahertz, 80386 CPU
Display device: grey scale, backlit, VGA
Adjustable screen brightness and/or contrast
8 Megabytes RAM, minimum
Serial, Parallel communications ports (allows network connection, for example)
Replaceable, rechargeable batteries with at least one hour of operational capability without power-saving options in effect
Battery charger/power supply
Weight less than 8 lbs.
Floppy disk drive (external)
External CD ROM drive

Optional:
FAX/Modem
Docking station--should include: card cage, keyboard, floppy drives; may also include: color Super VGA monitor, large capacity hard disk, tape backup drive
External monitor connector
Additional PCMCIA slots
Ruggedized carrying case/strap

Specification for equipment available within 1-2 years:

Pen pointing device
Attached keyboard input device that accommodates either a horizontal or a vertical display orientation
100 Megabyte or greater capacity hard disk
2 or more PCMCIA card slots
33-50 Megahertz, 80486 or 80586 CPU
Color Super VGA monitor
Adjustable screen brightness and contrast
16 Megabytes or more RAM
Serial, Parallel communications ports on the unit
Replaceable, rechargeable batteries with at least three hours of operational capability without power-saving options in effect
Battery charger/power supply

Weight less than 3 lbs.
Bar code and/or magnetic strip reader
Internal floppy disk drive
CD-ROM drive (possibly internal)
FAX/Modem
Wireless LAN ("WaveLAN")
Internal card slot (eg. for a network card)
Docking station--including, but not limited to: card cage, keyboard, floppy drives, color Super VGA monitor, large capacity hard disk, tape backup drive
External monitor connector
Ruggedized carrying case/strap

Chapter Four
Correlates of Individual Differences in Nondestructive Inspection Performance

4.0 INTRODUCTION

This chapter is divided into four sections. The first section, Background and Survey of Relevant NDI Research, discusses nondestructive testing and damage-tolerance design and research programs on NDI capabilities conducted by the Air Force, the nuclear power industry and the FAA. The General Survey of Inspection and Vigilance Research section reviews research related to individual difference variables in inspection and vigilance. The remaining two sections, Research Needs and Proposed Research, outline the direction and methods of the NDI performance research to be performed under the current FAA/AAM contract.

4.1 BACKGROUND AND SURVEY OF RELEVANT NDI RESEARCH

4.1.1 Nondestructive Testing and Damage-tolerance Design

According to Panhuise (1989), in most industries, the inspection requirements for various components are defined in a specification document that describes the sensitivity level of the inspection method as well as the rejectable flaw size. These inspection requirements, designed to control both the inspection process and the quality of the inspection results, define the detection/rejection requirements for each quality class material, the required procedures for meeting these requirements, and the required level of inspector training to meet requirements. However, several major, catastrophic failures of engineering systems (e.g., the F-111, space shuttle, nuclear reactors) led to the development of a new design method that assumes the existence of structural defects and then allows the designer to answer the following questions:

- What is the critical flaw size that will cause failure for a given component subject to service stress and temperature conditions?
- How long can a precracked structure be safely operated in service?
- How can a structure be designed to prevent catastrophic failure from preexisting cracks?
- What inspections must be performed to prevent catastrophic failure?

The answer to these questions forms the basis of nondestructive inspection (NDI), which involves damage-tolerance design and is centered on the philosophy of ensuring safe operation in the presence of flaws. Damage-tolerance design assumes that flaws exist as part of the normal manufacturing process. Flaw size is predicted to grow as a function of service usage and, in the case of aircraft, will reach a critical size after a certain number of flight hours. Damage-tolerance design further assumes that nondestructive evaluations, performed on a periodic basis, will be able to detect such cracks or flaws before they reach this critical size. The validity of this assumption was initially evaluated in the aerospace industry during the 1970's with findings that were profoundly disturbing, as described below.

4.1.2 Air Force Research

4.1.2.1 Reliability of Nondestructive Inspections (1978 Study)

The first, and most comprehensive field evaluation of NDI capabilities in the aerospace industry to date, was conducted by Lockheed in the 1970's in a major Air Force study that has since become known as the "Have Cracks, Will Travel" program (Lewis et al., 1978). This study, which began in 1974 and was completed in 1978, was designed to answer the following questions:

- What is the relative effectiveness of conventional NDI methods applied to structure (i.e., flaw detection probabilities relative to radiographic, ultrasonic, eddy-current, and penetrant inspections)?
- What is the Air Force field and depot capability in NDI? More specifically, what are the probabilities of flaw detection in structures by Air Force personnel and equipment?
- What differences, if any, exist in NDI capabilities from base to base?
- How effective are 7 Level Air Force NDI personnel in devising NDI procedures?
- What is the range of individual capabilities among all groups (all bases) and within each group (base)? In other words, what is the scatter factor attributed to individual differences and to differences between bases?

The approach taken was to obtain representative samples of six different types of aircraft structure with fatigue cracks ranging from .010 to 1.05 inches. These structure samples were presented to the NDI technicians in settings which closely approximated those encountered in routine field and depot operations. Some were placed in an overhead position to simulate NDI on a lower wing surface, others were in face-up and vertical orientations. Twenty-two facilities were involved in the study, with an average of 15 participants tested at each of the Air Logistics Centers and 6 at each field level base. Data obtained for each technician consisted of (a) the number of finds (hits), (b) the number of false calls, (c) the number of no-finds (misses), (d) the ratio of finds-to-total flaw count, and (e) that ratio in percent. Also obtained were (a) the total finds ratioed to the total number of technicians and (b) that ratio in percent. Each sample tested by a technician contained flaws of various sizes; no samples were included that contained no flaws nor were there any samples included in which all sites were flawed.

Military specification, MIL-A-83444 (USAF), "Airplane Damage Tolerance Requirements," states that in-service inspections are assumed capable of detecting cracks of specified lengths. A major finding of this study was the realization that the previously established 90-95 percent reliability criteria (90 percent probability of detection with a 95 percent confidence bound) for a .25 inch crack was not obtainable under normal field inspection. With the exception of dye penetrant inspection, the results indicated considerable difficulty in achieving a <u>50 percent</u> probability of detection for a <u>1/2 inch</u> crack size with a 95 percent confidence level.

There were no significant differences (with the exception of one depot) in NDI performance between individual installations, between individual Commands, or between field installations and depots. Nor were any significant differences observed between technicians using different manufacturers' equipment.

For purposes of the present paper, the most significant findings were with regard to individual differences between inspectors. While it comes as no surprise that the study found considerable differences in the

NDI proficiencies of individual technicians, the lack of any apparent relationship of such obvious variables as training and years of experience to performance is surprising. There was essentially no relationship between performance and any of the following variables:

- Skill level: Air force technicians at three skill levels (levels 3, 5, and 7, representing new entrants into NDI, the majority of practicing technicians, and advanced technicians, respectively) were compared with regard to their proficiency using each NDI method and on various types of structural samples. Although some positive relationships appeared to exist between skill level and performance, there were numerous instances of no apparent relationship. The authors were led to conclude that technician proficiency in crack detection ability was not significantly related to skill level.
- Formal education: High school graduates, when compared to those lacking a high school diploma, did not differ in NDI proficiency.
- Age: The performance of technicians under 25 years of age was contrasted with performance of those 40 years or older. As with skill level, the findings were mixed, with some comparisons showing a relationship and others not. The general conclusion, however, was that age was not systematically related to NDI proficiency.
- Years of NDI experience: As with age, two experience groups were compared: technicians with less than three years NDI experience were compared with those having more than ten years experience. Although age was generally confounded with experience, there was no evidence that experience per se was significantly related to NDI proficiency.
- NDI training: Hours of formal NDI training, as reported by the technicians, were examined for effect on performance by contrasting those with under 200 hours with those having over 500 hours. As with the previous comparisons, the amount of NDI training was found to be unrelated to proficiency.

To summarize the results of this study, the dominant finding was the failure to relate NDI proficiency to any of the personnel variables. However, the study was not without deficiencies, the most important of which was recognized by the authors. This was the failure to make use of any of the data collected on "false calls" in their analyses. Consequently, a technician could theoretically call every site visited a "flaw" and achieve 100% detection. The authors concluded, however, that, while some technicians showed extremely high false call counts and were therefore suspect, those instances of extremely high false call count were not numerous enough to cause the total data and/or findings to be suspect.

4.1.2.2 1978 NDI Reliability Workshop

Concomitant with the release of the "Have Cracks, Will Travel" study, a Lockheed-sponsored workshop was held to present the study's findings (along with some additional material not contained in the study report), and to solicit comments and recommendations from the workshop's participants (Lewis, Pless, and Sproat, 1978). Several aspects of this meeting relative to personnel considerations are worth mentioning:

- As a subsidiary aspect of the "Have Cracks" study, a technician selection/screening method was developed to determine whether simple flat plate test panels containing cracks could be used to assess proficiency, and hence predict technician performance on the study's structural test samples. Twenty-six technicians participated with mixed results. The job sample task predicted ultrasonic performance with reasonable accuracy, but failed to predict eddy-current performance. In spite of these mixed findings, the use of a job sample approach was viewed as having considerable potential for predicting technician performance.
- Although not mentioned in the results of the primary study, data were presented at the workshop showing eddy-current performance plotted against the number of eddy-current inspections performed per month. While the trend was weak, there was some evidence that poor performance tends to disappear with moderate frequency of inspections (more than ten times per month).
- The task group charged with developing selection criteria/recommendations devised the following profile for the ideal inspector. He or she should (a) have integrity, (b) have a sense of responsibility, (c) be mature enough to recognize his/her responsibilities, (d) take pride in accomplishments, (e) be self-motivated, (f) be self-disciplined, (g) have moral convictions, (h) have allegiance, and (i) be good at decision-making ability. Obviously, this profile would appear to characterize those traits desirable in employees in virtually any job. However, either stated or implied in this list are several qualities that have surfaced repeatedly in talks with inspectors, managers, and NDI instructors. Desirable qualities stated generally relate to personal integrity, motivation or interest in the job, and the ability to make decisions.

4.1.2.3 The Technician Proficiency Measurement Program

The high degree of variability in technician proficiency found in the "Have Cracks, Will Travel" study led the Air Force to consider possible forms of corrective action. As noted earlier, one of the subsidiary studies carried out under the "Have Cracks" project involved the assessment of a job sample test to predict technician performance. Although the findings were somewhat mixed, the approach was considered sufficiently promising to warrant further consideration. The result was the development of the "Air Force Technician Proficiency Measurement Program." This program was one in which practical tests involving nondestructive inspection of flawed aircraft structures, called test racks, were administered to technicians. The test racks, fabricated by the Lockheed-Georgia company, were made up of several specimen plates with simulated fatigue cracks of various sizes at randomly selected fastener sites. Technicians participating in the program were required to perform eddy-current or ultrasonic tests on the samples, with the resulting data scored in terms of hits, misses, false alarms, and true negatives.

Background for the development of this program, as well as a detailed description of the tests and procedures used, can be found in a paper by Boisvert, Lewis and Sproat (1981). Jayachandran and Larson (1983) have reported on the use of the proficiency measurement test in a study of 360 technicians distributed over 17 Air Force bases and 6 Air Force commands. Unfortunately, the study was concerned mainly with different methods for analyzing the data, and no attempts to assess the test's usefulness for either selection or training are reported. As with the original "Have Cracks" study, results of the job sample test again revealed a wide range of individual differences among technicians in NDI proficiency.

A subsequent study by Summers (1984) examined relationships between personnel information, as obtained from questionnaire items and performance (both ultrasonic and eddy-current) on the technician proficiency measurement test. NDI proficiency measures were obtained on 205 Air Force test participants (125 inspectors and 79 supervisors). (It is not clear as to whether this sample was taken from the study reported above by Jayachandran and Larson or whether it was an entirely different sample.) Of the respondents:

- 63% were military; 37% civilian.
- 87% were male; 13% female.
- 14% had less than a high school education.
- 50% had completed high school only.
- 25% had as much as 2 years of college.
- 11% had completed more than 2 years of college.

Answers to questionnaire items were tabulated and related to performance on the ultrasonic and eddy-current job sample tests. The following relationships were obtained.

<u>Negative findings:</u>

- Amount of formal schooling was not significantly related to job sample performance.
- Neither eddy-current nor ultrasonic test data showed a relationship to the amount of NDI training (Air Force or civilian).
- There was no indication of a relationship between performance and whether or not a technician was a volunteer for the NDI career field.
- Previous experience in metal working prior to NDI training was unrelated to NDI performance.
- No significant relationship was found between inspector performance and the degree of like/dislike for present job or for the NDI career field.
- Although self ratings of ability on eddy-current and ultrasonic performance were significantly correlated (r=.67), actual job sample performance was unrelated to an individual's self-rated ability.
- Performance on eddy-current and ultrasonic inspections was not related to the degree of comfort/discomfort technicians felt with equipment used in the inspection tests.
- Performance on the inspection tasks was not significantly related to local on-the-job training or existing resident NDI training.
- Neither amount of time spent on NDI tasks (in their normal job) nor time spent on individual NDI techniques (also relative to their present job) was related to ability to find flaws in the job sample test.
- Supervisor ratings of technician proficiency correlated no better with inspection test performance than did the technicians' self ratings.
- There was large variability in eddy-current and ultrasonic inspection performance across the sample -- among technicians and across bases and commands. In general, inspection results were too inconsistent for maintenance managers to have confidence in NDI capability.

Positive findings:

- There was a slight, but significant tendency for technicians with more than 2 years of college to have fewer false calls.
- The 17% of respondents who stated they were certified by the American Society for Nondestructive Testers performed somewhat better in making finds than did noncertified inspectors on both eddy-current and ultrasonic tests.
- Technicians who indicated an intention to re-enlist in the Air Force scored significantly higher in both hits and false calls than did those who had no such intention. (This suggests a motivational component that manifests itself as an increase in the frequency of positive responses.)

It is evident that the Summers study failed to identify clear-cut individual difference variables that correlate with inspection performance. In this respect, the study's conclusions did not differ appreciably from those of the earlier "Have Cracks, Will Travel" study. Neither study found any strong, consistent relationships between personnel variables and inspection performance. Unfortunately, the Summers study was essentially a summary final report and failed to provide much information on either the personnel data obtained or on the methods used in scoring inspection performance. Consequently, as with the earlier 1978 study, it is difficult to critique or evaluate the findings.

4.1.2.4 Reliability of NDI Applied to Aircraft Engine Components

The previously described "Have Cracks, Will Travel" study was directed toward an assessment of NDI reliability on aircraft structures under field conditions. However, it provided no information on the reliability or proficiency of Air Force inspectors performing NDI on engine components, leaving unanswered the question of whether the same levels of variability among inspectors found by Lewis et al. (1978) also applied to inspectors in engine overhaul facilities. Consequently, an Air Force sponsored study was conducted by Lockheed Aircraft to answer this question (Rummel et al., 1984). The approach was similar to that used by Lewis et al.: Sets of gas turbine blades, vanes and disks were constructed such that half the items in each sample set were nonflawed and half were flawed items, with flaws ranging in length from .01 to 0.5 inches. Inspectors at two Air Force logistic centers employed dye penetrant, magnetic particle, ultrasonic, and eddy-current methods to test for flaws.

The general findings of this study were similar to those found by Lewis et al.: (a) the overall reliability of nondestructive inspections used by the Air Force in engine overhaul was found to be below that which had been assumed or generally desired, and (b) variations in technician proficiency were observed and documented. With regard to this latter finding, however, the magnitude of the differences between technicians did not appear as great as that obtained by Lewis et al., and those differences that were observed seemed to be generally attributable to recency of training/experience. Consequently, no systematic attempts were made to relate technician proficiency to various subject or personnel variables.

4.1.2.5 Recommendations for Improving Air Force NDI Reliability

In 1987, Southwest Research Institute (SwRI) was contracted by the Air Force to investigate what might be done to improve technician proficiency. The approach taken in the resulting study (Schroeder, Dunavant, and Godwin, 1988) used NDI experts to (a) identify relevant areas of concern that could negatively impact the proficiency of Air Force NDI technicians, (b) seek possible solutions for the identified concerns, and (c) incorporate the most feasible, promising, and cost-effective potential solutions into recommendations for improving technician proficiency. Since the present review is concerned only with individual difference variables and their relationship to NDI performance, only concerns and possible solutions related to this topic will be reviewed here. From the numerous concerns raised by the experts, three were considered most relevant. These, along with proposed solutions, are given below:

A. Concern: Although NDI is a highly technical area, the Air Force has no intentional selection mechanism. Technicians come from the general manpower pool.

   Solution: As a short-term partial solution, it was suggested that samples be selected from the population who are higher in electronic and mechanical abilities, and then measure and compare their performance with personnel selected using the current approach.

B. Concern: There are a number of candidate selection variables proposed in the technical literature (e.g., ability to concentrate, patience, manual dexterity, intelligence, temperament, motivation), but virtually no systematic research into which variables predict good NDI technicians.

   Solution: (a) Sponsor and conduct research to establish predictors of proficient NDI personnel, and (b) analyze any new data using receiver operator characteristic (ROC) based measures of proficiency.

C. Concern: Much of the relatively little research that has been done is not meaningful, since measures of the predicted variable (proficiency) were not adequate.

   Solution: Analyze any new data using receiver operator characteristic (ROC) based measures of proficiency.

In a recent phone conversation with the principal author of this study, it was learned that, while the issues, concerns, and suggested solutions were received with considerable interest and enthusiasm by the Air Force, to the best of his knowledge, no research programs have been funded to implement the recommended solutions (J. E. Schroeder, personal communication, September 1, 1992). Schroeder stated that the problems the Air Force was having with differences in proficiency level of enlisted technicians has, to a large extent, been circumvented by hiring civilian technicians in their place. While this action may increase the overall level of inspection reliability, it is obviously a way of simply avoiding the more difficult problem of ascertaining reasons for differences among Air Force technicians, and then developing the necessary selection and/or training procedures to improve technician proficiency.

4.1.3 Nuclear Power Industry Research

Apart from the Air Force, the only other major organization to carry on a systematic research program in the human factors of nondestructive inspection would appear to be the nuclear power industry. As noted in a report by Triggs et al. (1986), NDI inspection in the nuclear power industry suffers from many of the same problems found in the Air Force studies. Although confining themselves primarily to ultrasonic inspection, they note that relatively high error rates in flaw detection are commonly obtained. In one of the reported studies, probability of detection for 30 flaws ranged from 0.0 to 1.0, with a mean of 0.37. Institution of new procedures resulted in a decrease in flaws missed. However, even under the best possible conditions, 34% of the flaws were still missed. In a second reported study, an NRC (Nuclear Regulatory Commission) analysis of six teams, who inspected 80 circumferential pipe welds requiring 1,500 operator judgements, found wide differences among teams and conditions of performance and a wide range of success rates. (No actual reference is provided for either of the above two studies cited by Triggs et al.)

Human factors research on variables related to NDI proficiency has not been much greater in the nuclear industry than in the Air Force. That which has been done has been largely conducted under the aegis of the Electric Power Research Institute (EPRI) located in Palo Alto, California. Under contractual support from EPRI, Harris has conducted several recent studies related to human factors aspects of NDI. In one of his more recent studies, information processing factors involved in ultrasonic flaw detection were investigated (Harris, 1990). Inspectors were hypothesized to employ some or all of the following factors in assessing signal characteristics: (a) Explicit hypothesis, (b) Test of explicit hypothesis, (c) Early conclusion, (d) Disregard of evidence, (e) If-then logic, (f) Explicit signal discrimination, (g) Identification of weld geometry, (h) Verification of signal, and (i) Recognition of a malfunction or abnormality. A stepwise multiple-regression analysis revealed that most of the predictive variability was contributed by five of the above factors -- early conclusion, test of an explicit hypothesis, if-then logic, disregard of evidence, and signal continuity. It was concluded that an inspection approach based on a well-defined information-processing strategy offered promise for improving inspection performance.

The study makes no reference to individual differences among inspectors. Differences are at least acknowledged, however, in a study of eddy-current inspection of steam generator tubes used in nuclear heat exchangers (Harris, 1991). Inspections of the test samples by experienced analysts showed considerable variation between inspectors in accuracy of detection. The large differences in analyst performance were not predicted by the qualification testing conducted in accordance with existing guidelines and current industry practice. The correlation between qualification test scores and percentages of indications correctly reported was only 0.17, not significantly different from zero. Beyond acknowledging this lack of relationship of qualification test scores to inspection performance, the study offered no further analysis of the obtained differences.

In another study of NDI personnel in nuclear power plants, an attempt was made to evaluate the characteristics of the most proficient inspectors (Beharavesh et al., 1988). Interviews were held with 57 persons involved, in one capacity or another, with ultrasonic inspection -- technicians, training supervisors, and vendor personnel. Characteristics of highly competent technicians, as determined from the frequency of characteristics mentioned in interviews were:

- Can handle pressure/stress
- Is conscientious/reliable/dedicated
- Is independent/autonomous/self-confident
- Is knowledgeable/skillful/experienced
- Is able to work well with others
- Is mentally and emotionally stable
- Has good attentional/perceptual/motor skills

In referring to the above list of characteristics, Harris (1988) notes that "these are the general characteristics of people who are the most competent in any job--airline pilot, assembler, police officer, taxi driver, football player, carpenter, computer programmer, rodeo clown, and others" and states that such characteristics are not sufficiently unique to serve as the basis for selecting persons who will become competent ultrasonic inspectors. The reader may remember that a similar list of characteristics, referred to earlier, was compiled by participants at the 1978 Air Force workshop on NDI reliability.

4.1.4 Federal Aviation Administration (FAA) Research

4.1.4.1 The State University of New York (SUNY) Program

Acting in response to the Aviation Safety Research Act of 1988, Public Law 100-951, the FAA's Office of Aviation Medicine has established a human factors research program to investigate aircraft maintenance and inspection practices, especially as these are applied to aging fleet repair, and to evaluate and recommend improvements. The current program, carried out largely under contract with Galaxy Scientific Corporation, is conducting research to improve practices in several related areas. These include, but are not limited to, the maintenance organization itself, maintenance inspection, advanced technology for training, and job aids. For purposes of the present paper, only that aspect of the total program dealing with inspection will be considered here. This presentation will be further narrowed to concentrate on that portion of the maintenance inspection research program dealing with task simulation of NDI.

As a subcontractor for Galaxy Scientific, Dr. Colin Drury at the State University at New York at Buffalo is conducting a substantial research program in aircraft inspection. In essence, work thus far has largely concentrated on development of task descriptions, task analyses, and a detailed error taxonomy. In addition, however, Dr. Drury has developed a simulated NDI task, using a SUN workstation, that incorporates the physical aspects and functional characteristics of an eddy-current NDI task. As was stated in the initial research study using this simulation, the task was not developed with the intention of measuring absolute values of the probability of detecting particular types and sizes of flaws, nor was it developed as a means of training inspectors for the actual tasks involved (Latorella et al., 1992). The intended use of the task, then, was and is to explore variables related to inspection performance and to isolate those that might have potential relevance to the operational environment. (This simulated NDI task will be described in greater detail in a later section of this paper as well as its intended use in proposed studies.)

A basic intent of the NDI study mentioned above was to evaluate the use of off-line performance feedback. Also included, however, were two personality tests previously shown to correlate with inspection performance. These were the Embedded Figures Test (a measure of field dependency) and the Matching Familiar Figures Test (a cognitive style measure of speed vs. accuracy in performance). The study failed to demonstrate improved performance as a result of off-line feedback, presumably because of the large between-subject variability. Interestingly, one of the two covariates based on the personality tests did show a relationship to performance. Thus, the covariate based on a composite index comprised of Matching Familiar Figures Test and visual acuity scores, was significantly related to both total task time and to the decision criterion used.

Drury clearly feels, with some justification, that intensive investigation of individual difference variables as possible correlates of inspection performance is likely to have a rather low probability of success (Drury, 1992). However, he has also noted that continued study of individual differences in aircraft inspection should not be disparaged, because the payoff for establishing a reliable and valid inspection test would be large (Shepherd et al., 1991). The sizable between-subject variability found in the above study by Latorella et al. (1992) is certainly consistent with Air Force and nuclear industry studies reviewed earlier, and the finding of a relationship between cognitive style and NDI performance supports the present author's belief that at least some of the variance in inspection performance is related to individual difference variables of potential use in selection. [This variable (cognitive style) and others will be considered latter in a section dealing with laboratory and field studies of individual difference variables and general inspection performance.]

4.1.4.2 The Sandia Program

Another research program in NDI has been funded through the Aviation Safety Division of the FAA Technical Center. One aspect of this program is the establishment of an Aging Aircraft NDI Development and Demonstration Center (AANC) at the Sandia Corporation. The essential purpose of this center is to support NDI technology, technology assessment, technology validation, data correlation, and automation adaptation as on-going processes. A second aspect of the program is to determine how well current equipment and procedures used in the field detect structural flaws (Spencer et al., 1992a). Only the field program will be briefly discussed here.

The stated objective of the field research study is to evaluate the reliability of eddy-current inspection procedures as they are done routinely at airline maintenance and inspection facilities (Spencer et al., 1992b). As described in this report, the planned experiment will be specific to inspection procedures used on Boeing 737 lap splice joints with sliding probe, oversize template, and rotating surface probe NDI eddy-current techniques. Panels of test samples will be developed with each panel containing differing frequencies of flaw length and density. Test samples will be evaluated by inspectors at different facilities and during different shifts. In evaluating factors that could affect reliability, the specific objectives of the study are to: (a) Assess Effects of Off-angle Cracks; (b) Assess the Effect of Inspecting Painted Versus Unpainted Surfaces; (c) Characterize the Reference Standards Used Within a Facility; (d) Assess Effects of Accessibility; (f) Access Inspection Time Effects; (g) Gather Facility Specific and Inspector Specific Data as Potential Explanatory Factors; (h) Provide Baseline (Laboratory Environment) Inspection Reliability Assessments; (i) Assess Effects Connected with Shift Work; and (j) Assess the Effect of Specimen Definition. This latter factor refers to assessing possible differences between test results conducted on "real" flaws and those generated by various artificial means.

It is anticipated that nine facilities will be visited, representing a range of facility characteristics. Data obtained from four inspectors at each facility will be analyzed in terms of probability of detection (POD) measures and measures derived from receiver operating characteristic (ROC) curves. Personnel data obtained on each inspector will include age, sex, physical condition, NDI experience, time since last performed eddy-current testing, amount of equipment-specific training, and perceived importance of NDI to management, as well as to the individual inspector. Each facility will be rated on lighting, temperature, atmospheric conditions, management practices, tools, general housekeeping practices, and noise level. Facility differences in procedures, equipment, training, and environment are thus considered to be part of the system being considered and will be analyzed as potential explanatory factors for observed variation in inspection results.

The Sandia study, then, is obviously quite different from the SUNY program in purpose and scope. However, it is surprisingly similar, both in purpose and methodology, to the previously described Air

Force "Have Cracks, Will Travel" study. The major difference between the two studies would appear to be that the Sandia study plans to incorporate false alarm data in their assessment of reliability while the earlier Air Force study failed to do so. Many of the same individual difference variables (e.g., age, years experience, level of NDI training) incorporated in Air Force studies (Lewis et al., 1978; Summers, 1984) are also included in the Sandia study. Like the earlier Air Force studies, the Sandia study will not attempt to experimentally control for any of these personnel variables, but rather will depend on correlational analyses to reveal possible relationships of each variable to NDI performance. The value of such an approach is, of course, dependent upon the extent of attribute variation in the samples that are available.

4.2 GENERAL SURVEY OF INSPECTION AND VIGILANCE RESEARCH

4.2.1 Inspection

The single most consistent finding of the Air Force and nuclear power industry NDI research programs and studies reviewed thus far has been the finding of sizeable and consistent individual differences among inspectors. Perhaps not surprising, this was also the most consistent finding reported by Wiener (1975) in his review of individual difference variables in inspection research carried out in university laboratories and in industrial settings. Unfortunately, like the Air Force and nuclear power plant studies, Wiener found little evidence of a consistent relationship of various individual difference (selection) measures to inspector performance. However, most of the studies reviewed were conducted in what Wiener refers to as the "pre-ergonomics era" (during and prior to World War II). These studies frequently used such questionable measures as supervisor's ratings as criteria. This was particularly true of the early studies employing aptitude tests. Of the aptitude tests employed, none was tailor-made for predicting inspection performance. The only aptitude test that has apparently been devised specifically for inspection is the Harris Inspection Test (HIT), a paper-and-pencil test that can be administered in 10-20 minutes (Harris and Chaney, 1969). As reported by Harris and Chaney, the test was successfully validated on six inspection tasks. Unfortunately, Wiener (1975) reports that a later study by Chaney and Harris found test results to be unrelated to inspection performance. (More recent use of this test, as well as others, will be given shortly when other studies by Drury and his colleagues are considered.)

Apart from visual tests, such as acuity which has obvious relevance to visual inspection, Wiener confines the remainder of his review to personality, gender, age, and intelligence as possible factors related to inspection proficiency. With the exception of age, Wiener found little or no research had been conducted that was specifically directed at the relationship of personality measures, gender, or intelligence to inspection performance. Age effects were found to have been studied by various investigators, with some evidence that inspection proficiency declines with age. However, conflicting findings led Wiener to conclude that any age-related differences in inspection are likely be small and of minimal significance. (It will be recalled that a similar finding was reported in the "Have Cracks, Will Travel" study.)

4.2.2 Vigilance

Vigilance research has often been considered in conjunction with inspection findings because the two areas have much in common. Both frequently involve sustained attention, decision making, and may involve visual search and scanning. Interestingly, both inspection and vigilance are characterized by sizeable individual differences and relatively consistent within-subject performance over time. Researchers attempting to account for individual differences in vigilance performance have often been as frustrated as those working in the area of inspection.

It is for many of these reasons that reviews of individual difference variables in inspection research, such as the one by Wiener (1975) just considered, generally include vigilance findings as well. Thus, Wiener reviews studies that have examined the relationship of gender, age, intelligence, and personality variables to vigilance performance. In general, the findings have been similar to those of inspection research, with gender, age, and intelligence showing either an inconsistent or lack of relationship to vigilance performance. Personality variables, studied within the context of vigilance have, however, been somewhat more successful in predicting performance. This is particularly true of the introversion-extroversion dimension, where introverts are hypothesized to perform better than extroverts (Eysenck, 1967). Although few of the studies reviewed by Wiener show a clear-cut superiority of introverts over extroverts, none show the opposite. [A recent review of extroversion and vigilance using a meta-analysis of studies covering a 30-year period generally supports the belief that introverts are superior in vigilance performance, but the effect size was found to be quite small because of a high incidence of inconsistencies (Koelega, 1992).]

A later review of individual differences in sustained attention or vigilance extends the earlier findings of Wiener and focuses more directly on personality variables (Berch and Kanter, 1984). The introversion-extroversion dimension is again shown to be rather consistently related to monitoring, although admittedly most of the studies reviewed were from the same time period encompassed by the Wiener review. Two additional dimensions not included in the Wiener paper were field dependence/independence and locus of control. Berch and Kanter cite examples of studies showing both field independence and internal locus of control to be related to superior vigilance performance. With regard to age, gender, and intelligence, these reviewers are in accord with Wiener in that there is little evidence to support a relationship of either gender or intelligence to monitoring performance. Some studies showed a relationship of age to monitoring performance, but perhaps only when certain conditions prevailed. These conditions are believed by Davis and Parasuraman (1982) to occur when: (a) detection of more than one signal is required; (b) the event rate is high; (c) visual search is involved; (d) an increased memory load is required for reporting or discriminating the critical signal. A number of these conditions were present in a study of complex monitoring performance comparing young, middle-aged, and older subjects (Thackray and Touchstone, 1981). Both the onset of attentional decline as well as its magnitude were found directly related to age.

The preceding, rather cursory examination of individual differences in both inspection and vigilance was intended to emphasize both the prevalence of wide subject variability found in the two related areas of research and to highlight the fact that neither area has been too successful in finding significant correlates of this variability. Within the past 10 years, however, several studies have been carried out to clarify reasons why the selection approach has thus far met with minimal success (Gallwey, 1982; Wang and Drury, 1989). The initial study by Gallwey (1982), based partially on an earlier task analysis of a visual inspection task by Drury (1975), used a variety of selection tests to predict inspection performance on a computer-generated symbol task containing multiple fault types. The task was designed to simulate a typical industrial inspection task containing elements of visual search, memory, judgement, and decision. The tests were chosen to tap different subtasks of the primary task. In general, the subtasks involved scanning an area to select a fixation region, examination of items within the region, comparisons with images in memory, and a decision to accept or reject item(s). The tests used included a measure of visual acuity, the Harris Inspection Test, the Eysenck Personality Inventory, a questionnaire on mental imagery, a card sorting task, two portions (the Embedded Figures Test and a composite test consisting of the Arithmetic, Digit Span, and Digit Symbol subtests) of the Weschler Adult Intelligence Scale (WAIS), the Embedded Figures Test, a measure of short-term memory, a measure of visual lobe size, and a measure based on part-task performance (single fault detection) of the primary task. Some of the more important findings were as follows:

- Single fault detection time was a very good predictor of detection time on the primary or multiple fault task.
- The WAIS composite score (a measure of attention-concentration) was effective in predicting probability of search and classification errors.
- Extroversion scores obtained from the Eysenck Test correlated significantly with search errors, with low scores on extroversion related to fewer errors.
- The Embedded Figures Test was found to be the best single predictor of inspection performance, being related to search time, search errors, and decision errors.
- Visual lobe size was a reasonable predictor of classification errors.

A second study comparing a variety of selection measures with inspection performance was reported by Wang and Drury, (1989). Based partially on some of the findings of the Gallwey (1982) study, these authors hypothesized that the skills and abilities of inspectors may indeed be task specific. If this is the case, and different inspection tasks require different skills/abilities, then the search for a general "inspection type" or a single selection task could prove to be a futile exercise. Thus, a more fruitful approach might be to select only those tests expected to correlate with the particular skills/abilities of a specific inspection task. In order to identify those skills/abilities, Wang and Drury first provided a seven-step task description of a generalized inspection task. These steps are:

(1) Orient the item to be inspected.
(2) Search the item.
(3) Detect a flaw or unusual phenomenon.
(4) Recognize/classify the phenomenon.
(5) Decide on status of item.
(6) Dispatch item to appropriate destination.
(7) Record information pertaining to the item.

Each of the above steps was then analyzed in terms of the skills (manual or perceptual) and abilities (attention, perception, memory, detection, recognition, judgement, classification) required. Based on the generalized task description, 11 tests, selected as potential measures of each of the above abilities, were incorporated as pretests in the study. Three different inspection tasks were chosen to represent different types of inspection. These were: (a) Circuit Pattern inspection, a pure search task; (b) Computer Generated Symbols inspection task, used previously by Gallwey (1982) and which incorporates both search and decision; and (c) Color Video Comparator inspection of printed circuit boards, representing a real inspection task currently used by electronics manufacturers. Each of 12 subjects performed all three tasks.

Separate factor analyses of pretest and performance scores revealed four pretest factors (labeled as Attention, Perception, Judgement, and Memory) and five performance factors (generally encompassing various measures of search time, search errors, and decision errors). Pearson correlations were also computed between the four pretest and five performance factors. Several significant relationships were obtained:

- The Attention factor was significantly correlated with search error and search time (the higher the attention score the fewer the search errors).
- The Judgement factor was significantly correlated with decision error (the higher the judgement score the fewer the decision errors).
- Time and error scores fell into different groupings, showing that speed and accuracy represent different aspects of performance
- Correlations for each of the three inspection tasks tended to cluster together within particular factors
- Tests loading on the Perception factor correlated with speed of visual search, but only on the computer-generated symbols task.

Although significant relationships were obtained, the patterns were far from clear. For example, none of the pretests demonstrated consistent predictive ability of search performance across the three inspection tasks. Thus, a perceptual or attention-concentration test that predicts search performance well on one task may not be a valid predictor on another. In summarizing their results, the authors conclude that:

> the best strategy in developing a valid inspection selection device would be to find out the specific mental requirements for the particular task by conducting a detailed task analysis, as well as by eliciting information from experienced inspectors. Then, based on these mental requirements, select a set of valid test items which can effectively measure those cognitive traits and, thus, produce one's own version of an inspection selection battery. (Wang and Drury, 1989, p. 189)

4.3 RESEARCH NEEDS

It was noted, in the 1988 Southwest Research Institute study of recommendations for improving Air Force NDI technician proficiency, that no research had been carried out with the specific intent of studying individual difference variables in NDI performance (Schroeder et al., 1988). Since the time of this 1988 study, there has apparently been only one study conducted that has at least examined a few individual difference variables in NDI performance. That study is the study referred to earlier by Latorella et al. (1992) in which two psychometric tests, the Embedded Figures Test and the Matching Familiar Figures Test, were correlated with performance on the SUNY simulated NDI eddy-current task. The finding that one of these tests, the Matching Familiar Figures Test, was significantly related to several performance measures suggests that at least some of the subject variance in performance can be accounted for, and that a more concerted effort, using tests covering a wider range of abilities, is warranted. Although research efforts have been rather unsuccessful thus far in devising predictors of general inspection or vigilance performance, this may be, as was indicated in the above quote from the Wang and Drury paper, that predictor measures are at least partially task specific. Consequently, and as these authors suggest, the most promising approach may be to select tests based on a detailed analysis of task behaviors for the task in mind, and thus produce a selection battery more likely to correlate with performance on the intended task. It is the intent of this research project to utilize this approach to develop useful predictors of NDI eddy-current performance.

4.4 PROPOSED RESEARCH

The research proposed here will incorporate elements of the recent findings of Drury and his colleagues, specifically with regard to NDI performance, with the findings of Thackray and others (e.g., Thackray et al., 1973, 1974; Thackray and Touchstone, 1980) who have examined correlates of sustained vigilance performance, with the intention of (a) isolating variables that successfully predict NDI task performance and (b) examining the interactions of these relationships with sustained performance on an NDI task.

The task to be used is the simulated NDI eddy-current task devised by Drury and his colleagues and described in studies by Drury et al. (1991) and Latorella et al. (1992). In essence, the task is implemented on a SUN SPARC workstation using a standard keyboard and optical three-button mouse as input devices. The display consists of four windows:

> Inspection Window. The left-central portion of the screen displays rows of simulated aircraft fuselage rivets. The subject uses the mouse to circle each rivet in order to classify it as defective or nondefective.
>
> Macro-view and Directionals. A macro-view in the upper left portion of the screen allows the subject to determine where, or what area on the total simulated fuselage, he is currently examining.
>
> Eddy-Current Meter. Defect indications are displayed in a simulated analog meter located in the upper right window of the screen. Deflections beyond a set point on the meter produce an audible alarm as well as a red flash on an indicator light. The particular meter value of the set point may be either subject or experimenter determined. Subjects judge whether deflections beyond the set point value are likely to be indicative of a defect.
>
> Lower-Right Window. This area of the display is used as a dialogue region in which subjects may use the mouse to exercise a number of task options (e.g., "zoom" to take a closer look at a rivet being inspected, stop task in order to take a break, display elapsed time, record subjective assessments).

The NDI simulation program generates an output file of the subject's performance which gives summary performance measures for the entire task. These include measures of time taken for various aspects of the task, number of hits, misses, and false alarms, total number of faults present, number of rivets classified and unclassified, and total number of rivets visited.

A SUN SPARC workstation has been procured by CAMI and the HOOPS graphics software installed. Arrangements will be made shortly with Dr. Drury at SUNY to procure and install the NDI simulation software.

Phase II of the present contract, then, is to conduct a pilot study using the SUNY NDI simulation with the intent of examining the relationships of a number of potential predictor variables to performance on this task. The first several months of this phase will largely be spent (a) in familiarization with the NDI task and its functional characteristics and (b) in selecting tests and measures that would appear to be promising measures of the relevant task behaviors involved. It would be premature at this time to specify exactly all of the behaviors to be assessed or the particular tests most relevant to each. However, based upon interviews with NDI inspectors and instructors, the findings of Drury and others relative to both general and NDI inspection, and upon the work of various researchers in the area of vigilance, measures of the following would appear potentially related to task performance and likely candidates for inclusion:

- Decision Making/Judgement
- Concentration/Attentiveness/Distraction Susceptibility
- Motivation/Curiosity/Perseverance
- Boredom Susceptibility
- Sustained Attention
- Mechanical/Electronics Interest and/or Aptitude

The experience of Drury and his colleagues in using the NDI task will be utilized extensively in formulating the testing/task protocol. Thus, it is anticipated that much of the subject's first day will be spent in (a) receiving an orientation/indoctrination in eddy-current testing and in the need for nondestructive testing in general, (b) taking the various psychometric pretests, including visual acuity tests, and (c) administering essentially the same training procedures used by Drury and his colleagues (Drury et al., 1991) in their report of a pilot study using this task. This will then be followed by one or two days of sustained performance on the task. (Actual periods of task performance will be determined after gaining familiarity and experience with task characteristics.) It is expected that 5 to 8 subjects will be tested in the pilot study. The pilot study will be completed on or before May 1, 1993 and a report of the findings submitted to Galaxy Scientific Corporation.

4.5 REFERENCES

Behravesh, M. M., Darimi, S. S., & Ford, M. E. (1988). Human factors affecting NDE technician performance. *Proceedings of the Ninth International Conference on NDE in the Nuclear Industry*, Tokyo, Japan.

Berch, D. B. & Kanter, D. R. (1984). Individual differences. In J. S. Warm (Ed.), *Sustained attention in human performance.* New York: Wiley.

Boisvert, B. W., Lewis, W. H., & Sproat, W. H. (1981). *Uniform qualifications of military and civilian nondestructive inspection personnel.* Report No. LS81 WP254-003, Lockheed-Georgia Company, Marietta, Georgia.

Davis, D. R. & Parasuraman, R. (1982). *The psychology of vigilance.* New York: Wiley.

Drury, C. G. (1975). Inspection of sheet materials -- Model and data. *Human Factors, 17*, 257-265.

Drury, C. G., Prabhu, P., Gramopadhye, A., and Latorella, K. (1991). *Nondestructive testing in aircraft inspection.* Report of a pilot study prepared under subcontract 89-1014-SC-3 to Galaxy Scientific Corporation, Mays Landing, New Jersey.

Drury, C. G. (1992). Inspection performance. In G. Salvendy (Ed.), *Handbook of industrial engineering, second edition.* New York: Wiley.

Eysenck, H. J. (1967). *The biological basis of personality.* Springfield, Illinois: Thomas.

Harris, D. H. (1988). *Human performance in nondestructive inspections and functional tests.* EPRI NP-6052, Project 2705-9, Prepared for Electric Power Research Institute, Nuclear Power Research Institute, Nuclear Power Division, Palo Alto, California.

Harris, D. H. (1990). Effect of human information processing on the ultrasonic detection of intragranular stress-corrosion cracking. *Materials Evaluation, 48,* 475-480.

Harris, D. H. (1991). Eddy current steam generator data analysis performance. In D. A. Steininger (Ed.). *NDE techniques for nuclear heat exchangers. Proceedings of the 1991 International Joint Power Generation Conference,* American Society of Mechanical Engineers, San Diego, California.

Harris, D. H. & Chaney, F. B. (1969). *Human factors in quality assurance.* New York: Wiley.

Gallwey, T. J. (1982). Selection tests for visual inspection on a multiple fault type task. *Ergonomics, 25,* 1077-1092.

Jayachandran, T. & Larson, H. J. (1983). *Statistical analysis for the nondestructive inspection (NDI) technician proficiency measurement program.* NPS-53-83-0008, Naval Post Graduate School, Monterey, California.

Koelega, H. S. (1992). Extraversion and vigilance performance: 30 years of inconsistencies. *Psychological Bulletin, 112* (2), 239-258.

Latorella, K. A., Gramopadhye, A. K., Prabhu, P. V., Drury, C. G., Smith, M. A. & Shanahan, D. E. (1992, October). *Computer-simulated aircraft inspection tasks for off-line experimentation.* Paper presented at the Annual Meeting of the Human Factors Society, Atlanta, Georgia.

Lewis, W. H., Pless, W. M. & Sproat, W. H. (1978). *Proceedings from the government/industry workshop on the reliability of nondestructive inspections.* Report No. SA-ALC/MME 76-6-38-2. Lockheed-Georgia Company, Marietta, Georgia.

Lewis, W. H., Sproat, W. H., Dodd, B. D. & Hamilton, J. M. (1978). *Reliability of nondestructive inspections -- Final report.* Report No. SA-ALC/MME 76-6-38-1, Lockheed-Georgia Company, Marietta, Georgia.

Panhuise, V. E. (1989). Introduction. In *Metals handbook,* 9th Ed., Vol. 17, *Nondestructive evaluation and quality control,* ASM International.

Rummel, W. D., Mullen, S. J., Christner, B. K., Ross, F. B., & Marthart, R. E. (1984). *Reliability of nondestructive inspection (NDI) of aircraft engine components.* Report No. SA-ALC/MM8151, San Antonio Air Logistics Command, Kelly Air Force Base, Texas.

Schroeder, J. E., Dunavant, D. W., & Godwin, J. G. (1988). *Recommendations for improving Air Force nondestructive inspection technician proficiency.* SwRI Project No. 17-7958-845, San Antonio Air Logistics Center, Air Force Logistics Command, Kelly Air Force Base, Texas.

Shepherd, W. T., Johnson, W. B., Drury, C. G., Taylor, J. C., & Beringer, D. (1991). *Human factors in aviation maintenance, phase I: Progress report.* DOT/FAA/AM-91/16, Office of Aviation Medicine, Washington, D. C.

Spencer, F., Borgonovi, G., Roach, D., Schurma.., D., & Smith, R. (1992a). *Generic reliability assessment programs at airline maintenance and inspection facilities.* Final Draft Report prepared for the FAA Technical Center, Atlantic City, New Jersey

Spencer, F., Borgonovi, G., Roach, D., Schurman, D., & Smith, R. (1992b). *Proposed reliability assessment experiment for eddy current inspection of lap splice joints in airline maintenance and inspection facilities.* Final Draft Report prepared for the F/ A Technical Center, Atlantic City, New Jersey.

Summers, R. H. (1984). *Nondestructive inspection: Improved capabilities of technicians: Final Report.* AFHRL-TP-83-63, Training Systems Division, Air Force Human Resources Laboratory, Lowry Air Force Base, Colorado.

Thackray, R. I., Jones, K. N., & Touchstone, R. M. (1973). Self-estimates of distractibility as related to performance on a task requiring sustained attention. *Ergonomics, 16,* 141-152.

Thackray, R. I., Jones, K. N., & Touchstone, R. M. (1974). Personality and physiological correlates of performance decrement on a monotonous task requiring sustained attention. *British Journal of Psychology, 65,* 351-358.

Thackray, R. I. & Touchstone, R. M. (1980). *An Exploratory Investigation of Various Assessment Instruments as Correlates of Complex Visual Monitoring Performance.* FAA-AM-80-17, Office of Aviation Medicine, Washington, D. C.

Thackray, R. I. & Touchstone, R. M. (1981). *Age-Related Differences in Complex Monitoring Performance.* FAA-AM-81-12, Office of Aviation Medicine, Washington, D. C.

Triggs, T. J., Rankin, W. L., Badalamente, R. V., & Spanner, J. C. (1986). *Human reliability impact on inservice inspection. Vol. 2. Review and analysis of human performance in nondestructive testing (emphasizing ultrasonics).* NUREG/CR-4436-V2, Battelle Human Affairs Research Centers, Prepared for U. S. Nuclear Regulatory Commission.

Wang, M. J. & Drury, C. G. (1989). A method of evaluating inspector's performance differences and job requirements. *Applied Ergonomics, 20.3,* 181-190.

Wiener, E. L. (1975). Individual and group differences in inspection. In Drury, C. G. and Fox, J. G. (Eds.), *Human reliability in quality control.* New York: Taylor and Francis.

Chapter Five
Investigation of Ergonomic Factors Related to Posture and Fatigue in the Inspection Environment

5.0 INTRODUCTION

In aircraft inspection and maintenance tasks, one of the most noticeable deviations from ergonomically optimum conditions is that tasks must be performed in restricted spaces which force awkward postures. Literature reviewed during Phase III indicates that tasks which possess excessive postural demands (i.e., cramped positions, maintenance of awkward postures) can produce fatigue and ultimately affect both performance and well-being (e.g., Corlett, 1983; Corlett and Bishop, 1978; Hunting, et al., 1980; Van Wely, 1970; Westgaard and Aaraas, 1984). The project reported here came from a task statement to propose a methodology to study extreme spatial conditions, created by restrictive or confined spaces, and their effect on human posture, performance, and stress.

Dependent upon the area of application, restricted or confined spaces have been defined in a variety of ways. For our purposes, a restrictive space will be defined as any area in which the spatial conditions result in decreases in performance or increases in operator workload, stress, or fatigue. Confined spaces are normally associated with whole-body restrictions which occur when an operator must enter an intervening structure to perform a task (e.g., cargo hold), thus creating a situation in which the entire body is confined to a specific area. However, restrictive spaces are also created in areas where the physical space is unlimited, but the immediate working area is restricted. These partial-body restrictions result in limited movement of a specific body part; for example, tasks aided by access devices (e.g., steps, scaffolding, cherrypickers) cause lower limb restriction, for the feet must reside within a limited area. Other examples include reaching arms through access holes and positioning various body parts in and around fixed aircraft components (e.g., viewing inside a small access panel). These partial-body restrictions may occur in addition to whole-body restrictions, as in interior inspection of the tail compartment which demands that the inspector climb into the area (whole-body restriction), as well as place their head and arms through narrow confines to check components (partial-body restriction).

A model is offered to guide research in the description and prediction of the effects of restrictive spaces and the associated postural, fatigue, and stress effects on performance and workload. Characteristics of the environment, operator, and task which act to define the *restrictiveness* of spaces are identified. The objectives of continuing research are to examine the operator compensations forced by restriction, and their ultimate effect on performance and workload; to develop techniques for measuring and alleviating the restrictiveness of spaces; and to demonstrate the use of these techniques.

The work reported here defines the methodology needed to accomplish these objectives.

5.1 RESTRICTIVE SPACE MODEL

The Restrictive Space model (Figure 5.1) attempts to systematically describe space, in terms of inputs, or factors, which define a physical or perceived space, and outputs which allow the effects of space to be understood and predicted.

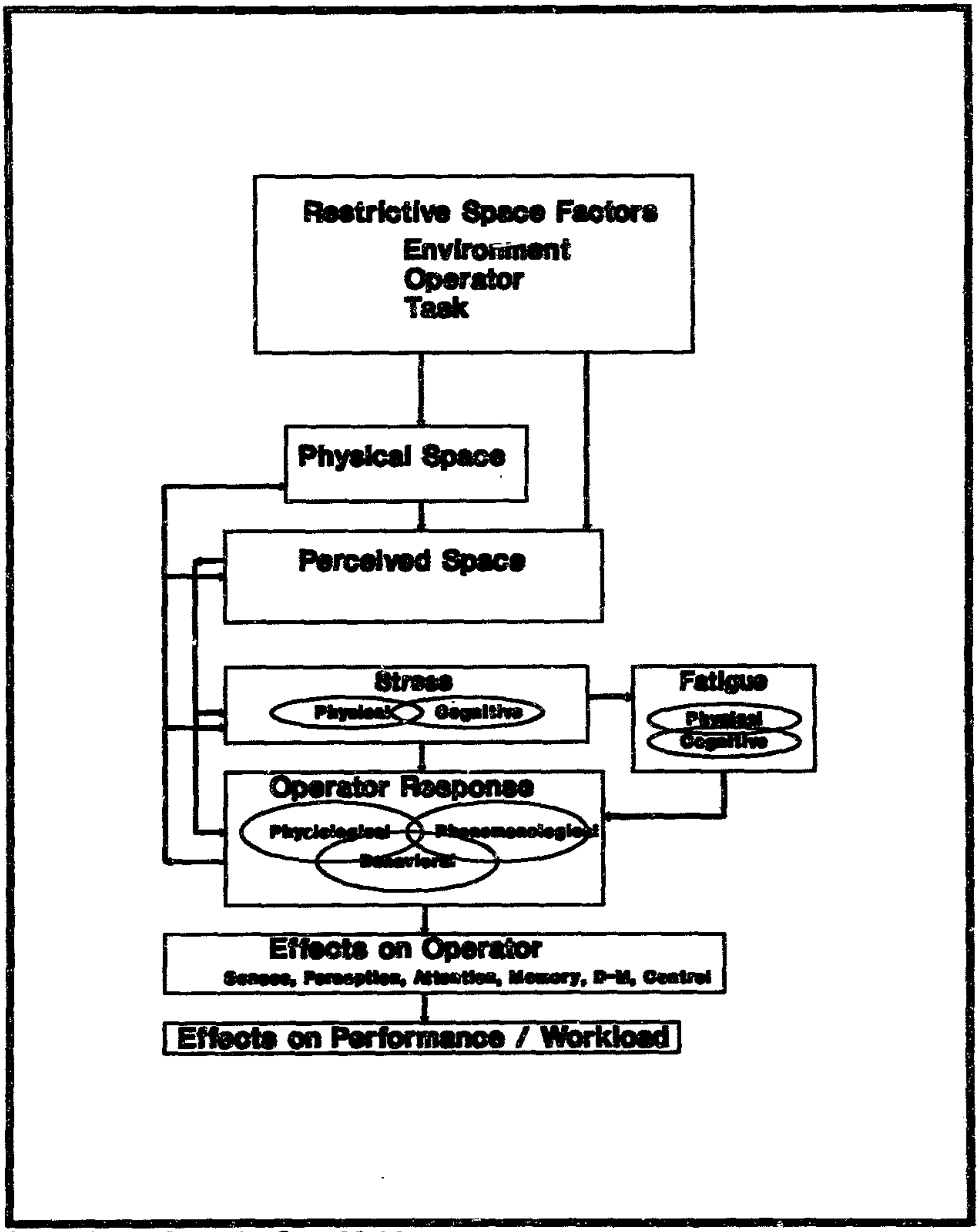


Figure 5.1 Restrictive Space Model

5.2 RESTRICTIVE SPACE FACTORS

Key factors posed by the task, environment, and operator which cause restriction and/or extreme postures have been identified and compiled (Table 5.1). This compilation of factors is not an exhaustive list and may be expanded during on-going investigation.

Table 5.1 Restrictive Space Factors

TASK	ENVIRONMENT	OPERATOR
Demands/Requirements Duration Equipment/Tooling Perceived Value	Affordances Area/Volume Lighting Number of People (acquaintance level, gender, status level) Surface Condition Resources Temperature Ventilation	Age Body Size Experience Flexibility Personality

5.2.1 The Task

The aircraft maintenance/inspection *task demands/requirements* logically define performance in restrictive spaces. Research performed during Phase I produced a generic task description which indicates the primary task demands involved in aircraft inspection (Shepherd, et al., 1991). Each of the primary task components will be discussed below in relation to restrictive space considerations, with the exception of Buyback Inspection. Although Buyback Inspection poses restrictive concerns, they are not unique from those to be considered during the primary inspection functions (i.e., search and decision-making) and thus will not be discussed independently.

5.2.1.1 Initiate

Before the inspection task even begins, the *perceived value* of a task may affect the tolerance to the space conditions. In general, if the operator does not value the job, cursory performance may occur, so that the operator can get out of the space quickly. Most tasks exhibit a speed/accuracy tradeoff (SATO), with faster performance increasing error probability. However, if the cost of mistakes is high, such as in aviation maintenance and inspection, performance is more deliberate. Furthermore, Shepherd, et al. (1991) indicates that inspectors are highly motivated to perform accurately, but a reduction of adverse environmental effects will help ensure that accuracy is never traded for speed.

5.2.1.2 Access

Access tasks consist of physically reaching the area to be inspected. All of these activities involve controlling the movement of the body or body part(s) within a restrictive space. In aircraft maintenance/inspection, this may be an unaided human task (e.g., area inspection of lower fuselage skin), aided by access devices (e.g., steps, scaffolding, cherrypickers), or require access through an intervening structure (e.g., inspection of wing fuel tank interiors through access holes).

Space has been found to be a critical parameter in the mathematical modeling of movement control. In many instances, the amount of space defines the accuracy requirements of a task, which may dictate the speed of performance. Numerous investigations have found a speed/accuracy tradeoff in human performance; as accuracy requirements are increased (i.e., decreased space), performance becomes slower. Mathematical models have been developed which accurately describe the relationship between space and both discrete movement time (e.g., Wickens, 1992) and continuous movement time (e.g., Bottoms, 1982; Drury, et al., 1987). Discrete movement involves moving from one location to another without having to consider the path of movement, while the path of movement is critical in continuous movement which requires moving accurately between two boundaries without exceeding a boundary. These models may be useful in describing and predicting the effects of restriction on movement control tasks in aviation maintenance and inspection.

For example, moving the hand to an access hole may be modeled as a discrete control task, while moving the hand through the access hole is a continuous motor control task, with performance time predicted based upon the accuracy required (i.e., the size of the access hole). Further changes in performance may be found dependent upon the posture adopted while the body part is restricted. Wiker, Langolf, and Chaffin (1989) reviewed research which indicated that there are only minimal differences in manual performance for work heights up to shoulder level. However, they found position and movement performance to decrease progressively when hands were postured above shoulder level, due to the production of movement with pretensed muscles which may serve to increase tremor and decrease maximum velocities.

Likewise, tasks in which the whole body is moved through an access hole may also be modeled as continuous control tasks. Restricted entries and exits have been found to affect ingress and egress times (Drury, 1985; Krenek and Purswell, 1972; Roebuck and Levedahl, 1961), as well as subjective assessments of accessibility (Bottoms, et al., 1979).

These models indicate that the speed chosen by an inspector increases until some limiting speed is reached. The point at which increases in space no longer result in performance being affected is the *performance boundary* (Drury, 1985). However, designing to this boundary does not ensure that increased operator stress, fatigue, or workload does not occur.

5.2.1.3 Search

Search requires the sensing, perceiving, and attending to information. Visual search requires the head to be at a certain location to control the eyes and visual angle. Thus, restricted areas frequently force inspectors to adopt awkward head, neck and back angles which induce stress and fatigue. In many instances, inspectors are forced to either search an area at less-than-optimum viewing angles or indirectly using a mirror. Although both methods can be utilized to produce acceptable performance, inspector workload and stress are increased, and performance is less efficient. These restricted space situations can occur in completely confined areas (i.e., in an interval structure) or in an area in which the space is physically unlimited, but the immediate working space is restricted by the task demands (e.g., wing inspection).

Manual tactile inspection and non-destructive inspection (NDI) techniques, such as Eddy Current or Ultrasonic, require precise motor control on the surface to be inspected. In other words, the task demands are restricting the movement control. In addition, the *tooling and equipment* associated with Eddy Current and Ultrasonic inspection can physically restrict the area. The motor control models discussed above can be related to these situations. Specifically, these models predict that motor control tasks will be performed less efficiently (i.e., speed and accuracy), as the space conditions and postures demanded become more extreme.

5.2.1.4 Decision-Making

Decision-making requires that potential defects located during search be evaluated to decide whether it should be reported based upon specified standards. Comparison standards, which allow direct comparison of the potential defect with a standard at the point of inspection, have been found to improve decision-making (Galaxy Scientific Corporation, 1992). However, restricted areas may prohibit any extraneous material from being easily accessible in the immediate working area (e.g., workcard illustration), thus forcing decisions to be made without comparison standards (increased memory load), or additional time to obtain information from the workcard (fairly rapid task), a manual (a longer task), or a supervising individual. Moreover, as described earlier, viewing angles may be less-than-optimum, further decreasing sensitivity and increasing the difficulty of decisions. Thus, restricted spaces can force the decision-making task to be more memory-intensive, longer, and more difficult.

Conversely, pressures for cursory decision-making may occur, so that the operator can get out of the space quickly. Decision-making tasks exhibit a speed/accuracy tradeoff (SATO), with speeded performance associated with inaccurate decision-making. However, inspectors are highly motivated to perform accurately (Shepherd, et al., 1991), thus it is predicted that accurate decision-making performance would not be compromised by even the most extreme of space conditions, although the workload and stress may increase.

5.2.1.5 Respond

The respond task demands that detected defects be marked and documented. As discussed above, restricted areas may not allow additional material such as non-routine repair forms in the workspace. Thus, the inspector must remember all defects within an area until they are later documented on the appropriate forms. This situation can create a high memory load on inspectors and presents the potential for an inspector to forget to note a defect.

5.2.1.6 Repair

Many repair tasks require mechanics to be in a confined or restricted area for prolonged periods of time. *Task duration*, which forces longer periods of time in a restrictive area, could psychologically affect the perception of space. Habitability literature, concerned with the study of manned underwater vessels and space vehicles, indicates that internal space requirements vary as a function of duration (Blair, 1969; Price and Parker, 1971). Furthermore, Cameron (1973) indicates duration to be the primary variable associated with fatigue effects.

In addition, extreme space conditions only allow a limited number of inefficient postures to be adopted, thus physical working capacity may be reduced in restrictive spaces, as indicated by research in the area of manual material handling (Davis and Ridd, 1981; Mital, 1986; Ridd, 1985; Rubin and Thompson, 1981; Stalhammer, et al., 1986). Under unlimited space conditions, operators are able to adopt efficient postures or switch postures and use other muscle groups, enabling primary muscle groups to be rested (Drury, 1985). However, frequent breaks from restrictive areas, common during maintenance/inspection activities, allow relief from sustained task performance and allow the primary muscle groups to be rested.

5.2.2 The Environment

The physical *volume* of space obviously alters the workplace. A majority of the research in this area has focused on investigating the effect of the amount of space on task performance and was discussed above in Section 5.2.1.

Lighting and surface condition may create extreme spatial conditions. For example, poor lighting can demand a certain posture to be adopted for task performance, by forcing a specific visual angle. An oily surface can act in much the same manner, by limiting the postures which an operator is willing to adopt to avoid oil-soaked clothing. Other variables may act to exacerbate the perception of a restrictive environment (e.g., *extreme temperatures, poor ventilation*).

Conversely, other environmental characteristics may moderate the effects of a space, such as restrictions which provide *affordances* (i.e., support). For example, Davis and Ridd (1981), Ridd (1985), and Rubin and Thompson (1981) found that when restrictions acted as supports for manual material handling in restricted spaces, lifting capacity increased. Thus, given a restricted situation, interesting interactions may exist, for task performance may be aided by some restrictions and degraded by others.

Social aspects of the environment may limit space. As the *number of people* within a given area increases, the amount of space for a single person decreases. Although a majority of the inspection work is performed by a single individual (Shepherd, et al., 1991), many maintenance tasks require more than one person. If uncomfortably close spacing is required between individuals, tolerance to the environment may be limited. In addition, the acquaintance level, gender, and status level of individuals within the environment may mediate this response (Little, 1965). For example, a mechanic may experience a more intense reaction if an inspector is within the restricted space during task performance. If there are many individuals within the same area, performing the same tasks, the available *resources* may become limited, the space may be perceived to be more restrictive and people may become frustrated (e.g., specialized tooling not available, thus making the task more difficult).

5.2.3 The Operator

Body size may act to limit the amount of physical space, which in turn may increase the restriction. Roebuck and Levedahl (1961) found body dimensions to be somewhat predictive in the analysis of aircraft escape times through restrictive door and window exits. Body dimensions may also indirectly affect posture, for smaller individuals may be able to adopt postures more conducive to space reduction. Restrictive spaces may force individuals to adopt unnatural postures; thus, smaller individuals may be more able to do this. Bodily *flexibility* may also be associated with the ability to adopt unnatural postures caused by restricted spaces.

Age has been found to have an effect on task performance, primarily due to a deterioration in the physical and cognitive function of older individuals. However, this effect may be reversed when experience is important (Czaja and Drury, 1981). These effects may be particularly relevant in aviation inspection, since most inspectors are senior personnel.

Experience, or familiarity with a situation, may act to moderate the stress response. Previous exposures, practice, or conditioning may reduce uncertainty, or influence the perceived demands, constraints, and strategy selections (Sutherland and Cooper, 1988). Personal communication with inspectors has revealed that repeated exposures, as well as attitude, reduced the stress response during whole-body restrictions.

Enduring *personality characteristics and cognitive style* do have an effect on some subtasks demanded during maintenance and inspection performance. Historically, there have been attempts to select personnel based upon personality scales and aptitude tests thought to be relevant to various jobs. However, Wiener (1975) indicates that most of these efforts have not been fruitful and endorses job design, training, and motivation as better alternatives.

5.3 PHYSICAL AND PERCEIVED SPACES

The above factors can directly affect the spatial conditions. The workspace has physical characteristics which can be easily defined and investigated, but the physical space is also perceived by the operator. Thus, the effective workspace is partially created by physical elements and partially by perceived elements. Thus, the effective workspace within a fixed physical space is not necessarily constant but is dependent upon an individual's constantly-changing perceptions. The effects of this effective space must be inferred, as direct observation is not logically possible.

5.4 STRESS

It is logical to model these restrictive space effects within a traditional stress framework, where the extreme space conditions act as a stressor. Context-dependent examination of the space-affecting factors allows the specific stress-inducing situation to be defined, so that the subjects' perceptions may be determined to assist in interpreting behavior (Meister, 1981). Thus, field investigation is important for understanding the specific response to restricted spaces in aircraft maintenance/inspection activities. However, controlled laboratory studies allow more precise data to be collected without disrupting the actual maintenance/inspection activities. Therefore, both field and laboratory studies will be needed to understand the effects of restricted space. In an effort to operationally define stress within the context of restrictive space, the following definitions will be employed (Alluisi, 1982; Pratt and Barling, 1988):

> **Stressor** - The environmental, operator, and task characteristics which comprise the space and impinge on the individual. In this context, the physical and perceived spaces are the stressors.
>
> **Stress** - A state within the individual caused by the perceived magnitude of the stressor. The existence and interaction of the various environmental, operator, and task characteristics will dictate the intensity of the stress.

Task performance in restrictive spaces normally includes both physical and cognitive demands; the stress induced by these demands will be differentiated to more clearly define and understand the various stress responses. *Physical stress* is directly perceived by the involved physical subsystems within the individual (e.g., biomechanical, physiological) due to a discrepancy between the environmental/task demands and the individual's physical ability to meet the demands. It is perceived by an individual through a specific, or localized, experience of discomfort. Thus, response can be specifically aimed at eliminating, or alleviating, the stressor when possible. There will also be an overall physiological response to bodily requirements caused by the restriction. For example, restriction may cause postural stress and discomfort in various muscle groups, which results in increases in heart rate and blood pressure (Astrand and Rodahl, 1986).

Cognitive stress creates a cognitive state resulting from an individual's perception of the discrepancy between the perceived environmental/task demands and their perceived ability to meet those demands (Cox, 1990, 1985). It is this mismatch which eventually determines the stress reaction, thus the operator perceptions play a key role. This stress is experienced as negative emotion and unpleasantness (Cox, 1985; Sutherland and Cooper, 1988), and may be difficult to localize.

It is hypothesized that whole-body confinements, as opposed to partial-body restrictions, are more apt to produce cognitive stress effects. Inspectors may feel that they have less control to adapt, or adapt to, the perceived space. For example, when totally enclosed within an area, there may be fewer opportunities to eliminate the stressor (e.g., frequent rest breaks outside the space). Both whole-body and partial-body restrictions are hypothesized to cause physical stress effects, particularly postural, due to the body positions which are demanded. However, these physical stress effects will most likely lead to cognitive stress effects, if task completion is compromised. In summary, the effects of stress on human performance provide the basis for investigation. These effects include increased arousal, increased processing speed, reductions in working memory, reduced attentional capacity and attentional narrowing, and changes in the speed and accuracy of performance (Hockey and Hamilton, 1983; Hockey, 1986; Reynolds and Drury, 1992; Wickens, 1992).

5.5 FATIGUE

As discussed above, task performance under extreme spatial conditions can present both physical and cognitive stress, which in turn can induce physical or cognitive fatigue. *Physical fatigue* may be defined as a state of reduced physical capacity (Kroemer, et al., 1990). Work can no longer be continued because the involved physical subsystems are not capable of performing the necessary functions. For example, a posture can no longer be maintained due to exceeding the endurance limit of the muscles (Rohmert, 1973).

Cognitive fatigue is a term normally associated with stress and may be broadly defined as a generalized response to stress over time. The effects may reside as a psychological state within the individual or extend to affect performance. Symptoms of fatigue include restricted field of attention, slowed or impaired perception, decreased motivation, subjective feelings of fatigue and task aversion, and decreased performance in the form of irregularities in timing, speed, and accuracy (Bartlett, 1953; Grandjean and Kogi, 1971).

5.6 OPERATOR RESPONSE

The operator response is a function of the perceived space, and the associated stress and fatigue effects. In most instances, this response cannot be described by one variable but is manifested in various physiological, psychophysical and behavioral patterns.

An individual may respond to, or cope with, a stressful situation in order to lessen the effect of, or eliminate, the stressor (Cox, 1985). A dependency may exist between the different modes of response (i.e., psychophysical, physiological and behavioral). Any mode(s) of response may in turn elicit another mode(s) of response (Meister, 1981). For example, while performing maintenance or inspection in a cramped area of an aircraft, there may be an initial physiological response to the postural demands such as lack of blood flow to the leg muscles, which in turn causes a behavioral response (e.g., posture shifting) and/or subjective response (e.g., perceived discomfort). In addition, in the context of restrictive space, a response may alleviate one component of the stress response, while causing another. Continuing the example, a change in posture may reduce the physiological response, but the new posture may make the task more difficult to perform, causing feelings of frustration.

5.7 EFFECTS ON OPERATOR

In order to describe, or possibly predict, the effects of operator response on performance and workload, there is a need to understand the effects of stress and fatigue on the operator. These effects were cited previously in their respective sections (Sections 5.4 and 5.5). If performance is affected, a specification of the affected subsystem and why it is affected may be possible. For example, *perception* may be affected by the inability to obtain an adequate visual angle, *attention* may be distracted by discomfort due to postural stress, or *decision-making* may be speeded up in an effort to finish the task and eliminate the stressor (i.e., leave the environment).

5.8 A RESTRICTIVE SPACE FRAMEWORK TO MEASURE THE EFFECTS ON PERFORMANCE/WORKLOAD

Performance and workload will ultimately be affected by any changes in operator function forced by the spatial conditions and associated stress and fatigue. Drury (1985) advances a 3-level framework which attempts to describe task performance with respect to physical space. The following proposed framework includes an additional zone to better predict the effects of space and awkward postures on inspector stress and workload as well as performance. This framework presents four zones which specifically define performance, workload, and stress (**Table 5.2**).

Table 5.2 Performance, Workload, and Stress Defined within Restrictive Space Framework

ZONE	PERFORMANCE	WORKLOAD	STRESS
0	None possible	$W=0$	$S=0$
1	Proportional to space	$W_{task + compensation(s)}$	$D_{task + compensation(s)} > HOC$
2	Acceptable	$W_{task + compensation(s)}$	$D_{task + compensation(s)} > HOC$
3	Acceptable	W_{task}	$D_{task} < HOC$

5.8.1 Zone 0 - Anthropometrically Restricted Zone

The task cannot be accomplished, as the space conditions, or postures, are too extreme for the operator to function. The boundary between zone 0 and zone 1 is normally determined by anthropometric data (i.e., human dimensions). These minimum criteria are only used if space is a critical commodity (e.g., aircraft). Under normal conditions, larger spaces are recommended. The limitations in using this type of data are that it is normally based on static sitting/standing and does not account for normal working postures, does not include any allowance for special equipment, and represents a young population. Hence, anthropometrically-defined spaces must underestimate minimum space requirements (Drury, 1985). There are computer-aided systems, such as CREWCHIEF (McDaniel and Hofmann, 1990) which account for some of these limitations. However, one manufacturer, which has developed and utilizes a similar computer-aided human modeling system, admits that,

> ...[these] systems [have] limits, and some mock-ups still will be required. "Human models....can't do all the interface work,"...

Nevertheless, even if these 'minimum allowance models' could ensure that individuals can work in a given space, they do not account for fatigue, workload, or stress effects

5.8.2 Zone 1 - Performance Restricted Zone

Task performance is possible, but performance is not optimum because the spatial conditions/posture still interfere with the task. This zone ranges from allowable access for task performance up to acceptable task performance. As the space increases, performance increases. The total workload is equal to the workload associated with the task plus the workload associated with the operator compensations caused by the workspace. Similarly, there is increased stress present in this zone, for the task demands exceed the operator capabilities. Workload and stress most likely decrease within the zone, because as the spatial demands decrease, the compensations should decrease.

5.8.3 Zone 2 - Workload/Stress Restricted Zone

Task performance is acceptable, at least in the short term, but operators' workload and stress are increased because of compensating for the limited space and/or extreme postures. As space increases within this zone, operator compensation(s) or responses should decrease, thus causing the total workload and stress to decrease.

5.8.4 Zone 3 - Unrestricted Zone

This zone allows acceptable task performance without additional operator compensation; thus, there is no additional workload or stress imposed by the spatial conditions.

5.9 RESTRICTIVE SPACE METHODOLOGY

Experimentation will utilize the restrictive space model to assist in understanding and describing the relationships between the spatial conditions and the operator compensations, fatigue, stress, and ultimately performance and workload. The restrictive space framework will be used to guide the categorization of restrictive spaces and describe the effects on stress, workload, and performance. This research will include field investigation in conjunction with a series of laboratory experiments to investigate the effects of restrictive space on visual inspection performance, which accounts for 90 percent of all inspection activities in aircraft inspection (Shepherd, et al., 1991). Extensions to NDT inspection, which involves the additional restriction of working with equipment, are possible at a later date.

Knowledge of the effect of awkward postures and restrictive spaces on the human operator, reviewed in the previous sections, will be applied within the following methodology to give:

1. A recognition guide which allows users to predict which tasks will have a performance decrement and/or stress increase due to the spatial/postural demands.

2. A set of interventions, keyed to task, operator, and environmental factors, which will reduce the spatial demands and operator workload, stress, and fatigue.

5.10 ON-SITE EVALUATION

5.10.1 Task Description

The task analysis procedure developed during Phase I (Shepherd, et al., 1991) will be adapted and applied for use in assessing restrictive spaces. Detailed descriptions of a representative sample of tasks which possess restrictions and awkward postures will be obtained. This step will include having human factors analysts work with inspectors during the completion of workcards. While obtaining task descriptions, emphasis will be placed on documenting environmental, operator, and task factors identified in the previous section which create, or exacerbate, restricted spaces or extreme postures.

5.10.2 Behavioral Measures

Extreme spatial conditions can limit the number of postures adopted or force unnatural postures. An adapted version of Branton and Grayson's (1967) postural recording scheme will be utilized to measure whole body postures. The number of postures adopted and the frequency of each posture will be obtained. These measures have been successfully applied to the assessment of postural demands in various work tasks (Bhatnager, et al., 1985; Branton and Grayson, 1967; Zhang, et al., 1991) other than aircraft inspection.

5.10.3 Psychophysical Measures

Physiological monitoring, which presents many difficulties and limitations, will be limited to the laboratory, thus an emphasis will be placed on utilizing psychophysical techniques in field studies. These techniques are attractive, particularly for field use, for they are unrestrictive, require minimal instrumentation, and thus easy to use/administer, and give valid and reliable [illegible] which can be related to other non-aviation tasks.

Feeling Tone Checklist (FTC). This scale will be utilized to measure fatigue effects over time. It is an interval scale which has been found to be a valid and reliable measure of subjective feelings of fatigue (Pearson, 1957).

Body Part Discomfort Chart (BPD). This is the most noted technique utilized to obtain postural discomfort data (Corlett and Bishop 1976). This chart categorizes the body into a number of functional areas to allow the assessment of individual body areas. A 5-point ordinal scale will be utilized to solicit operators' *BPD* ratings.

NASA - Task Load Index (TLX). This is a multi-dimensional rating scale which measures six workload-related factors (e.g., mental demand, physical demand, temporal demand, performance, effort, and frustration) and their associated magnitudes to form a sensitive and diagnostic workload measure (Hart and Staveland, 1988).

5.10.4 Experimental Protocol

A representative sample of aircraft inspection tasks, which include both whole-body and partial-body restrictions, will be selected for field investigation. Postural data will be collected throughout task performance. The FTC and BPD will be administered before, during, and after task performance. Ideally, the same inspectors will be compared across tasks; although, this may not always be possible due to the scheduling demands on the hangar floor. In addition, the TLX will be administered after task performance.

5.11 LABORATORY EVALUATION

5.11.1 Experiments

Experimentation will involve a series of studies investigating single and multiple whole-body restrictions. Investigation will focus on examining the effect of spatial restriction/extreme postures on performance, workload, stress, and fatigue. These experiments will focus on the effects of restrictions and their interactions in three planes: side-to-side (lateral), head-to-feet (vertical), and front-to-back (sagittal) restrictions. Experimentation will be driven by the restrictive space framework, to demonstrate the various restrictive space zones. Based upon this data, predictions can be made of the effects of various spatial conditions on performance, workload, stress, and fatigue.

5.11.2 Tasks

An aircraft inspection task will be used to simulate the inspection environment during laboratory investigation. In addition, a neutral inspection task will be utilized to provide more easily-interpreted results, shorter training time, and more sensitive measures of search and decision-making. Both tasks are computer-based.

Aircraft Inspection Task. The visual inspection task is simulated on a SUN SPARC Station 1 Workstation. The task requires the inspector to search for multiple defects frequently found on an aircraft, including missing, damaged, pooched/dished, and loose rivets, dents, and rivet cracks. The defects may be classified as critical or noncritical, dependent on the severity of the defect.

This simulator includes a windowing function (Galaxy Scientific Corporation, 1992), which results in only a small area being fully illuminated, within a large inspection field. Only within this window can faults be detected and indicated. The entire inspection field is viewed by successive movements of this window. This windowing function forces what are known as field of view (FOV) movements (Drury, in press). This will allow the search strategy process measures described in Section 5.11.4 (e.g., fixation time, sequential distribution of fixation, etc.) to be collected. This is an attractive alternative for the measurement of eye movement parameters, in contrast to conventional techniques which require sophisticated and restricting instrumentation to be attached to the subject.

Neutral Inspection Task. The experimental inspection task will include a random arrangement of background and target characters (Barnes, 1984). The task allows the two primary components of inspection, search and decision making, to be measured separately. The software will be adapted to include an inspection window so that eye movement parameters may be obtained.

5.11.3 Independent Variables

The current research will focus on investigating the effect of a subset of restrictions described in the previous section (**Table 5.1**).

The Environment. The *physical amount of space* will be altered in order to determine the effects of various restrictive environments. The *affordance* of the various restrictions (i.e., support provided) will be changed indirectly through volume/area alterations and be dependent upon the experimental space conditions. For example, some restriction(s) will be extreme enough that they allow subjects to lean against them during task performance. These behaviors will be noted.

The Operator. To control for *body size*, anthropometric measurements will be utilized to standardize the amount of space to each subject (e.g., space is equivalent to percentage of various body measurements). Thus, clearance conditions will be equivalent for each subject. In an effort to obtain face validity, a sample representative of the current inspector population will be selected. *Age* will be used as a covariate to control for any possible age effects. The level of *experience* in the restricted space will be controlled by utilizing a between-subjects experimental design which ensures that each group of subjects is only exposed to one restrictive environment.

Task specific pretests which measure different *cognitive styles/personalities* may provide some predictive power and partially explain the inherent variability between individuals so that the results can be better understood. Four pretests, which are relevant to the task context and environmental conditions, will be utilized in this study: Embedded Figures Test (EFT), Matching Familiar Figures Test (MFFT), Locus of Control (LC), and a claustrophobia screening test.

The Task. The inspection tasks will allow two operator *demands* to be investigated: search and decision-making. Each experimentation period will include baseline, task, and recovery periods, to be described more fully in the experimental procedure section below (Section 3.3). In addition, the effects during the baseline, task and recovery periods will be measured over time (i.e., *duration*).

5.11.4 Dependent Variables

Behavioral. An adapted version of Branton and Grayson's (1967) postural recording scheme will be utilized to measure whole body postures. A computerized version of this system will be developed, posture will be indirectly observed (i.e., videotape), and positions directly input. Thus, a continuous record of all positions, their frequencies, and durations can be obtained. This gross assessment does not provide data on the magnitude of the postural deviations, only that postural changes occur. Thus, the specific magnitude of the trunk, neck, and head angles will also be measured for each posture. The frequency of each posture, their duration, and the number of posture changes will be measured and related to postural severity and discomfort (Bhatnager, et al., 1985; Branton and Grayson, 1967; Drury, et al., 1988; and Zhang, et al., 1991).

Physiological. Heart rate and respiration rate will be obtained. These measures were chosen for the following reasons: (1) they present minimal intrusion on task performance, and (2) they are sensitive to changes in physical stress (e.g., Astrand and Rodahl, 1986), cognitive stress (e.g., Kak, 1981), and workload (e.g., O'Donnell and Eggemeier, 1986). As indicated earlier, physiological monitoring possesses many limitations. Restrictive environments present additional inherent difficulties. However, a limited number of measures will be obtained in an attempt to capture this primary stress response.

Psychophysical. In addition to the FTC, BPD and TLX described in Section 5.11.3, a modified version of the Stress-Arousal Checklist (MSACL) (Mackay, et al., 1978) will be utilized to measure stress and arousal levels experienced in restrictive spaces (Cruickshank, 1984).

Task Performance. Inspection is a two-stage process which demands search and decision-making. Visual search proceeds as a series of fixations at specific points in the visual search field. These fixations are separated by saccades which occur when an individual moves his/her eyes to a new location. Factors which affect search performance include search strategy, speed/accuracy tradeoff (SATO), and stopping policy.

Search strategy is defined by the overall pattern of eye movements, in our case FOV movements. The following parameters will be measured: fixation time, spatial distribution of fixations, sequential distribution of fixations, and interfixation distances. These subtle process measures may be sensitive to the fatigue and stress effects described previously. These effects may be exhibited by a change in the number or rate of fixations or a more random search path (Latorella, et al., 1992).

Speed/Accuracy Tradeoff (SATO) can be assessed by the performance measures: search time taken to detect a fault, search errors (i.e., failing to locate a fault), and stopping time (i.e., time taken to search for a defect before giving up). There is evidence to suggest that individuals change their operating point, with respect to speed and accuracy, under various stressful conditions (Hockey, 1986).

Decision-making is required if a defect is detected. This decision process can be modelled by signal detection theory (SDT), which describes how humans detect signals in noise (Wickens, 1992). Within this SDT structure, three factors can affect decision-making: sensitivity, criterion, and SATO.

Sensitivity is a measure of discriminability, the perceived difference between the observed flaw and standard, and may be affected by the extraneous noise introduced by restrictive environments.

Decision Criterion is the internal standard chosen by an inspector for reporting a fault and can be affected by the defect rate, cost of errors, and time on task (Galaxy Scientific Corporation, 1992). Shifts in criterion may be found as time on task increases due to missed signals and corresponding changes in signal expectancy. Moreover, criterion changes have been found to be caused by stress-inducing situations (Wickens, 1992).

SATO is the amount of time taken to make a decision and can affect sensitivity. In restrictive environments, signal integration may be affected by speeded performance. Research was reviewed earlier which indicated that stress can cause a strategic change in performance resulting in an increase in speed and errors (e.g., Hockey, 1986).

The following measures will be obtained in order to assess changes in sensitivity, criterion, and SATO: decision time (i.e., time to make decision, after defect detected), misses (i.e., deciding not to indicate a defect which is classified as defective) and false alarms (i.e., deciding to indicate a defect which is not classified as defective).

5.11.5 Experimental Procedure

Subjects will perform the inspection task under unrestricted and restricted space conditions in order to measure *changes* in performance between the two conditions. The task periods will be segmented into several portions and separated by breaks, which is characteristic of the maintenance and inspection task organization. A conventional stress design will be employed, thus allowing baseline measures to be obtained and any aftereffects assessed.

Performance measures will be obtained during each task period. In addition, videotape analysis allows the frequency, duration, and severity of each posture, and the frequency of posture changes to be obtained continuously. The BPD, MSACL, and FTC will be administered at the beginning and end of each task period, while the TLX will be obtained at the end of each task period. The physiological measures will be obtained throughout the baseline, task and recovery periods.

5.12 ANALYSIS OF RESULTS

The results of the on-site evaluations and laboratory experiments will be combined and analyzed to derive operational definitions of the zone boundaries. This will allow the recognition and prediction of tasks which will have performance decrements and/or workload/stress increases due to restrictive spaces or extreme postures. Knowledge of human factors models of human inspection and the functioning of individual human subsystems (i.e., senses, perception, attention, memory, decision-making, feedback, and control) can be utilized to identify the subsystems affected by the identified restrictions.

The additional demands/compensations forced by restrictions can be determined and the effects predicted and described by the restrictive space framework presented in Section 5.0. Anthropometric models, and subsequently population percentiles, will be utilized to quantify and operationally define these effects in terms of absolute space dimensions within this framework.

5.13 DEVELOPMENT OF CHECKLIST/INTERVENTION GUIDE

Based upon the results, a recognition checklist will be developed which classifies and describes the effects of restrictive spaces and extreme postures on inspection tasks into one or more zones. Procedures aimed at alleviating the reduced performance and increased workload/stress in Zones 1 and 2 will be devised and compiled. The task demands, and associated compensations forced by restrictions, can be compared with known human capabilities to provide interventions aimed at the identified environmental, operator, and task restrictive space factors in order to reduce operator workload/stress and improve performance. These intervention strategies will be used in the development of a guide.

5.14 REFERENCES

Alluisi, E.A. (1982). Stress and stressors, commonplace and otherwise. In E.A. Alluisi, and E.A. Fleishman (Ed.), *Human performance and productivity, Vol 3: Stress and performance effectiveness.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Astrand, P. and Rodahl, K. (1986). *Textbook of work physiology.* New York: McGraw-Hill.

Barnes, R.E. (1984). *Inspection performance of mentally retarded persons.* Unpublished doctoral dissertation, State University of New York at Buffalo.

Bartlett, F. (1953). Psychological criteria of fatigue. In W.F. Floyd and A.T. Welford (Eds.), *Symposium on fatigue.* London: H.K. Lewis and Co.

Bhatnager, V., Drury, C.G. and Schiro, S.G. (1985). Posture, postural discomfort, and performance. *Human Factors, 27*(2), 189-200.

Blair, W.C. (1969). Human factors in deep submergence vehicles. *Marine Technology Society Journal, 3*,(5), 37-46.

Bottoms, D.J., Barber, T.S. and Chisholm, C.J. (1979). Improving access to the tractor cab: An experimental study. *Journal of Agricultural Engineering Research, 24*, 267-284.

Bottoms, D.J. (1982). The tractor driver's steering control task. *Ergonomics, 25*, 31-39.

Branton, P., and Grayson, G. (1967). An evaluation of train seats by observation of sitting behavior. *Ergonomics, 10*, 35-51.

Cameron, C. (1973). A theory of fatigue. *Ergonomics, 16*, 633-648.

Corlett, E. N. (1983). Analysis and evaluation of working posture. In Kvalseth, T. O. (Ed.), *Ergonomics of workstation design*. London: Butterworths.

Corlett, E.N. and Bishop, R.P. (1976). A technique for assessing postural discomfort. *Ergonomics, 19*, 175-182.

Corlett, E.N. and Bishop, R.P. (1978). The ergonomics of spot welders. *Applied Ergonomics, 9*, 23-32.

Cox, T. (1985). The nature and measurement of stress. *Ergonomics, 28*, 1155-1163.

Cox, T. (1990). The recognition and measurement of stress: Conceptual and methodological issues. In J.R. Wilson, and E.N. Corlett (Eds.), *Evaluation of human work*. London: Taylor and Francis.

Cruickshank, P.J. (1984). A stress and arousal mood scale for low vocabulary subjects: A reworking of Mackay, et al., (1978). *British Journal of Psychology, 75*, 89-94.

Czaja, S. and Drury, C. G. (1981). Aging and pretraining in industrial inspection. *Human Factors, 23*, 473-483.

Davis, P. R. and Ridd, J. E. (1981). The effect of headroom on acceptable lifting capacity. *Ergonomics, 24*, 239.

Drury, C.G. (1985). Influence of restricted space on manual materials handling. *Ergonomics, 28*, 167-175.

Drury, C.G. (in press). Exploring search strategies in aircraft inspection. In D. Brogan (ed.), *Visual search* (2nd ed.). London: Taylor and Francis.

Drury, C.G., Goonetilleke, R.S., and Maurice, M. (1988). *Automation, stress, and performance in industrial inspection*. Final Report on PHS Grant 5 R01OH024, Impact of Automation on Industrial Inspector Stress.

Galaxy Scientific Corporation (1992) (in press). *Human Factors in Aviation Maintenance Phase Two: Progress Report*.

Drury, C.G., Montazer, M.A. and Karwan, M.H. (1987). Self-paced path control as an optimization task. *Systems, Man, and Cybernetics, 17*, 455-464.

Grandjean, E., and Kogi, K. (1971). Introductory remarks. In K. Hashimoto, K. Kogi, and E. Grandjean (Eds.), *Methodology in human fatigue assessment*. London: Taylor and Francis.

Hart, S.G, and Staveland, L.E. (1988). Development of NASA-TLX: Results and empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), *Human mental workload*. North Holland: Elsevier Science Publishers.

Hockey, G.R.J. (1986). Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. In K.R. Boff (Ed.), *Handbook of perception and human performance* (Vol 2). New York: Wiley and Sons.

Hockey, G.R.J. and Hamilton, P. (1983). The cognitive patterning of stress states. In R. Hockey (Ed.), *Stress and fatigue in human performance*. New York, NY: John Wiley and Sons.

Hunting, W., Grandjean, E. and Maeda, K. (1980). Constrained postures in accounting machine operators. *Applied Ergonomics, 11*, 145-149.

Kak, A.V. (1981). Stress: An analysis of physiological assessment devices. In G. Salvendy and M.J. Smith (Eds.), *Machine pacing and occupational stress*. London: Taylor and Francis.

Krenek, R. F. and Purswell, J. L. (1972). Automobile escape worthiness--an approach to a predictive model. *Proceedings of the Human Factors Society's 16th Annual Meeting*, 46-57.

Kroemer, K.H.E., Kroemer, H.J. and Kroemer-Elbert, K.E. (1990). *Engineering Physiology* (2nd ed.). New York: Van Nostrand Reinhold.

Latorella, K.A., Gramopadhye, A.K., Prabhu, P.V., Drury, C.G., Smith, M.A. and Shanahan, D.E. (1992). Computer-simulated aircraft inspection tasks for off-line experimentation. *Proceedings of the Human Factors Society's 36th Annual Meeting*, 92-96.

Little, K. B. (1965). Personal space. *Journal of Experimental Social Psychology, 1*, 237-247.

Mackay, C., Cox, T., Burrows, G., Lazzerini, T. (1978). An inventory for the measurement of self-reported stress and arousal. *British Journal of Social and Clinical Psychology, 17*, 283-284.

McDaniel, J.W., and Hofmann, M.A. (1990). Computer-aided ergonomic design tools. In H.R. Booher (Ed.), *MANPRINT: An approach to systems integration*. New York: Van Nostrand Reinhold.

Meister, D. (1981). The problem of stress definition. In G. Salvendy and M.J. Smith (Ed.), *Machine pacing and occupational stress*. London: Taylor and Francis.

Mital, A., (1986). Subjective estimates of load carriage in confined and open spaces. In W. Karwowski (Ed.), *Trends in ergonomics/human factors III.* North Holland: Elsevier, 827-833.

O'Donnell, R.D., and Eggemeier, F.T. (1986). Workload assessment methodology. In K.R. Boff (Ed.), *Handbook of perception and human performance* (Vol. 2). New York: John Wiley and Sons.

Pearson, R.G. (1957). Scale analysis of a fatigue checklist. *Journal of Applied Psychology, 41*, 186-191.

Pratt, L.I., and Barling, J. (1988). Differing between daily events, acute and chronic stressors: a framework and its implications. In J.J. Hurrell, L.R. Murphy, S.L. Sauter, and C.L. Cooper (Eds.), *Occupational stress, issues and developments in research.* London: Taylor and Francis.

Price, H. E. and Parker, J. F., Jr. (1971). Forecast of human factors technology issues and requirements for advanced aero-hydro-space systems. (Contract #N00014-69-C-0327). Arlington, VA: Office of Naval Research.

Reynolds, J.L., and Drury, C.G. (1992). The effects of physical exertion on task performance in modern manufacturing: A taxonomy, a review, and a model. Paper submitted for publication in Ergonomics.

Ridd, J.E. (1985). Spatial restraints and intra-abdominal pressure. *Ergonomics, 28*, 149-166.

Roebuck, J.A., and Levedahl, B.H. (1961). Aircraft ground emergency exit design considerations. *Human Factors, 3*, 174-209.

Rohmert, W. (1973). Problems in determining rest allowances. *Applied Ergonomics, 4*, 91-95.

Rubin, T. and Thompson, S. J. (1981). The effect of restricted workspaces upon maximal exertions in the sagittal plane. *Ergonomics, 24,* 240.

Shepherd, W. T., Johnson, W. B., Drury, C. G., Taylor, J. C. and Berninger, D. (1991). *Human factors in aviation maintenance phase 1: Progress report.* (Report No. DOT/FAA/AM-91/16). Springfield, VA: National Technical Information Service.

Stalhammer, H. R., Leskinen, T. P. J., Kuorinka, I. A. A., Gautreau, M. H. J. and Troup, J. D. G. (1986). Postural, epidemiological, and biomechanical analysis of luggage handling in an aircraft luggage compartment. *Applied Ergonomics, 17,* 177-183.

Sutherland, V.J., and Cooper, C.L. (1988). Sources of work stress. In J.J. Hurrell, L.R. Murphy, S.L. Sauter, C.L. Cooper (Eds.), *Occupational Stress - issues and developments in research.* London: Taylor and Francis.

Van Wely, P. (1970). Design and disease. *Applied Ergonomics, 1*, 262-269.

Westgaard, R.H., and Aaraas, A. (1984). Postural muscle strain as a causal factor in the development of musculoskeletal illnesses. *Applied Ergonomics, 15*, 162-174.

Wickens, C.D. (1992). *Engineering Psychology and Human Performance* (2nd ed.). New York: Harper Collins.

Wiener, E. L. (1975). Individual and group differences in inspection. In Drury, C. G. and Fox, J. C. (Eds.), *Human reliability in quality control.* London: Taylor & Francis, 19-30.

Wiker, S.F., Langolf, G.D., and Chaffin, D.B. (1989). Arm posture and human movement capability. *Human Factors, 31(4)*, 421-441.

Zhang, L., Drury, C. G. and Woolley, S. M. (1991). Constrained standing: Evaluating the foot/floor interface. *Ergonomics, 34*, 175-192.

Chapter Six
Evaluating the Visual Environment in Inspection: A Methodology and a Case Study

6.0 INTRODUCTION

Visual inspection accounts for almost 90% of all inspection activities; thus, it is imperative that the task be performed in the most suitable work environment. Studies in aircraft inspection have shown that poor illumination, glare and other adverse lighting conditions could be important reasons for "eye strain" or visual fatigue. Visual fatigue causes a deterioration in the efficiency of human performance during prolonged work. The purpose of this study is to develop a methodology which allows adequate lighting equipment to be selected in order to provide an improved visual environment.

Much of the recent literature on lighting requirements is concerned with costs of providing the light, whether purchase costs, operating costs or maintenance costs. However, the purpose of lighting is to allow rapid and effective human performance. The costs of personnel time and the potential cost of even a single human error are orders of magnitude higher than the costs of providing the lighting. Thus, in this study, adequacy of lighting is the major criterion for lighting choice.

The sections below provide an outline of the sequence of steps which were followed to demonstrate and ultimately comprise the advanced methodology. Initially, the basic principles of lighting and lighting system design are related to aircraft inspection. Thereafter, through site visits, the existing visual environment in aircraft inspection is assessed. An evaluation was then undertaken at a single facility in order to acquire detailed data and to demonstrate how to perform a human factors investigation of a visual environment. This investigation included photometric evaluations of the ambient and task lighting as well as input from inspectors at four different facilities.

Concurrently, alternative portable and personal lighting sources were evaluated at the same facility and in the laboratory. Recommendations are then offered based upon the information obtained. This step illustrates the utility of using an organized approach to structure the various components which comprise a visual environment in order to allow adequate light sources to be suggested. Finally, the methodology which encompasses all the preceding steps is formally advanced.

6.1 LIGHT CHARACTERISTICS/LIGHTING SYSTEM DESIGN

Four fundamental light characteristics (i.e., light level, color rendering, glare and reflectance), the principles of specialized lighting, and the basic requirements of lighting design need to be considered in relation to aircraft inspection.

6.1.1 Light Level

The recommended illumination depends upon the type of task and whether the visual task is of high or low contrast. General lighting requirements for different tasks can be found in Eastman Kodak (1983) and Illuminating Engineering Society (IES) (1987). Vision can be improved by increasing the lighting level, but only up to a point, as the law of diminishing returns operates. Also, increased illumination could result in increased glare. Older persons are more affected by the glare of reflected light than younger people, and inspectors are often senior personnel within a maintenance organization.

According to IES (1987), direct, focussed lighting is recommended for general lighting in aircraft hangars. Inspection of aircraft takes place in an environment where reflections from airplane structures can cause glare so that low brightness luminaries should be installed. Often, additional task lighting will be necessary when internal work, or shadowed parts around the aircraft, result in low illumination levels.

Table 6.1 presents the required illumination levels for aircraft maintenance and inspection tasks (IES, 1987). Generally, most maintenance tasks require between 75 foot-candles (f-c) and 100 f-c, although more detailed maintenance tasks may require additional illumination. General line inspections (e.g., easily noticeable dents) may only require 50 f-c; however, most inspection tasks demand much higher levels. From the site observations of actual defects, it is apparent that many difficult inspection tasks may require illumination levels up to or exceeding 500 f-c. Based upon the current IES standards, it is recommended that the ambient light level in a maintenance hangar be at least 75 f-c in order to perform pre- and post-maintenance/inspection operations and some general maintenance/inspection tasks without the necessity for additional task lighting. Furthermore, adequate illumination levels may be obtained in a majority of inspection tasks and many maintenance tasks through the utilization of task lighting.

Table 6.1 Levels of Illumination Required in Aircraft Inspection/Maintenance (IES, 1987)

TASK	F-C
Pre-/post-maintenance and inspection	30-75
Maintenance	75-100
Inspection	
Ordinary	50
Detailed	100
Fine	200

6.1.2 Color Rendering

Color rendering is the degree to which the perceived colors of an object illuminated by various artificial light sources match the perceived colors of the same object when illuminated by a standard light source (i.e., daylight). The color rendering of task lighting is important for inspection because "change in color" of sheet metal is often used as a clue to detect corrosion, wear or excessive heating. The difference in the spectral characteristics of daylight, incandescent lamps, fluorescent lamps, etc., have a large effect on color rendering. Such effects are described in detail in IES (1984). **Table 6.2** presents some of the commonly used lighting sources and their characteristics (adapted from Eastman Kodak, 1983).

Table 6.2 Commonly Used Lighting Sources

TYPE OF LIGHT SOURCE	COLOR	COMMENTS
Incandescent	good	Commonly used, but prone to deterioration over time. High energy lost, but convenient and portable. Lamp life about 1 year.
Fluorescent	fair to good	The efficiency and color rendering capabilities vary greatly depending upon tube type. Problems of flicker may have an annoying effect while performing inspection. Can be dangerous with rapidly cycling machinery. Lamp life 5-8 years.
Mercury Vapor	very poor to fair	Green/blue colors are unusual, and output drops rapidly with age. Lamp life 9-12 years.
High Pressure Sodium Lamp	fair	Monochromatic yellow light. High efficiency lamp ranging from 80-100 lumens per watt. Lamp life 3-6 years.
Low Pressure Sodium Lamp	poor	Highly efficient light source but yellow in color. Lamp life 4-5 years.

6.1.3 Glare

Direct glare reduces an inspector's ability to discriminate detail and is caused when a source of light in the visual field is much brighter than the task material at the workplace. Thus, open hangar doors, roof lights, or even reflections from a white object such as the workcard can cause glare. Glare can also arise from reflections from the surrounding surfaces and can be reduced by resorting to indirect lighting. The lighting system should be designed to minimize distracting, or disabling glare, using carefully designed combinations of area lighting and task lighting.

6.1.4 Reflectance

Every surface reflects some portion of the light it receives as measured by the surface reflectance. High reflectance surfaces increase the effectiveness of luminaires and the directionality of the illumination. Specular, or mirror-like, reflectance should be avoided as it produces glare. Diffuse reflection, for example, from a semi-matte surface is preferred. Thus, for an aircraft hangar, it is important that the walls and floors are of high diffuse reflectance (i.e., light paint, patterned plastics) so that they help in reflecting light and distributing it uniformly. This is more critical under the wings and fuselage where there may not be adequate lighting, due to aircraft shadows. **Table 6.3** presents recommended surface reflective values to assist in obtaining an adequately uniform visual environment.

Table 6.3 Recommended Diffuse Reflective Values (Adapted from IES, 1987).

SURFACE	REFLECTANCE
Ceiling	80 to 90%
Walls	40 to 60%
Equipment	25 to 45%
Floors	not less than 40%

6.1.5 Specialized Lighting

During visual inspection of an aircraft fuselage the inspector is looking for multiple defects, including corrosion, ripples, hairline cracks in the metal components, dents in the fuselage, missing rivets, damaged rivets ("pooched," "dished" rivets), and rivet cracks.

It is possible that no one single lighting system is suitable for detecting all defects. Therefore, the use of specialized lighting systems which make each class of defect more apparent may be necessary. However, the use of special light systems implies that the area must be examined for each class of defects sequentially rather than simultaneously, which could involve time and expense. For example, the diffused nature of general illumination tends to wash out the shadows while surface grazing light relies upon showing shadows to emphasize objects that project above or below the surface. Task visibility is distinctly better for surface topography with grazing light even though a lower level of illumination is used. An example of this scenario is the inspection of the fuselage for ripples. Ripples are easier to detect using surface-grazing lighting because general illumination tends to wash them out. However, normal-incidence lighting may mask important textural and color differences. The lighting should be compatible with the visual objective regarding the form and texture of the task object. Grazing light reinforces an impression of the texture while normal incident light allows the discrimination of color and surface, but minimizes the perception of surface variations.

6.1.6 Design Requirements For Lighting

Literature on visual search has shown that the speed and accuracy with which the search process can be accomplished is dependent on the conspicuity of the defect which in turn is dependent on size of the defect, defect/background contrast, and lighting intensity (Drury and Fox, 1975).

Lighting design also has broader requirements to fulfill. In order for the inspection to be successful, the lighting should be such that the following tasks can be performed satisfactorily and preferably optimally: inspecting (visual search) the aircraft structure for defects, reading the workcard/instructions, moving around the aircraft (using the scaffolding, or equipment, e.g., cherrypicker), and special purpose lighting should not interfere with any other parallel task (e.g., access or maintenance) in progress.

The inspection task is frequently difficult because of the heavy perceptual load present. In designing the lighting system, the objective must be to reduce visual fatigue caused by poor illumination and poor contrast. In designing lighting systems, one must consider the minimum lighting requirements for each task and subtask, the type of artificial light sources that can be used to illuminate the work surface, the amount of task lighting that can be provided and the available methods to minimize glare. These factors must be balanced with implementation and operating costs (IES, 1987); however, the total cost of installing, running and maintaining lighting is a small fraction of the cost of either the employment of personnel or of rectifying lighting-induced human errors.

6.2 THE EXISTING VISUAL ENVIRONMENT IN AIRCRAFT INSPECTION

6.2.1 Classification Of Light Sources

The lighting sources employed in aircraft inspection include *ambient lighting* which is comprised of daylight, area and specialized lighting (built into aircraft); and *task lighting* which includes portable lighting (set up at inspection site) and personal lighting (e.g., flashlight). The ambient lighting represents the minimum lighting level available in a task while task lighting represents the maximum lighting level, both from lighting devices set up to cover an inspection area, and from personally-carried lighting. Note that to provide adequate lighting for any task it should be possible to reduce glare from ambient lighting and use the task lighting in a focussed manner to illuminate the task without causing unnecessary glare.

6.2.2 Site Observations

In the first phase of this research program many inspection/maintenance sites were visited (Shepherd, et al., 1991). Detailed Task Analyses were performed on numerous inspection activities, resulting in a list of examples of poor human factors design. Each example represents an opportunity to improve the human/system fit, and hence, increase job performance with decreased work stress.

The conclusions to be drawn from these observations are that ambient lighting in some cases can range from inadequate to poor for performing inspection tests, which could result in visual fatigue and deterioration of performance. Moreover, task lighting was not adequate, lighting equipment was not always portable, and the lighting level was well below the IES recommended level of 75-100 f-c in most visual aircraft inspection tasks (IES, 1987). These conclusions are substantially the same as found by the FAA's Office of Flight Standards Aging Fleet Evaluation Program which measured the visual environment at nineteen sites performing "D" checks (Thackray, 1992).

6.3 EVALUATION OF A VISUAL ENVIRONMENT

As a demonstration of how to perform a human factors study of lighting in a facility, an investigation of the visual environment at a representative maintenance hangar was performed. The hangar was due for closure, so that findings would be applied only in other hangars. This study included an evaluation of the ambient lighting, task lighting, and perceived lighting characteristics based upon input from inspectors.

6.3.1 Evaluation Of Ambient Illumination, Luminance, and Reflectance

The evaluation measured the illumination and luminance levels produced by the ambient light sources only. Lighting characteristics of the personal and portable lighting were considered separately. Procedures were performed according to the IES Lighting Handbook (IES, 1984). The illumination levels indicate the amount of light falling over the area (in f-c), while luminance levels represent the quantity of light reflected off the various surfaces (in foot-lamberts (f-l)).

The illumination and floor luminance levels were obtained in two different aircraft bays, bay #1 (with an aircraft present) and bay #2 (without an aircraft present). Each bay area was divided into zones and several readings were taken within each zone at night with the hangar doors closed. Average illumination and luminance values were calculated by aircraft area (**Figure 6.1**). Floor reflectance values, the amount of light reflected off the floor compared to the amount of light falling on the floor, (i.e., floor luminance/illumination) were calculated and given in **Figure 6.2.**

The average illumination levels varied dramatically between areas. **Figure 6.1** indicates that the areas under the fuselage and wings had considerably lower illumination than the open areas (i.e., where no aircraft was present). This is a concern, for many visual inspection tasks occur in these poorly lit areas (i.e., under the wings and fuselage). The floors are presently a natural grey color (cement), thus resulting in low average floor luminance and reflectance levels across all areas. The floors should be painted a lighter color (e.g., white), which would improve the overall illumination levels, especially under the wings and fuselage. However, any paint used should be non-glossy to eliminate specular reflections from the floor surface. For new hangars, or major renovations, lighter colored flooring could be installed.

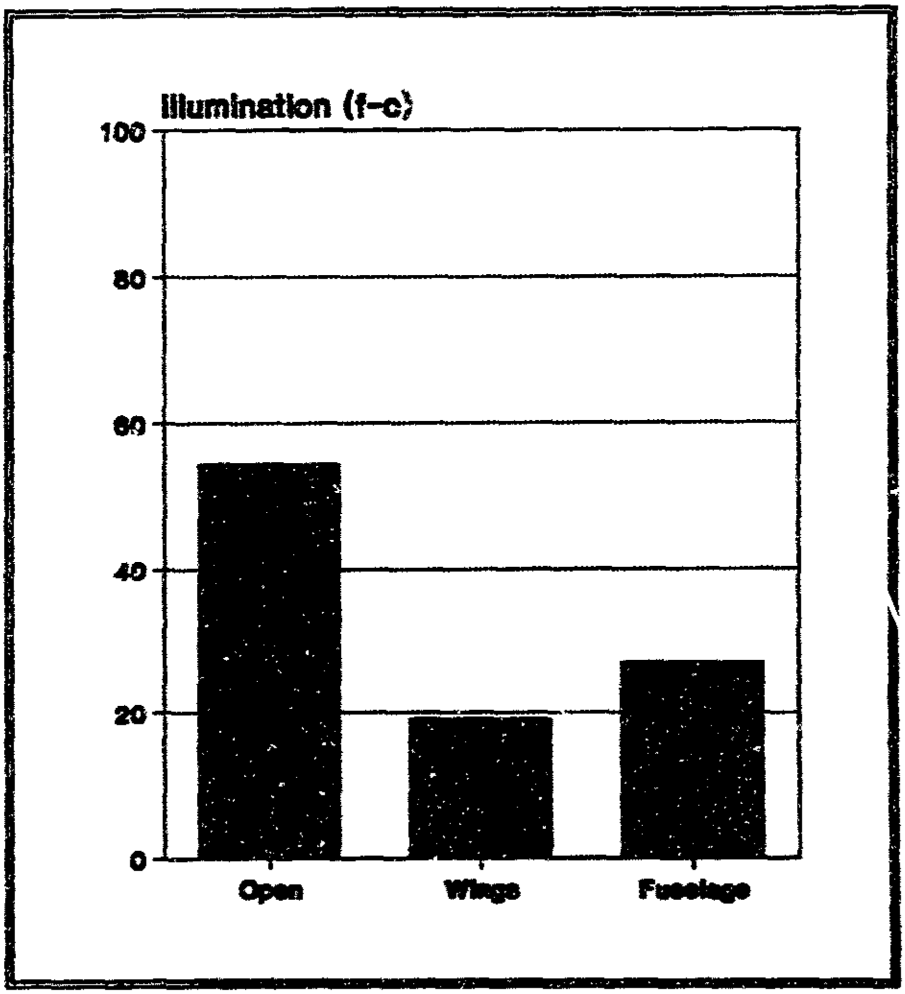


Figure 6.1 Illumination by Aircraft

6.3.2 Evaluation Of Task Illumination, Luminance, and Reflectance

A representative sample of aircraft visual inspection tasks was selected from various locations on a Fokker F-100: air conditioning access (A/C), cargo compartments (cargo), exterior fuselage-nose, nosewell, and wheelwell. A lighting evaluation (i.e., illumination, luminance, and reflectance levels) was performed with the results shown in **Table 6.4**. The light environment for each task includes the contribution of the ambient levels in conjunction with any additional task lighting.

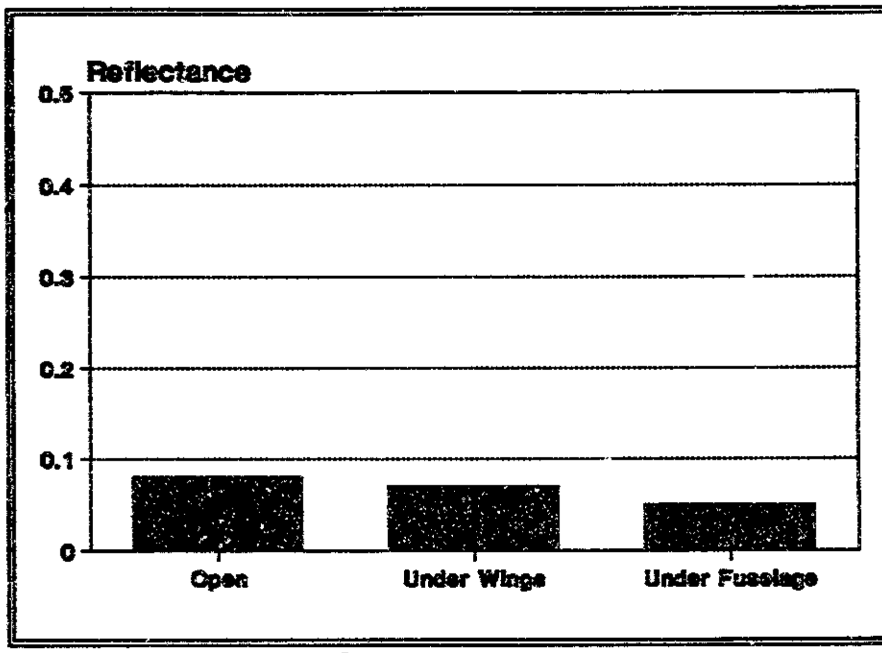


Figure 6.2 Perceived Glare by Aircraft Area

Values were obtained from various locations in each task area under actual inspection conditions; that is, while the task lighting of choice (i.e., personal/portable) was utilized. Generally, the *average* task illumination levels were adequate, with the exception of the nosewell. However, large variabilities existed in these levels within each area, primarily dependent upon whether it was possible to aim the lighting equipment at the point of inspection. In many instances, areas were difficult to access with the lighting equipment, thus not allowing adequate levels of light. Task lighting was necessarily the primary light source in all task areas, for the ambient illumination levels were inadequate. Thus, the accessibility of the area and the portability of the task lighting affected the light level at a majority of inspection points.

Table 6.4 Task Light Environment and Illumination by Task Area

AREA	TASK CONDITIONS	
	Light Environment	Illumination (f-c)
A/C	3D-Cell Flashlight	X = 115 SD=138.0
Cargo	Headlamp plus General	182 SD=72.4
Fuselage-Nose	3D-Cell Flashlight plus General	97 SD=46.0
Nosewell	3D-Cell Flashlight plus General	42 SD=22.6
Wheelwell	3D-Cell Flashlight plus General	102.1 SD=40.0

6.3.3 Inspector Perceptions

In addition to the detailed measurements obtained at one facility, inspectors' perceptions of the visual environment were assessed at several other facilities within the partner airline. Psychophysical rating was obtained from 51 inspectors and maintenance personnel from four other sites, to allow a detailed assessment of the perceived quantity and quality of the general and task lighting. Inspectors and maintenance personnel were asked to evaluate the lighting characteristics of the visual environment (e.g., contrast, glare, flicker, color rendering), as well as the adequacy of the lighting equipment (e.g., ease of handling, light level and focus control).

Psychophysical rating was obtained on the visual environment and combined by aircraft area: upper exterior areas (above wing chord line), lower exterior areas (below wing chord line), and interior areas. Generally, according to the frequency distributions, the perceived light levels and contrast ranged from adequate to good in the upper exterior areas, but there were many instances of perceived glare. Conversely, the perceived light levels and contrast were frequently rated as inadequate in the lower exterior and interior areas, but there was less perceived glare (**Figures 6.3, 6.4 and 6.5**). Color rendering was perceived to be adequate by most personnel, although this distrib ıtion was skewed towards inadequate in the lower exterior and interior areas (**Figure 6.6**).

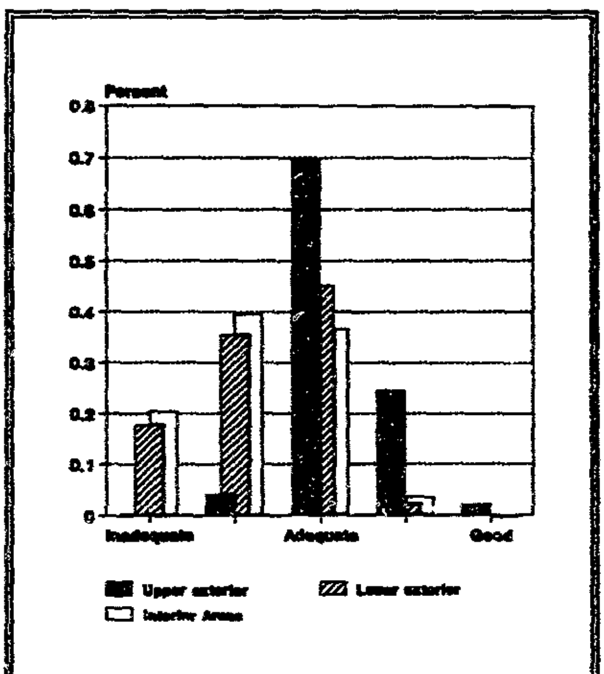


Figure 6.3 Perceived Light Level by Aircraft Area

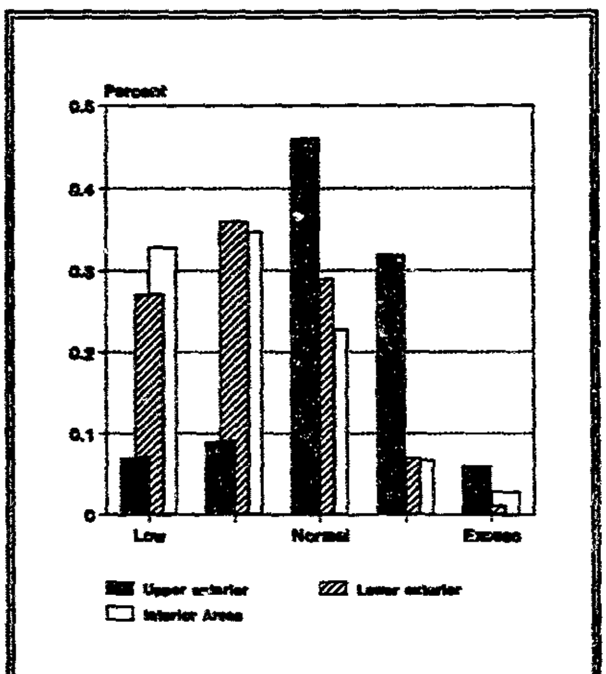


Figure 6.4 Perceived Contrast by Aircraft Area

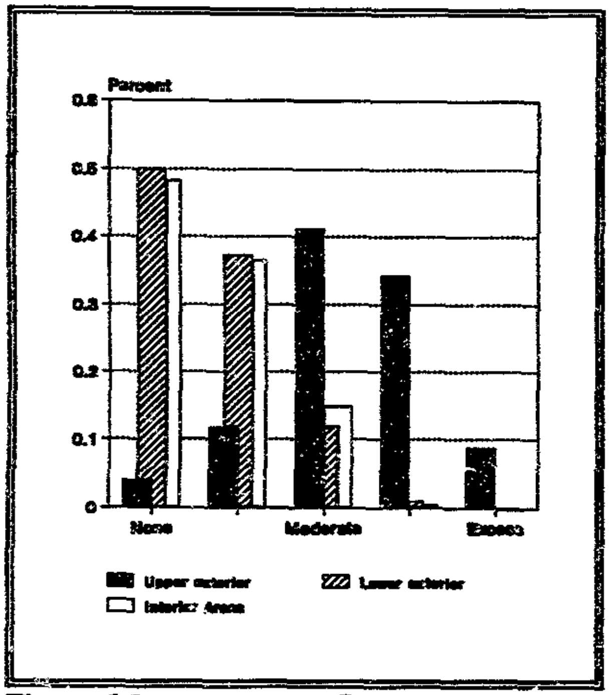


Figure 6.5 Perceived Glare by Aircraft Area

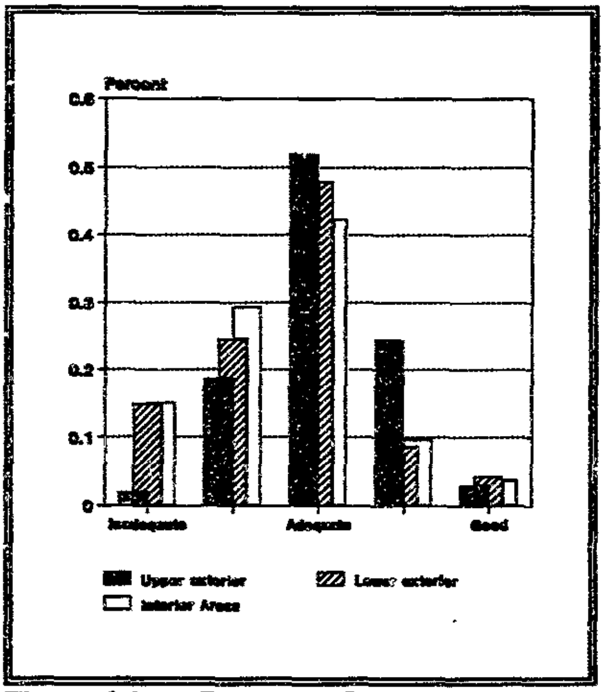


Figure 6.6 Perceived Color Rendering by Aircraft Area

In the upper exterior areas, a majority of personnel indicated a reliance on primarily general lighting (over 90%), with a smaller dependence on daylight and personal lighting (Figure 6.7). Portable lighting seems to be rarely used. In contrast, in the lower exterior and interior areas, personal lighting is the primary light source (over 90%), with general and portable lighting being utilized somewhat. Daylight contributes minimally to the visual environment in the lower exterior and interior areas. This is presumably the reason why color rendering was perceived to be worse in these areas for artificial light is the primary source.

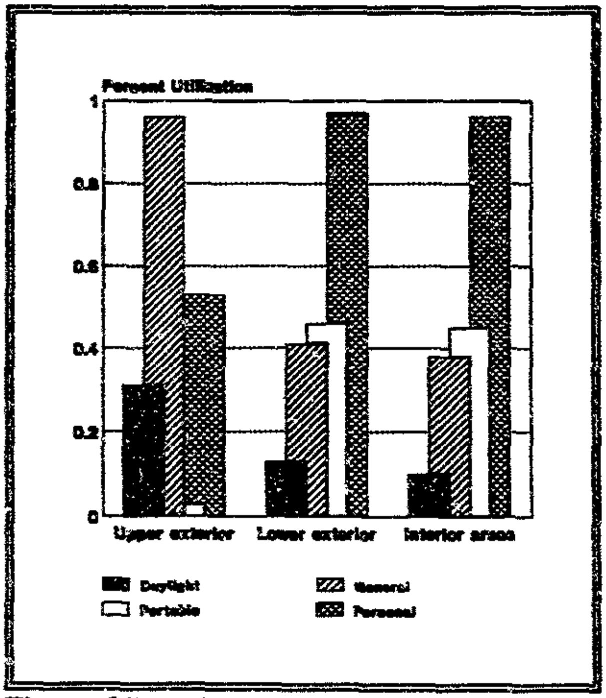


Figure 6.7 Lighting Utilization by Aircraft Area

A majority of personnel indicated that both personal and portable lighting equipment produce adequate light levels. There were varied perceptions with respect to handling, although a majority felt personal lighting was adequate and portable lighting was inadequate. Likewise, a majority of personnel feel the focus ability of personal lighting was good, while the aiming ability of portable lighting was inadequate.

These perceptions may indicate why personal lighting is relied on more than portable lighting (Figure 6.7); it is easier to handle and control. A need exists for better portable lighting to decrease reliance on personal lighting in restricted spaces.

Finally, general comments and concerns related to personal and portable lighting systems and the visual environment were obtained. The comments are ranked according to the frequency with which inspectors and maintenance personnel indicated the importance of the various factors (Table 6.5). The major considerations fall within the categories of lighting, ease of handling, durability, work shift, hangar maintenance, flexibility, and miscellaneous attributes.

Light output/brightness of the visual environment was the biggest concern of personnel. The ability to control the light output to reduce glare were also of particular concern. Color rendering and contrast were of lesser importance. A few personnel indicated the need to investigate alternative light sources (e.g., lasers). Surprisingly, flicker normally associated with fluorescent lights was not indicated as a major concern, although fluorescent lighting was not the primary light source utilized across the population sampled. Specific to personal lighting, personnel indicated the effect of low-quality batteries and bulb type on the quality of light.

The ease with which equipment was handled was of particular concern. Personnel indicated that light sources which are difficult to use will not be utilized, obviously affecting the visual environment. The weight/size of the lighting equipment was found to be the primary determinant of ease of handling, with accessories a secondary factor. The set-up required for portable lighting was found to directly affect utilization. Electric-powered personal lighting, as opposed to battery-powered, was found to be less portable due to the need for power cords. Power sources, such as battery-packs, may be a promising alternative for improving the portability of electric-powered lighting equipment.

The general durability (i.e., daily wear and tear) of lighting equipment was a consideration, for it affects the visual environment. For example, personnel indicated that if their flashlight lens was scratched, the light output decreased. The safety requirements met by the lighting equipment [e.g., Occupational Safety Health Administration, (OSHA)] is another issue which was indicated, and becomes critical in hazardous areas (e.g., fuel tanks). Lighting equipment which meets specialized safety requirements, as evaluated in this study (e.g., explosion/vapor-proof), needs continuous investigation to ensure compliance with changing standards.

Personnel indicated that the workshifts (i.e., day/night) resulted in drastically different visual environments, possibly dictating different lighting needs. General hangar maintenance can also affect the visual environment. For example, as discussed earlier, light paint on the floor, walls, and ceilings causes light to be reflected and creates a brighter work environment. Furthermore, these surfaces, in addition to the lights themselves, must be free of dirt and grime in order to reflect/produce adequate light.

Often the flexibility of light sources is important for performing inspection, particularly with respect to specific task demands and fault types (e.g., light of grazing incidence may be necessary to highlight ripples while light perpendicular to the surface may be necessary for detecting other common faults). Flexibility is more easily provided by personal lighting equipment (e.g., flashlight and headlamps) rather than by portable and direct lighting which are more suited to meet general lighting requirements.

Table 6.5 Relevant Lighting Considerations Based Upon Inspector Perceptions

CHARACTERISTICS	PERSONAL	PORTABLE	FREQUENCY
LIGHT	1. Output/brightness	1. Output/brightness	27
	2. Glare/brightness control	2. Glare/brightness control	6
	3. Distribution/focus	3. Distribution/aim	5
	4. Color rendering	4. Color rendering	1
	5. Contrast	5. Contrast	2
	6. Alternative sources	6. Alternative sources	2
	7. Flicker	7. Flicker	1
	8. Power source (battery type)		4
	9. Bulb type		3
			Total 51
EASE OF HANDLING	1. Weight/size	1. Weight/size	15
	2. Accessories	2. Accessories	2
		3. Set-up	3
	3. Power source		2
			Total 22
DURABILITY	1. General	1. General	5
	2. Safety requirements	2. Safety requirements	5
	3. Bulb life	3. Bulb life	0
	4. Battery life		8
			Total 18
WORK SHIFT	1. Light (day/night)	1. Light (day/night)	9
	2. Shiftwork	2. Shiftwork	2
			Total 11
MAINTENANCE	1. Paint	1. Paint	7
	2. Hangar cleanliness	2. Hangar cleanliness	2
			Total 9
FLEXIBILITY	1. Task demands	1. Task demands	2
	2. Fault types	2. Fault types	2
			Total 4
OTHER ATTRIBUTES	1. Cost	1. Cost	5
	2. Access	2. Access	4
	3. Availability	3. Availability	2
	4. Individual differences	4. Individual differences	1
			Total 12

Items such as cost were indicated to be a concern. For example, several personnel indicated that rechargeable flashlights are superior to other types of less expensive lights. Safe disposal of used batteries may become an increasingly important reason for choosing rechargeable lighting systems. In many situations personnel purchase their own equipment, as opposed to using less adequate equipment supplied by the company.

As discussed above (Section 6.3.2), access to an area can dictate the quality and quantity of light on a surface. There should be an effort not only to improve the portability of lighting equipment, but more importantly, to consider the human inspection process at the aircraft design stage. The availability of various lighting sources was a concern, and may be dependent upon the company supply, hangar design (e.g., availability of electric outlets around the aircraft), or accessibility. Finally, common to all inspection tasks, individual differences must considered.

6.4 EVALUATION OF ALTERNATIVE LIGHTING SOURCES

An evaluation of lighting sources was performed to identify systems which possess features which may contribute to the existing visual environment of aircraft inspection/maintenance operations. This evaluation included an investigation of available systems, and both laboratory and field evaluation of the selected sources.

6.4.1 Laboratory Evaluation

A number of both personal and portable lighting systems was selected to represent the types currently being used in inspection and alternative sources available in catalogs. Several attributes of these selected personal and portable lighting systems were investigated in a controlled environment (i.e., light source, weight, focus/aiming control, durability, safety requirements, accessories, and light output/distribution). The results of this investigation can be found in Reynolds, et al., 1992.

6.4.2 Field Evaluation

A sample of the lighting systems which appeared to hold promise in the laboratory evaluation are presented in **Tables 6.6** and **6.7**, and were further investigated during actual task performance.

Table 6.6 Specifications of Selected Portable Lighting Equipment (center illumination measured at: *0.5m or **2.0m)

Light Source		Wt. (lbs.)	Aiming Control	General Durability	Safety Requirements	Accessories
Handlamp	Fluorescent 13 Watts 85 f-c*	0.8	No	Adequate	NEC #410-35(a)(b),36, 42(a)(b),44,45 OSHA #1926.405(E)(F)(I)(j)(ii)(iii)(A)(B)(V) UL listed	Hook
Portable Lamp	Fluorescent 27 Watts 164 f-c**	10	No	Adequate	NEC #410-35(a)(b),36, 42(a)(b),44,45 OSHA #1926.405(E)(F)(I)(j)(ii)(iii)(A)(B)(V) UL listed	Hangars, Magnets
Standing Lamp	Halogen 500 Watts 1200 f-c**	9	Yes	Adequate	UL listed for indoor/outdoor use.	Handle Stand

Table 6.7 Specifications of Selected Personal Lighting Equipment (*center illumination measured at 0.5m)

Light Source		Wt. (lbs.)	Focus Control	General Durability	Power	Safety Requirements	Accessories
2D-Cell flashlight	Krypton Bulb 340 f-c*	1.5	Yes	Adequate	Battery	Explosion-proof MIL-STD-810C	None
3D-Cell flashlight	Krypton Bulb 1100 f-c*	2.0	Yes	Adequate	Battery	Explosion-proof MIL-STD-810C	None
4D-Cell flashlight	Krypton Bulb 1900 f-c*	2.4	Yes	Adequate	Battery	Explosion-proof MIL-STD-810C	None
Headlamp	Incan-descent Bulb 1500 f-c*	1.9	Yes	Adequate	Battery	None indicated	Headband Battery clamp

The following provides a summary of the results obtained from the laboratory and field evaluations.

• There are two different kinds of lights: inspection and work lights. Inspection lights (i.e., dynamic sources) must provide easy handling, for inspection normally demands frequent movement in and around the aircraft. In addition, the lights must provide a focussed beam of light which can be controlled to reduce glare. Work lights (i.e., static sources) need not be as portable as inspection lights, for they are normally used in one place for a period of time (i.e., generally 30 minutes or more).

• The flashlights provide adequate light, durability, and focus control to reduce glare. They are also easily portable, which suits most inspection tasks. The light outputs and distributions of the flashlights increase with the size of the light (i.e., 2D to 4D). The larger lights have more batteries; however, they are also heavier. The focus ability of the flashlights provides either an intense focussed beam or less illumination over a larger area.

• The headlamp provides adequate light and focus control to reduce glare. It produces a comparable amount of light as the 4D-Cell flashlight, although it is lighter and allows hands-free portability. However, it meets no additional safety requirements, thus possibly limiting its use in some environments. The actual weight of the lamp is less than the indicated weight, for the batteries are separated from the light source (0.3 lbs.).

• The handlamp is not well suited for many inspection tasks because the power cord reduces its portability and it does not provide a highly focussed beam. However, this light can serve as a small portable light source. It produces less light over a smaller area than the other portable lights, but gives off minimal heat and can fit into small access areas. It is very durable and meets OSHA and National Electrical Code (NEC) safety requirements related to general electrical codes.

• The portable lamp is a good static light source. It can be hung, using the provided strap or magnet, or placed (e.g., under a wing) in the work area for overall, heat-free light. Furthermore, these lamps meet OSHA and NEC safety requirements related to general electrical codes.

• The standing lamp provides a large amount of light over a large area. It can be used to illuminate large static work areas. However, it gives off heat, and thus could not be used for interior inspections or in small areas, limiting its use to open, exterior areas. In addition, it is UL listed for indoor/outdoor use, possesses up/down aiming control, is light-weight, and has a handle for easy portability and set-up.

• The color rendering characteristics of the standard incandescent lamps (i.e., headlamp), krypton lamps (i.e., flashlights), and halogen lamps (i.e., standing lamp) are superior. The fluorescent lights generally provide adequate color rendering characteristics, dependent upon the chemical composition of the liner, and are more energy efficient, producing less heat than incandescent lights.

6.5 RECOMMENDATIONS

Based upon the above evaluation of the visual environment at the tested facility and the selected sample of lighting sources, initial recommendations are presented. The task demands, the restrictiveness of the space to be inspected, the ambient light conditions, and the lighting requirements are considered (Table 6.8). Recommendations are advanced for the specific task environments evaluated earlier (Section 6.3.2), and only consider the sample of lighting sources selected for detailed field evaluation (Section 6.4.2). Caution should be exercised in generalizing these recommendations to other task situations and light sources, although the methodology presented here can be used to determine the applicability of each light source in new situations. As discussed previously (Section 6.3.2), the ambient illumination levels in all the task areas were inadequate for satisfactory performance. Thus, there must be some reliance on personal or portable lighting in each area.

For each task area, the task demands dictate the required illumination, the focus/aiming, and the required handling. A majority of inspection tasks require dynamic sources, to allow for frequent movement in and around the area; whereas, maintenance tasks may be adequately illuminated by static sources. Although inspection tasks are the primary focus in this study, recommendations will also be made for static sources for they can be useful in contributing to the ambient light level in many areas.

Based on the task demands and corresponding illumination requirements, it is observed that each of the recommended lighting systems meets the illumination requirements as observed in **Table 6.8** and **Table 6.6** through **6.7**. Furthermore, personal lighting not only provides the necessary illumination but also greater flexibility in terms of maneuvering the light source (e.g., a flashlight can be used both at grazing incidence to detect ripples and at normal incidence to study corrosion).

The restrictiveness of the area to be inspected was rated on a two-point scale (i.e., restricted/unrestricted). Thus, restricted areas (i.e., A/C access, nosewell, and wheelwell) require the light source to be manipulated around obstructions in a cramped area, in order to provide an adequate visual environment. As was discussed previously (Section 6.3.2), large variability existed in the light levels in these areas, dependent upon the accessibility of the light source. The cargo and exterior fuselage-nose areas are considered unrestrictive, for the light source is not obstructed by the environment. Any size personal lighting may be used here without compromising the visual environment.

The inspection of the A/C access area requires a dynamic light source which possesses focus and aiming ability and provides an average level of 100 f-c of illumination. The 2D-Cell flashlight and the headlamp are recommended for they meet these requirements and are small enough to be manipulated around the area. In addition, the handlamp is recommended as a static light to increase the general light level in the area in order to reduce the reliance on personal lighting. It is small and can be hung or placed in the area, and does not give off heat.

Similarly, inspection of the cargo and exterior fuselage-nose areas also require dynamic light sources with easy controllability. In addition, an average illumination level, when combined with the ambient light level, of 100 f-c in the cargo area and 200 f-c in the exterior-nose area is required. The areas are not restricted, thus any size flashlight or the headlamp could be used as a personal lighting source. The standing lamp could be aimed up from the outside of the aircraft, or the portable lamp could be hung/placed in the area, to provide overall light.

Finally, inspection of the nosewells and wheelwells requires dynamic, focussed average illumination levels of 100 and 200 f-c, respectively. The areas are somewhat restrictive, thus requiring the smaller flashlights or headlamp for better handling. The handlamp and portable could be hung/placed in tight locations in these areas, while the standing lamp could be aimed up into these areas for general overall lighting.

Table 6.8 Lighting Recommendations for Various Tasks Based Upon the Task Demands, the Lighting Requirements, and The Size of the Area

AREA	TASK DEMANDS	SPACE	AMBIENT ILLUM. (f-c)	LIGHTING REQUIREMENTS			RECOMMENDATIONS
				Illum. (f-c)	Focus/ Aiming	Handling	
A/C	Check for leaks and security of items	Restricted	0	100	Yes/ Yes	Dynamic Static	2D-Cell flashlight Headlamp Handlamp
Cargo	Check for security of items, missing fasteners, spillage and corrosion, dents. Check cargo door for dents. Inspect internal brackets, hinge arms, fittings, handle housing. Check if material is lodged in latches or attached to stops. Inspect for security of attachments.	Unrestricted	12	100	Yes/ Yes	Dynamic Static	Flashlights Headlamp Standing lamp Portable lamp
Fuselage-Nose	Perform a detailed visual inspection of area looking for dents, corrosion, missing fasteners.	Unrestricted	44	200	Yes/ Yes	Dynamic Static	Flashlights Headlamp Standing lamp Portable lamp
Nosewell	Inspect nose landing gear torque-link center for play. Check for security of items and leaks.	Restricted	2.5	100	Yes/ Yes	Dynamic Static	2D,3D-Cell flashlights Headlamp Handlamp Standing lamp Portable lamp
Wheelwell	Inspect main landing gear, landing gear assembly, for corrosion and cracks. Inspect for security of joints, safety pin for shear, hinges of door for wear and play.	Restricted	3	200	Yes/ Yes	Dynamic Static	2D,3D-Cell flashlight Headlamp Handlamp Standing lamp Portable lamp

6.6 GUIDE FOR VISUAL ENVIRONMENT EVALUATION

A methodology by which to evaluate and design a visual environment may be advanced based upon the techniques employed in the above demonstration project. A four-step methodology is presented below.

1. *Evaluate existing visual environment.* The first step requires an investigation of the visual environment in order to obtain an understanding of the existing conditions and to focus the investigation on problem areas. Ambient and task lighting conditions and task analyses should be performed in order to determine the task demands and associated visual requirements. In addition, personnel should be consulted to obtain additional information regarding the light characteristics and utilization and adequacy of the currently used lighting sources.

2. *Evaluate existing and alternative lighting sources.* An evaluation of the existing and alternative lighting sources is performed in order to identify the capabilities of each source. Manufacturers' catalogs can be consulted to determine the current status of lighting source technology. These alternative sources, in addition to the sources currently being used, can be evaluated. Evaluations performed to date, including the present one, have used various criteria to judge visual environments (e.g., light output, glare, luminance, etc.). There is a need for standard criteria which allow visual environments in aircraft maintenance/inspection operations to be evaluated in a consistent manner and which insure that important components of the process are not over-looked. An attempt has been made to identify the most important components which need to be considered in the evaluation of an aircraft inspection/maintenance visual environment. Considering the operator perceptions and other factors discussed earlier (Sections 6.1 and 6.3.3), a guide has been developed to indicate important considerations in the selection of adequate lighting sources (**Table 6.9**). Requirements are given for both personal and portable lighting.

3. *Selection of lighting sources.* Once steps 1 and 2 are completed, lighting sources can be selected based upon a comparison of the lighting requirements with the various lighting sources. An investigation of the existing visual environment (step 1 above) will allow the determination of the lighting requirements to be based upon the task demands. These results can be directly compared with the capabilities of the various lighting sources (step 2 above), to determine which lighting sources provide the most appropriate visual environment for each task analyzed.

4. *Evaluate and address general visual environment factors.* In addition to attending to the specific task conditions, there are factors relevant to the overall environment which need to be addressed. Based upon the operator perceptions and other factors discussed earlier (Sections 6.1 and 6.3.3), a guide has been developed to indicate relevant considerations in the design of an adequate visual environment (**Table 6.10**). The assessment of these considerations should result in additional improvements in the overall visual environment.

This methodology does not provide guidelines which dictate how to design a visual environment. Instead, it provides a flexible process which may be followed to allow each practitioner to tailor the methodology to meet their individual needs. For example, this demonstration emphasized consideration of lighting requirements, handling, and space restrictions in advancing recommendations. However, dependent upon each facility's needs and associated tasks, other factors identified in this study (steps 1 and 2) may be given stronger consideration (e.g., safety requirements, power sources).

Table 6.9 Lighting Source Design Considerations

CHARACTERISTICS	PERSONAL	PORTABLE
LIGHT	1. Output/brightness 2. Glare/brightness control 3. Distribution/focus 4. Color rendering 5. Contrast 6. Alternative sources 7. Flicker 8. Power source (battery type) 9. Bulb type	1. Output/brightness 2. Glare/brightness control 3. Distribution/aim 4. Color rendering 5. Contrast 6. Alternative sources 7. Flicker
EASE OF HANDLING	1. Weight/size 2. Accessories 3. Power source	1. Weight/size 2. Accessories 3. Set-up
DURABILITY	1. General 2. Safety requirements 3. Bulb life 4. Battery life	1. General 2. Safety requirements 3. Bulb life
FLEXIBILITY	1. Task demands 2. Fault types	1. Task demands 2. Fault types
OTHER ATTRIBUTES	1. Cost 2. Space 3. Individual differences	1. Cost 2. Space 3. Individual differences

Table 6.10 General Visual Environment Design Considerations

CHARACTERISTICS	VISUAL ENVIRONMENT
Light	1. Light level 2. Glare 3. Distribution 4. Color rendering 5. Contrast 6. Flicker
Work Shift	1. Light (day/night) 2. Shiftwork
Maintenance	1. Paint 2. Hangar cleanliness
Other Attributes	1. Access devices 2. Availability of lighting sources

6.7 CONCLUSIONS

This evaluation provides a methodology by which various light sources can be matched to different tasks, based upon consistent criteria. This methodology includes an evaluation of the general and task lighting environments, the task demands, and alternative lighting sources. In addition, the major factors which need to be considered in the design of an adequate visual environment for aircraft inspection are identified in an initial attempt to standardize this evaluation process. The techniques utilized to assess the visual environment at a typical facility may be incorporated into a formal methodology which may be utilized to investigate visual environments and guide selection of lighting equipment at other aircraft inspection sites.

6.8 REFERENCES

Drury, C. G. and Fox, J. G. (Eds.) (1975). *Human reliability in quality control*. London: Taylor and Francis.

Drury, C. G., Prabhu, P. V. and Gramopadhye, A. K. (1990). Task analysis of aircraft inspection activities: Methods and findings. *Proceedings of the Human Factors Society 34th Annual Meeting*, 1181-1185.

Eastman Kodak (1983). *Ergonomic design for people at work* (Vol. 1). Beshmart, CA: Lifetime Learning Publications.

Illuminating Engineering Society (1987). *IES lighting handbook, application volume*. New York: Illuminating Engineering Society.

Illuminating Engineering Society (1984). *IES lighting handbook, reference volume*. New York: Illuminating Engineering Society.

Reynolds, J. L., Gramopadhye, A. and Drury, C. G. (1992). Design of the aircraft inspection/maintenance visual environment. Unpublished final report. (Contract No. DTFA01-92-R-08651). Federal Aviation Administration.

Shepherd, W., Johnson, W. B., Drury, C. G., Taylor, J. C. and Berninger, D. (1991). *Human factors in aviation maintenance, phase 1: Progress report* (Report No. DOT/FAA/AM-91/16). Springfield, VA: National Technical Information Service.

Thackray, R. I. (1992). *Human factors evaluation of the work environment of the operators engaged in the inspection and repair of aging aircraft* (Report No. DOT/FAA/AM-92/3). Washington, D.C.: FAA Office of Aviation Medicine.

Chapter Seven
Design of Workcards

7.0 INTRODUCTION

The workcard is the primary document that controls an inspection task. It has, therefore, a great influence on inspection performance. Costs, due to undetectable faults or faulty detection, when weighed against the cost of providing quality documentation, make a strong case for developing optimum documentation and a methodology (coupled with a set of guidelines) for designing such documentation. This study develops such a methodology, based on the application of human factors knowledge to the analysis of aircraft inspection tasks, and demonstrates its use in two practical applications. The methodology developed, being highly generic, can also be extended for design of information for portable computer-based workcards, as well as hypermedia-based documentation for inspection and maintenance tasks. The project was performed in close cooperation with a partner airline to ensure that the results addressed airline concerns.

7.1 A TAXONOMY OF ISSUES IN DOCUMENTATION DESIGN

A taxonomy for design of usable documentation was developed using the inspection task analysis data from Phase 1 of this program (Shepherd, et al., 1991) and the literature on the human factors of information presentation. This taxonomy has four basic categories of design issues:

1. Information Readability
2. Information Content
3. Information Organization
4. Physical Handling and Environmental Factors

7.2 INFORMATION READABILITY

Information readability is the crux of any visually displayed material. All other issues become meaningful only after this primary issue has been addressed. Design issues affecting information readability are the typographic layout of the information and the language structure, namely, sentences, words and letters.

7.2.1 Typographic Layout

Typographic layout involves the use of vertical spacing, lateral positioning, paragraphing and heading positioning, etc. All the principles of typography cannot be satisfied when the space available is limited. In such cases, the use of secondary typographic and spatial cues becomes essential. Typographic cueing refers to use of variations in the appearance of the text in order to provide a visual distinction, e.g., boldfacing, italics, underlining, color coding, capital cueing, etc. Spatial cueing refers to the spatial layout of the typographic material, e.g., justification of margins, line spacing, etc. Advances in computer technology and word processing provide us with new tools such as full justification of typographic material, which improves reading speed considerably as compared to an irregular margin (Campbell, Marchetti and Mewhort, 1981).

7.2.2 The Sentence, the Word and the Letter

The arrangement of print on paper supplies information about sentence, word and paragraph boundaries. Every printed language has conventions familiar to readers, and disruption of reading results when these conventions are violated (Haber and Haber, 1981). This finding suggests that readers routinely use print arrangement as a source of visual information. In addition to the context, the shape alone of the word itself may prove to be useful in word recognition/ identification. Carroll, Davies and Richman (1971) demonstrated this using very high frequency words from text (e.g., "the," "and," "it"). However, when the text is presented in all capitals, little or no word shape information is present, indicating a waste of an information resource. Since words are basically composed of letters, each of which has a distinct identity and name, a part of the visual information in reading must include the visual features of the individual letters of the alphabet. Based on the feature description models, the entire English alphabet can be described by a total of eight feature descriptions (Haber and Haber, 1981). Type faces like Helvetica have no irrelevant features for visual processing where as type faces like Times have redundant features like serifs which need additional processing.

7.3 INFORMATION CONTENT

Information content refers to issues like origin, appropriateness, accuracy, completeness and comprehension of both textual as well as graphical information. The workcard designer has to understand the way in which information on the workcard is going to be used and the influence it will have on user strategies. Two of the more important issues in this area concern the appropriate information content and the presentation of graphic information.

7.3.1 Appropriate Information Content

To reduce and eliminate user strategy biases and consequently improve the usability, the information should incorporate the following qualities (Swander and Vail, 1991):

- It should be accurate.
- It should be complete, including information regarding: What is to be done, where, how, in what sequence, which specific items to pay attention to.
- It should be up to date with revisions and updates.
- It should be easy to use and comprehend.
- It should be written in a consistent and standardized style and syntax.
- It should be clear and unambiguous.
- It should be specific and contextual, e.g., pertaining to the particular aircraft being inspected.
- It should be flexible, i.e., to support both the expert as well as the novice user.
- It should use only approved and proper acronyms.
- It should have logical and uncontradictory statements.

7.3.2 Graphic Information

Plain text can be uninviting to read and can, at other times, involve high cognitive costs of interpretation. The same objective can be achieved at lower cognitive costs by use of graphic information provided that the graphic information is designed and presented in an appropriate manner. At times textual information becomes difficult to comprehend, especially when conveying spatial information. In such cases, graphics can present the information more clearly. However, high fidelity graphics can involve high cognitive costs of interpretation and may have negative effects due to clutter. Hence, items not relevant to the task should be eliminated to avoid clutter.

7.4 INFORMATION ORGANIZATION

The primary rule of information organization is to classify information into relevant and clearly distinguishable categories (Sutton, 1991). Another important issue is the flexibility of information usage so that information can be used by both the novice as well as the expert. This aspect of information usage has led to the concept of information layering explained in the following sections.

7.4.1 Classification of Information

Information in any workcard can be clearly classified as: directive information (procedures and methods for achieving certain goals), references to additional information, warnings, cautions and notes. These classes of information should follow a standard prioritized order within the document itself, e.g., warning should precede cautions and notes. Since directive information forms the major portion of workcard information, it is explained below in more detail.

Inaba (1991) suggests that directive information should not include more than two or three related actions per step, keeping in mind the limitations of the human short term memory. All directive information can be broken into three logical parts: the command verb, the objects, and the action qualifier. The command verbs used should have no synonyms, to reduce the level of ambiguity. The objects need to be broken down into further subgroups to prevent action slips. The action qualifier should be distinct from the other two, and may begin with a standard article like "for." Given below is the generic format plus a specific example of the three sub-groups differentiated by typeface:

Generic Form		Example	
Command Verb:	*- Object 1*	Check:	- all hydraulic lines
	- Object 2		- control cables
	- Object 3		- pulleys
	for *action qualifier 1,* *action qualifier 2* and *action qualifier 3*		for wear, fraying, damage and corrosion

7.4.2 Information Layering

A novice inspector may require elaborate information at every stage of inspection for an action qualifier; an expert, on the other hand, might require brief information. The information organization should be such that it caters to the needs of both. The prime goal is to make the information more flexible and more context sensitive (Jewette, 1981).

Multiple levels can be built into the information organization, for example, having the main ideas at the first level, followed by elaboration of each of the main ideas at the second level, and finally detailed descriptions at the lowest level. A number of methods can be adopted for presenting multi-layered information in hard copy format: use of distinctly separate layers (e.g., a checklist followed by a detailed information sheet); indented paragraphing (Jewette, 1981); use of color, graphical anchors or boxes; use of different print sizes and styles; use of symbolic nomenclatures, e.g., "A," "B," "1.1," etc. Also, at the lowest level, other tools such as italics, boldface, underlining, brackets, footnotes, appendices, etc., can be used.

In addition to the obvious advantages to the user in terms of flexibility of usage, multi-level writing has some distinct advantages to the writer. It is easy to write, as it has a preset framework within which to write. It is less dependent on fancy phraseology. Sequencing and rearranging of information becomes an easier task, with less planning requirements. The amount of redundancy in the information too is considerably lower. Finally, multi-level writing involves the use of explicit statements of intention in a format dictated by the framework and is hence less error prone.

7.4.3 Other Organizational Issues

Ideally speaking, both text and graphics should be presented on the same page or facing pages, but for reasons of cost effectiveness and system limitations this may not be feasible at all times. The page size should be treated as a naturally occurring module within a document, in the physical sense, i.e., care should be taken to see that each page starts with a new task and that tasks do not carry forward across multiple pages). Each module has all the information necessary to achieve a goal (i.e., completing a task or subtask). Thus, the inspector does not have to read across multiple pages to assimilate the information needed to complete the subtask. The information should be organized according to a rational task order, which may either be the most rational way of doing that task or may be the order followed by most inspectors, due to practical reasons discovered during workcard usage.

7.5 PHYSICAL HANDLING/ENVIRONMENTAL FACTORS

A workcard which satisfies all of the above principles of information design but is not physically compatible with the task at hand will be of little use as people will be reluctant to use it. Handling and usage is a critical factor and will remain so even with automated job-cards using pen-based or laptop computers. Providing a simple workcard holder can at times solve this problem. Depending on the task, however, a specialized design of a workcard holder may be essential to improve the usability of the documentation.

Non-compatibility with the working environment can encompass a number of factors:

- physical handling difficulty due to unwieldy size
- excessively heavy, cannot be held continuously
- environmental degradation due to wind, rain and snow
- incompatible with the other tools used in the workplace, e.g., lighting equipment, hand tools, etc.
- improper lighting conditions, need for a localized reading light

7.6 SUMMARY AND GUIDELINES

This taxonomy which is comprised of four basic issues that address the human factors concerns of information presentation (Sections 7.2 to 7.5) provides us with a framework for design of usable documentation. This framework is generic and can be extended to a set of guidelines for design of paper-based documentation for aircraft inspection tasks as discussed below. The guidelines in **Table 7.1** attempt to summarize the issues brought out in the previous sections in the form of assertive and usable statements.

Table 7.1 Guidelines for Design of Paper-based Documentation for Aircraft Inspection

1.	**INFORMATION READABILITY**
a.	Typographic Layout
1.	Resort to use of primary typographic spatial cues like vertical spacing, lateral positioning, paragraphing and heading positioning as much as possible.
2.	If space usage is premium, then resort to use of secondary cueings; e.g., boldfacing, italics, underlining, color coding and capital cueing in a decreasing order of preference.
3.	Use full justification of the textual material.
4.	Use a consistent typographic layout throughout the document.
b.	Sentence, Word and Letter
5.	Use of sentence conventions: • Boundary conventions • initial capitalization • final punctuation marks • extra space • question mark at end of question • exclamation mark • Direct speech conventions • quotation marks • paragraphing for change of speaker
6.	Use of word conventions: • Do not use all capitals format, use both upper and lower case. • Hyphen indicates word division at end of line. • Space before and after word. • Initial capitalization for proper nouns.
7.	Use of letter conventions: • Use a typeface like Helvetica that has no redundant features. • Avoid using a generic dot-matrix typeface.

Table 7.1 Guidelines for Design of Paper-based Documentation for Aircraft Inspection cont.

c.	Printing Quality Standards
8.	Develop and implement standards for changing printer ribbons, toner boxes, etc., to ensure a consistent print quality at all times.
2.	INFORMATION CONTENT
a.	Appropriate Content
9.	Information provided should be supportive of the inspector's personal goal to "read quickly and also understand the information", to ensure its usage and eliminate personal biases.
10.	It should have certain consistent and common elements to foster generalizations across contexts.
11.	It should be accurate.
12.	It should be complete, i.e., it should include information regarding what is to be done? Where? How? In what order/sequence? Which specific terms to pay attention to? References to additional sources of information?
13.	It should be up-to-date with revisions and updates.
14.	It should be easy to use and comprehend and hence should be clear and unambiguous.
15.	It should be specific and contextual, i.e., pertaining to the particular aircraft being inspected.
16.	It should be written in a consistent and standardized syntax.
17.	It should be flexible for both expert as well as novice inspectors.
18.	Eliminate use of all illogical and self contradictory statements.
19.	Use only certain approved acronyms and proper nouns and provide a glossary if called for.
20.	Try to achieve a balance between brevity, elaboration and redundancy of information.
b.	Graphic Information
21.	All spatial information is to be presented in graphical format. Avoid use of textual format for presenting spatial information.
22.	The text should assist the graphics and not vice versa.
23.	Avoid use of high fidelity graphics to eliminate clutter. Simple line drawings are superior in most cases.
24.	Use a consistent format for figure layout and numbering.
25.	Use ordinary numbers, e.g., line 1, 2, 3, etc., when referring to figures and avoid use of complicated reference numbers, e.g., T07-4032-001.
26.	Use consistent view-direction information, i.e., use either the UP-AFT icon or the UP-FWD icon, not anything else.
27.	The figure views should be as the inspector sees it, from a fixed distance/scale, e.g., 5 feet viewing distance. Avoid use of perspective part drawings as figures.
28.	All figures and attachments should have back-references to the workcard page/task which originally referred to the figure.
29.	Use standard and correct technical drawing terminology, e.g., avoid use of terms "section" and "view" interchangeably.
30.	Use typographic differentiation between figure titles, part names, crack locations, notes, etc. This differentiation should highlight the importance that one needs to give to each of these, e.g., figure number, crack location, notes, part names, etc., in decreasing order of importance calls for boldface cueing for figure numbers.
31.	Use standard drawing layout conventions, e.g., location of sectional views with reference to main views.
32.	Provide different graphics for spatially mirror imaged tasks, to reduce the cognitive costs of image inversion, e.g., avoid use of same graphics for both left and right wing inspection tasks.
33.	Differentiate close-up views from distant views by giving appropriate scaling information.

Table 7.1 Guidelines for Design of Paper-based Documentation for Aircraft Inspection cont.

3.	**INFORMATION ORGANIZATION**
a.	Classification of Information
34.	Distinguish between directive information, reference information, warnings, cautions, notes, procedures and methods.
35.	There should be a code for identifying the importance of a particular category of information over others, e.g., warnings, cautions, notes, procedures, methods, directive information, references in decreasing order of importance.
36.	Directive information should be broken into the command verb (e.g., check), the objects (e.g., valves, hydraulic lines) and the action qualifiers (e.g., for wear, frays).
37.	Each chunk of directive information should not include more than two or three related actions per step to eliminate action slips (e.g., "remove 10 bolts, remove cover" is acceptable but "check brake valves, brakes, tires and cables" is not acceptable as one chunk and calls for further break down).
38.	There should be a clear differentiation between general and specific directive information for tasks, e.g., general tasks usually call for a less detailed inspection over a large but less critical area.
b.	Information Layering
39.	Provide multiple levels of information to cater to the needs of both expert as well as novice inspectors, providing more elaborate information for novices and more concise information for experts performing the same task.
40.	Develop a standard framework for distinguishing between and writing multiple layered information. Such a framework should eliminate dependency on fancy phraseology for communication and provide a structure to write into.
c.	Other Organizational Issues
41.	The task information should be ordered/sequenced in the natural order in which the tasks would be carried out by most inspectors, e.g., according to the spatial location of the tasks as internal tasks and external tasks.
42.	The page should act as a naturally occurring information module, i.e., it should contain a fixed number of tasks and avoid carryover of tasks across pages. Each task that begins on a page should preferably end on that page too.
4.	**PHYSICAL HANDLING AND ENVIRONMENTAL FACTORS**
43.	The size of the workcard pages should be handy, e.g., avoid using large technical drawings.
44.	The entire workcard should not be excessively heavy. It should be such that it can be held continuously for an extended period by an average inspector.
45.	If the use demands exposure to environmental agents like wind, rain, snow or even harsh and oily floor conditions, adequate precautions should be taken to avoid excessive degradation.
46.	It should be compatible with the other tools that an inspector uses in the workplace, e.g., hand tools, boroscope, lighting equipment, etc.
47.	Provide a localized reading light in poor lighting conditions.
48.	Provide a specialized workcard holder to enable wiring in most positions.
49.	Provide standard writing tools (pens, pencils, etc.) that enable writing in all positions, even against vertical surfaces.

7.7 CASE STUDIES IN WORKCARD DESIGN

Aircraft inspection checks are scheduled at periodic intervals, ranging from routine flight line checks and overnight checks, through A-, B- and C-checks, to the heaviest, the D-check. Among these, two extreme representative conditions were considered as demonstration case studies. The A-check is a more frequent but less detailed inspection, while the C-check is a less frequent but more detailed inspection. The taxonomy for document design was used to develop workcards for both these inspection tasks.

7.8. A-CHECK CASE STUDY

7.8.1 Task Description

The maintenance supervisor assigns the A-check workcard to the technician. Normally two technicians are assigned to an aircraft, and inspection is carried out in the open, often under poor and varying environmental conditions. Normally, the maintenance technician completes a number of the inspection and testing tasks before beginning work on reported discrepancies. The technician has to perform and sign off each of the 201 items mentioned in the workcard, in the scheduled time. A sample page from the current workcard is shown in Figure 7.1.

```
-----------------------------------------------------------------------------------------------
03FEB92 07,51                WORK   CONTROL     CARD
REV DATE: 09-04-90           MFG P/N: JO7-3200-21-02-001          M&E P/N: 07-521-1-0017
                             PROC, A/C, 3200, B727, ATA32,                        LANDING GEAR
*----------------------------------------------------------------------------------------------
* POS*                       JOB DESCRIPTION                         * MECH*      * INSP*
*----------------------------------------------------------------------------------------------
* 01 * TITLE: LANDING GEAR
*----------------------------------------------------------------------------------------------
*        ATTACHMENTS : T07-3240-1-0001, T07-32-1-0005                   N/A           N/A
*----------------------------------------------------------------------------------------------
*        1. CHECK SYSTEMS CONDITION
*        A. LANDING GEAR AND WHEEL WELLS
*          (1) CHECK FOR GENERAL CONDITION, DAMAGE AND
*          EVIDENCE OF FLUID LEAKS                                                    N/A
*----------------------------------------------------------------------------------------------
*          (2) CHECK FOR BROKEN ALDEL CLAMPS ON HYDRAULIC LINES IN THE WHEEL WELLS AND ON
*          THE LANDING GEAR ASSEMBLIES AND CHECK FOR CHAFING HYDRAULIC
*          LINES.
*          NOTE : DISCREPANCIES IN THE ABOVE ITEM SHOULD BE CORRECTED PRIOR TO DISPATCH
*                        OF THE AIRCRAFT.                                             N/A
*----------------------------------------------------------------------------------------------
*          (3) CHECK BRAKE DEBOOST VALVES, BRAKES AND TIRES FOR BEING WITHIN LIMITS.
*          NOTE : SEE ATTACHMENT FOR PEX EFFECTED ACFT DEBOOST VALVE CHECK.            N/A
*----------------------------------------------------------------------------------------------
```

Figure 7.1 Sample Page from the Current Workcard

The maintenance technicians who perform the A-checks range in age from 23 to 55 years, with experience on A-checks varying between 1 year to 25 years. All 201 signoffs within the A-check can be classified into 18 subtasks, which fall into two general categories of tasks: "inspection tasks" and "testing tasks." The inspection tasks are those of visual inspection, to ascertain conformance to predetermined standards. Testing, on the other hand, involves determination of the proper functioning. Both inspection and testing can be further classified into "internal" and "external" tasks, depending on whether the task is to be performed on the interior or exterior of the aircraft.

7.8.2 Methods

Field visits were conducted at various A-check inspection sites. These visits included direct observations of the task, observational interviews, and personal interviewing of inspectors (inexperienced as well as experienced), technicians, and supervisors. Inspector perception regarding workcard usability was obtained from various A-check inspection sites within the airline.

7.8.3 Results

The taxonomy for documentation design was used to identify the issues relating to the current workcard design for the A-check as presented in **Table 7.2**. This case study demonstrates how such a taxonomy can be used to analyze existing documentation and points out the key issues that need improvement. What emerges from the inspector responses about workcard usage is a moderate level of satisfaction with the current workcard, but a number of users who need different information. There was a substantial agreement that the current ordering of information was incorrect and that the sign-off procedure was not performed after every step. **Table 7.3** summarizes the conclusions from inspector responses.

This study indicated that the technicians had strong views and were willing to report them when given a formal opportunity. An analysis of the task sequence preferences obtained from the inspector responses was undertaken. Based on these responses, an optimal task sequence was developed, which again is in agreement with the four basic task divisions of the A-check (inspection/test, internal/external).

7.8.4 Workcard for A-Check: Proposed Design

Based on the issues identified in **Table 7.2** and the taxonomy, a design for the workcard for A-checks has been proposed. This design comprises two parts: the design of the information/ paperwork, and the design of a workcard holder. The proposed workcard for the A-check has a two level hierarchical layering of information, as discussed. The top level is in the form of a checklist (**Figure 7.2a**), with brief task descriptions for each of the 201 signoffs, a place for the signoff itself and comments. This is the part that forms the work completion document. At the lower level is the detailed information in the form of a bound copy (**Figure 7.2b**), which remains the same until a new revision or update comes up. The directive information is broken into the command verb, the objects, and the action qualifier as illustrated. Note that identification information, rarely used by the inspector, is located on the far right.

A design was proposed for the workcard holder using the issues of **Table 7.2** under the heading of "Physical Handling/Environmental Factors." The top layer holds the checklist portion (19 pages) which can be clipped on every time before going out for an inspection, and the inner compartment holds the detailed information sheets, which remain in there until revised. The top layer opens on a hinge which houses a small reading light to allow reading in poor lighting conditions. The holder also has paper retainer clips which aid usage in windy conditions. The prototype is shown in **Figure 7.3**.

Table 7.2 A-Check Workcard: Issues Identified Within the Taxonomy

1. INFORMATION READABILITY	
A. Typographic Layout	• no consistent typographic layout • layout discontinuous, breaks within pages • no usage of secondary typographic cueing, e.g., boldface, etc. • *no use of full justification of typographic material*
B. Sentence, word, and letter	• non-conformability with printing conventions • use of all capitals format, resulting in a low reading speed • use of a 5x7 dot matrix typeface, hence no choice of any standard typeface
2. INFORMATION CONTENT	
A. Appropriate content	• some inaccuracy in the information • incomplete information for certain tasks • language difficult to use and comprehend • syntax not standardized • directive information ambiguous • generalization across aircraft types is a cause of confusion • not flexible for use by both novice and expert inspectors • use of difficult acronyms • *logical errors and contradictory statements* • redundancy and repetition • not consistent with user training • does not foster generalizations across tasks, as every task is described differently
B. Graphic Information	• system unsupportive of graphics • spatial information conveyed through text, *results in the use of* complex and lengthy sentences which are difficult to comprehend
3. INFORMATION ORGANIZATION	
A. Information Classification	• no categorization or classification of tasks • notes, cautions, methods, directions, etc. not in any prioritized order • *no demarcation between directive information, references, notes,* methods, etc. • directive information is not broken up into command verb, objects, and action qualifiers • directive information includes more than two or three related actions per step • general as well as specific information chunked together • *external as well as internal tasks not properly demarcated, mixed*
B. Information Layering	• no layering of information • not conducive to expert as well as novice usage • difficulty in writing such unstructured information
C. Other organizational issues	• no use of naturally occurring page modules for fitting in information • *improper sequencing of tasks*
4. PHYSICAL HANDLING & ENVIR.	
	• physical handling difficult due to unwieldy size • excessively heavy, cannot be held continuously • usage in extreme environments difficult • not compatible with the other tools used along with, during the task • inadequate lighting conditions • *no holder or place for holding the workcard while usage* • all these factors force them to carry out the external inspection without the workcard, relying only on memory

Table 7.3 A-check Workcard Usage: Interpretations of Inspector Responses

Q.No.	Interpretation
1.	- 66% of the inspectors find the present workcard a useful source of information.
2.	- 60% of the inspectors refer to the workcard while doing the A-check, either usually or always.
3.	- Most people feel that the readability of the current workcard is either fair or good.
4.	- There is no unanimous opinion amongst the inspectors, as to whether they prefer a concise or detailed workcard.
5.	- Almost half the inspectors prefer a smaller size workcard, while the other half feel that the current size is about right.
6.	- Most inspectors feel that the information provided on the workcard is only sometimes sufficient to carry out the A-check task.
7.	- Almost 50% of the inspectors feel that the current workcard is moderately easy to understand.
8.	- Most inspectors face problems either sometimes or always in physically using the workcard while working.
9.	- 65% of the inspectors do not carry out the A-check activities in the same way as listed out in the workcard.
10.	- 80% of the inspectors say that they have felt the need for more information that was not provided on the workcard, either sometimes or always.
11.	- There is no unanimous opinion amongst the inspectors, as to whether they use the A-check accountability list provided at the beginning of the current workcard.
12.	- 50% of the inspectors sign off the completed tasks on the workcard at the end of the entire inspection.

S.#	Description	P#	Signoff	Comment
A	**Landing gear and wheel wells**	—	—	—
1.	General condition, damage, fluid leaks	5		
2.	Wheel wells, landing gear assemblies, hydraulic lines	5		
3.	Brake deboost valve limits	5		
4.	Brake limits	6		
5.	Tires, for wear, damage, fluid leaks	6		
6.	Nose wheel cap attached bolts	6		
7.	Nuts for bottoming of last tread	7		
8.	Shock struts for, normal extension, general condition, leaks	7		
B	**Tire pressure check**	—	—	—
1.	Tire pressure	7		
C	**Main gear doors**	—	—	—
1.	Doors, operating cable, cranks & arms, general condition	8		
D	**Main gear wheel well down lock viewing windows**	—	—	—
1.	Indicator stripes for clarity and legibility	8		

Figure 7.2a Proposed Design for A-Check Workcard: Checklist

2A&B LANDING GEAR (Nose & Main)	*Work Control Card : Landing Gear SC1CIAAC*
Check systems condition	
A. Landing gear and wheel wells	
(1) *Check* : general condition damage and evidence of fluid leaks	
(2) *Check* : - wheel wells - landing gear assemblies - chafing of hydraulic lines for broken aldel clamps on hydraulic lines on the *Note* : Discrepancies on the above item must be corrected prior to the dispatch of the aircraft.	*03Feb92 07:51* *Rev date 090490* *No. 02189 MFG P/N:J07 3200-21-020-001* *Proc, A/C 3200, B727, ATA32, M&E P/N :*
(3) *Check* : - Brake deboost valves for being within limits *Note* : See attachment for PEX effected ACFT deboost valve check	*07-0521-1-0017* **Page 5**

Figure 7.2b Proposed Design for A-Check Workcard: Detailed Information

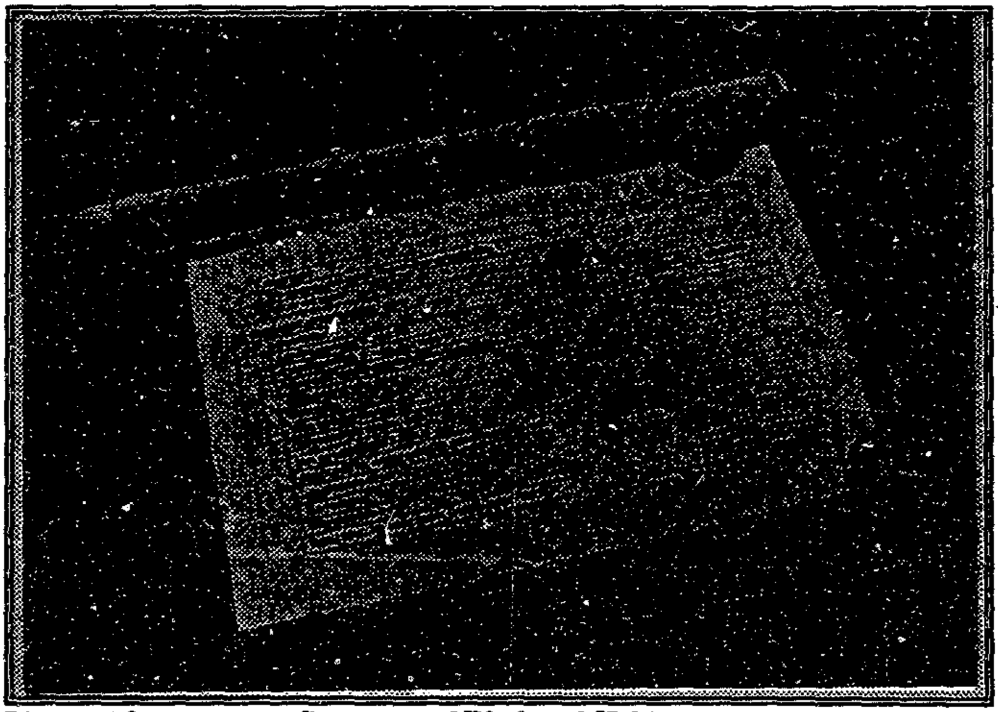


Figure 7.3 Prototype of Workcard Holder

7.9 C-CHECK CASE STUDY

7.9.1 Task Description

A typical C-check workcard consists of seven item groups. Most items refer to attachments which must be procured prior to inspection, these being figures to be referred to during the tasks. Unlike the A-check which is performed by just one maintenance technician (overnight), the C-check is a task which extends across shifts, involving a number of inspectors, working simultaneously on various tasks. The C-check is usually carried out in a hangar, unlike the A-check: the task lighting does, however, vary depending on the task. The time scale is typically not as short as that involved in an A-check.

The inspectors performing C-checks range in the age between 25 to 60 years, with a 1 to 35 years of experience on C-checks. Also, since only a portion of the C-check task is given to an inspector at a time, inspectors are expected to perform different portions of the C-check at various times depending on what has been scheduled for them. This demands a total expertise on all tasks, but with some time elapsing between repetitions of a specific task. Considering the number of tasks involved in a typical C-check, it was decided to analyze and demonstrate particular portions representative of most inspection tasks. After discussions with inspectors and supervisors, two tasks were selected for this case study: Left Wing Inspection and the Right Wing Inspection. The current C-check workcard is very similar in layout to the A-check workcard shown in **Figure 7.1**.

7.9.2 Methods Used for the Study

A field visit was conducted at a C-check inspection site. Visits included direct observations of the left and right wing inspection task, observational interviews, and personal interviewing of both experienced as well as inexperienced inspectors, technicians, and supervisors. Inspector perception about the current C-check workcard was obtained from all C-check inspection sites within the airline.

7.9.3 Results

The taxonomy for documentation design was used to identify some of the issues relating to the current C-check workcard as presented in **Table 7.4**. The information readability and organization issues are very similar to those for the A-check. The information content issue, however, is different as far as the requirements of graphic information are concerned. **Table 7.5** summarizes the conclusions. As with the A-check, most C-check inspectors seem to be troubled with the issue of information content, pointing at a scarcity in the information and need for more and better quality graphic information. As far as the issue of information organization was concerned, most users felt that there was no clear differentiation between general and specific information.

Table 7.4 C-Check Workcard: Issues Identified within the Taxonomy

1. INFORMATION READABILITY	
A. Typographic Layout	• no consistent typographic layout • layout discontinuous, breaks within pages • no usage of secondary typographic cueing, e.g. boldface, etc. in both text and graphics • no use of full justification of typographic material
B. Sentence, Word, and Letter	• non-conformability with some of the printing conventions. • *use of all capitals format, resulting in a low reading speed* • no room for selecting an appropriate typeface • use of a 5x7 dot matrix typeface
2. INFORMATION CONTENT	
A. Appropriate content	• some level of inaccuracy in the information • incomplete information for certain tasks and lack of information on spatial location • language difficult to use and comprehend • syntax not standardized • directive information ambiguous • generalization across aircraft types is a cause of confusion • use of difficult acronyms • logical errors and contradictory statements • redundancy and repetition • does not foster generalizations across tasks, as every task is described differently
B. Graphic Information	• *no figure numbering, even though the workcard refers to specific figure numbers,* • fosters guessing and speculation for interpretation • *no consistent layout of figures, use of a mixed layout and no demarcation* • no consistency in view directional information (e.g., use of both UP-AFT & UP-FWD) • *non-contextual figure views, or views as the inspector sees it, just perspective part drawings* • *no information to aid spatial location of parts* • no back references to the workcard page/task which refers to the figure • *improper usage of technical drawing terms (e.g., "section" and "view" used inter-changeably)* • no *typographic differentiation between: figure titles, part names, crack locations, notes, etc.* • no use of standard drawing conventions (e.g., location of sectional views) • same graphics for both left and right wing tasks, mentally inverting the figures causes high cognitive workload • some figures use high fidelity graphics, causes confusion and clutter • no consistency of scaling in graphics, close up views not differentiated from distant views
3. INFORMATION ORGANIZATION	
A. Information Classification	• no categorization or classification of tasks • notes, cautions, methods, directions, etc. not in any prioritized order • no demarcation between directive information, references, notes, methods, etc. • directive information is not broken up into command verb, objects, and action qualifiers • directive information includes more than two or three related actions per step • both general as well as specific information chunked together • general and specific tasks not properly demarcated
B. Information Layering	• *no layering of information* • *not conducive to expert as well as novice usage* • difficulty in writing such unstructured information
C. Other organizational issues	• no use of naturally occurring page modules for fitting in information • improper sequencing of tasks • no consistency in the number of signoffs across the task
4. PHYSICAL HANDLING & ENVIR	• size of the attachments different from the workcard size, causes inconvenience in usage • inadequate lighting conditions in certain work areas • no holder or place for holding the workcard while usage

Table 7.5 C-Check Workcard: Interpretations of Inspector Responses

Q.No	Interpretation
1.	- 48% of the inspectors find the present workcard a useful source of information.
2.	- 93% of the inspectors refer to the workcard while doing the C-check, either usually or always.
3.	- 42% of the people feel that the readability of the current workcard is either fair or good.
4.	- There is no unanimous opinion amongst the inspectors, as to whether they prefer a concise or detailed workcard, 55% of the people prefer more detailed information.
5.	- 75% of the inspectors feel that the physical size of the current workcard is about right.
6.	- 42% inspectors feel that the information provided on the workcard is never sufficient to carry out the C-check task.
7.	- 40% of the inspectors feel that the current workcard is moderately easy to understand.
8.	- 60% inspectors faced problems sometimes in physically using the workcard while the remaining were equally divided in their views.
9.	- 33% of the inspectors do not carry out the C-check activities in the same way as listed out in the workcard.
10.	- 98% of the inspectors say that they have felt the need for more information that was not provided on the workcard, either sometimes or always.
11.	- There is no unanimous opinion amongst the inspectors, as to whether they use the C-check accountability list provided at the beginning of the current workcard.
12.	- 42% of the inspectors felt that there were too many signoffs on the current workcard, but 30% felt that there were too few.
13.	- 88% of the inspectors had missed noticing workcard revisions either sometimes or at least on some rare occasions.
14.	- Only 32% of the inspectors signoff each completed task individually after it is done.

7.9.4 Workcard for C-check: Proposed Design

The issues highlighted by the inspector responses and those identified in **Table 7.4** were used to produce with a demonstration design for the left and right wing inspection tasks of the C-check. Unlike that of the A-checks, the workcard for the C-check calls for a single layered design with an additional set of figures and graphics in the form of attachments. **Figures 7.4a** and **7.4b** show the proposed design for the left wing inspection task. Within the main workcard itself (**Figure 7.4a**), information is organized in a format encompassing three tasks per page, each task description consisting of two parts, graphic and text. The graphic is an iconic representation of the wing with the location of the body stations at which the task is to be carried out. The text part is the directive information, broken up into the command verb, the objects and the action qualifier.

Figure 7.4b illustrates the basic layout for redesign of the graphics attachments. It consists of an iconic representation of the wing in the top right corner with the location of the task to be carried out. Also shown are the body stations and a footnote containing the directive information from the workcard. The directive information refers to that particular figure and the page number on which it appears. Even though there is an attachment number, each figure has a single digit figure number for ease of referencing. The actual figure is a low complexity graphic representation of the portion being inspected in a view as the inspector would see it, rather than the isometric part views as presented in the existing attachments. Use of typographic differentiation between figure titles, part names, crack locations, notes, etc., is also included. A consistent scale has been used for all primary graphics, using a viewing distance of around 4 to 5 feet from the inspector. Also, different figures are provided for both the left and the right wing inspection tasks, unlike the present attachments which use the same figures for both. This allows proper body station numbering on the figure and also helps in reducing the cognitive workload of having to mentally invert the figures before interpreting them.

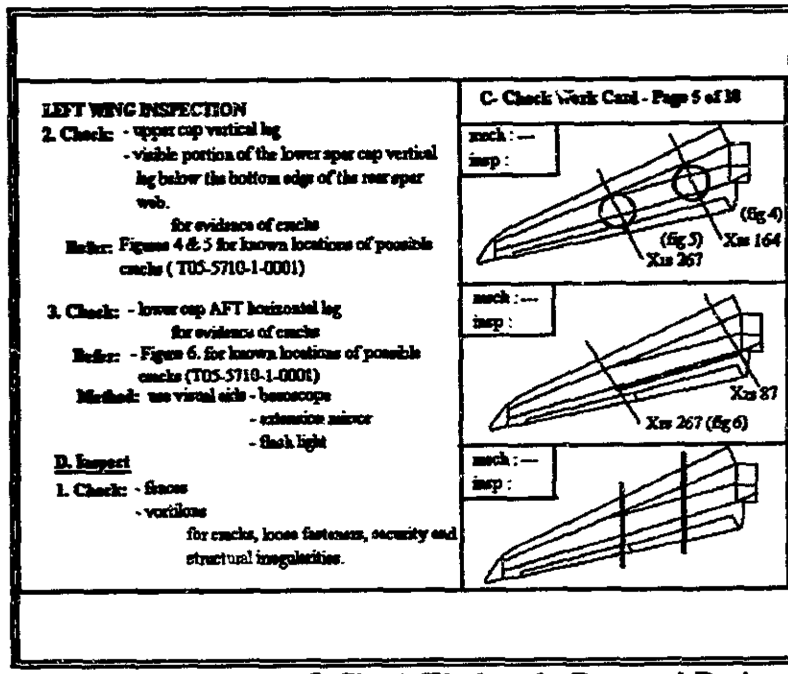


Figure 7.4a C-Check Workcard: Proposed Design

7.10 FIELD EVALUATION OF PROPOSED DESIGNS FOR C-CHECK WORKCARD

To test the validity of the issues identified and the proposed taxonomy, an empirical evaluation was carried out with the airline partner to evaluate the proposed design of the C-check workcard in relation to the current computer-generated C-check workcard.

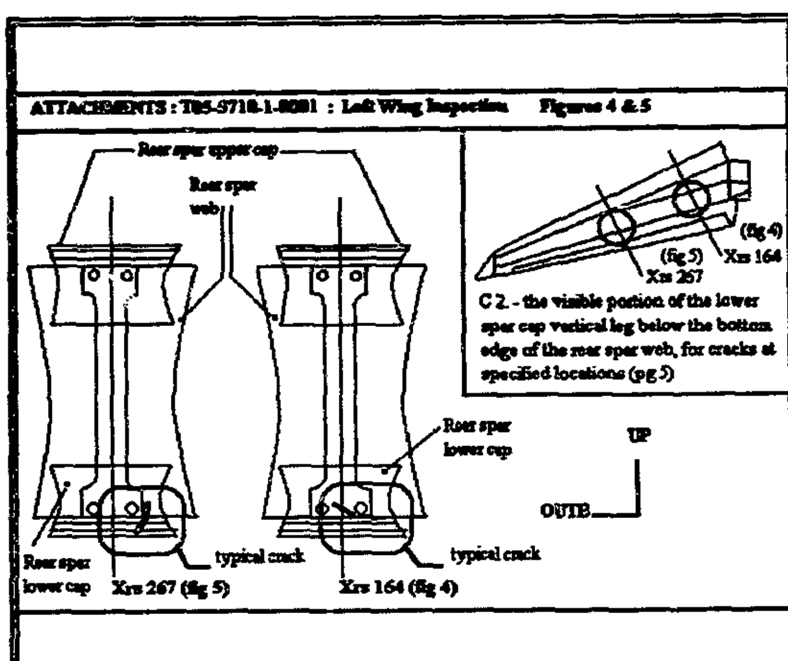


Figure 7.4b C-Check Workcard: Proposed Design - Graphics Attachments

7.10.1 Experimental Design

The eight inspectors used for this experiment were asked to perform two tasks from the wing inspection portion of the C-check. One of the tasks involved using the current computer-generated workcards and the other involved using the proposed design of the C-check workcard. In all, four combinations of the workcard were used, one combination for the left and the other for the right wing inspection. Table 7.6 lists the order in which the experiment was carried out. After each task the inspector was asked to rate the workcard used for that task on 14 issues, each relating to those brought out in the taxonomy. The same set of issues was used for both the tasks.

Table 7.6 Experimental Design

Inspector #	Task 1	Task 2
I1	Current W/C - Left Wing	Proposed W/C - Right Wing
I2	Current W/C - Left Wing	Proposed W/C - Right Wing
I3	Current W/C - Right Wing	Proposed W/C - Left Wing
I4	Current W/C - Right Wing	Proposed W/C - Left Wing
I5	Proposed W/C - Right Wing	Current W/C - Left Wing
I6	Proposed W/C - Right Wing	Current W/C - Left Wing
I7	Proposed W/C - Left Wing	Current W/C - Right Wing
I8	Proposed W/C - Left Wing	Current W/C - Right Wing

7.10.2 Results

The inspector responses to the rating task were analyzed using Wilcoxon Matched-Pairs Signed-Ranks Test for two Related-Small Samples, since each subject acted as his/her own control. **Table 7.7** summarizes the results along with the average rated scores for the current and the proposed workcards. The results point in favor of the proposed design on each of the 14 points. One of the issues addressed the more general matter of whether separate attachments for the left and the right wing inspection tasks would be useful. Most inspectors thought that such a format would be extremely useful.

7.11 RECOMMENDATIONS AND CONCLUSIONS

Both the A-check and the C-check case studies showed that substantial redesign of the existing workcards is required. This is true whether they are to be replaced by new hard-copy workcards, or by a portable computer system. The taxonomy of documentation design presented here provides the framework required for investigating documentation in field conditions, using direct observation and user feedback in a structured manner to develop improved designs. The results of the in-field empirical evaluation of the proposed C-check workcards proves the validity of the methodology as well as the proposed taxonomy. At present, sample workcards are being prepared for other tasks by the airline partner to ensure that the benefits are applied as widely as possible.

Table 7.7 Summary of the Results

Q.#	Issue addressed	8-Point Rating Scale End-points		Current W/C mean rating (SD)	Proposed W/C mean rating (SD)	Wilcoxon Matched Pairs Signed-Ranks Test
		0	8			
1	Readability of the Workcard	Terrible	Excellent	4.0 (1.2)	6.6 (0.7)	$p < 0.005$
2	Continuity of information flow	Terrible	Excellent	3.4 (1.1)	6.0 (1.0)	$p < 0.01$
3	Ease of information location	Very difficult	Very easy	2.5 (1.0)	4.7 (1.0)	$p < 0.005$
4	Chance of missing information	Always	Never	4.6 (1.1)	5.8 (0.8)	$p < 0.025$
5	Ease of understanding	Very difficult	Very easy	2.7 (0.7)	5.9 (1.6)	$p < 0.005$
6	Ease of location w.r.t. body-stations	Very difficult	Very easy	2.2 (1.3)	5.5 (1.4)	$p < 0.005$
7	Ease of relating figure numbers	Very difficult	Very easy	3.4 (1.9)	6.1 (1.1)	$p < 0.01$
8	Amount of information provided	Too little	Too much	4.2 (1.6)	3.6 (0.7)	n. s.
9	Ease of readability of attachments	Terrible	Poor	4.1 (1.4)	6.4 (1.2)	$p < 0.025$
10	Relating graphics with wing structure	Very difficult	Very easy	2.8 (0.9)	6.1 (1.7)	$p < 0.025$
11	Consistency of presentation	Terrible	Poor	4.0 (1.4)	6.8 (0.8)	$p < 0.01$
12	Compatibility of attachments with Workcard	Terrible	Poor	3.7 (1.6)	6.1 (0.7)	$p < 0.025$
13	Amount of graphics provided on the attachments	Too little	Too much	2.5 (1.0)	4.0 (0.5)	$p < 0.025$
14	Overall ease of usability of the Workcard	Terrible	Excellent	3.3 (0.9)	5.6 (0.4)	$p < 0.005$

7.12 REFERENCES

Campbell, A. J., Marchetti, F. M. and Mewhort, D. J. K. (1981). Reading speed and text production: A note on right-justification techniques. *Ergonomics, 24,* 633-640.

Carrol, J. B., Davies, P. and Richman, B. (1971). *The american heritage word frequency book.* Boston: Houghton-Mifflin.

Drury, C. G. (1989). Information Environment in aircraft inspection. *Proceedings of the Second International Conference on Human Factors in Aging Aircraft,* 92-109.

Haber, R. N. and Haber, L. R. (1981). Visual components of the reading process. *Visible Language XV, 2,* 147-182.

Inaba, K. (1991). *Converting technical publications into maintenance performance aids.* XYZYX Information Corporation.

Jewette, D. L. (1981). Multi-level writing in theory and practice. *Visible Language, XV, 1,* 32-40.

Prabhu, P., Gramopadhye, A. and Drury, C. G. (1991). Human factors in workcard design: an initial presentation of the issues. Unpublished technical report prepared for Galaxy Scientific Corporation/FAA/OAM.

Shepherd, W. T., Johnson, W. B., Drury, C. G., Taylor, J. C. and Berninger, D. (1991). *Human factors in aviation maintenance, phase I: Progress report.* (Report No. DOT/FAA/AM-92/16). Springfield, VA: National Technical Information Service.

Sutton, I. S. (1991). Writing operating procedures. *International Conference on Hazard Identification and Risk Analysis, Human Factors and Human Reliability in Process Safety,* 293-308.

Swander, A. J. and Vail, R. E. (1991). Building procedures for repeatability and consistency in job performance. *International Conference on Hazard Identification and Risk Analysis, Human Factors and Human Reliability in Process Safety,* 339-352.

Wright, P. (1991). Quality: An excellent idea? *Conference on Quality in Documentation.* Centre for Professional Writing: U.K., 11-24.

Chapter Eight
Training for Visual Inspection of Aircraft Structures

8.0 INTRODUCTION: THE STATE OF VISUAL INSPECTION TRAINING

Although much has been written about training for aircraft maintenance in the past several years (e.g., Shepherd and Parker, 1990), very little applies directly to the acquisition and enhancement of visual inspection skills. Typically papers concern either the overall structure of training program (Skinner, 1990; Desormiere, 1990) or the technology of training delivery systems (Payne, 1990; Kurland and Huggins, 1990; Goldsby, 1991; Rice, 1990). Visual inspection has skill, rule and knowledge-based components and, as such, is less amenable to rule-based diagnostic procedures. Such procedures have received widespread study in avionics (Johnson, 1990; Johnson, et al., 1992) and nuclear power (Kello, 1990) and can represent the technological leading-edge of training delivery. However, for visual inspection, visits by the State University of New York (SUNY) team to many airlines have revealed a relatively uniform approach. Training content is seen as either knowledge or skills, with knowledge imported in the classroom and skills through on-the-job training (OJT). Despite the overall effectiveness of inspection, our site visits have revealed a strong desire to find enhanced ways to implement visual inspection training.

Examples of high quality inspection training can be seen in many airlines, but in all there is the same classroom/OJT split of delivery methods.

- On-the-job training (OJT) is the preferred way of imparting training to new inspectors. Most of the classroom training targets pre-flight checks, non-destructive test (NDT) or orientation. One airline (Lutzinger, 1989) has such a program where 5% (approximately 15-17) of their inspectors are trained daily. They also have a program called C-3 under which, when a supervisor discovers an aircraft discrepancy missed during an earlier inspection, he codes this item C-3. These C-3 items are used to point out to inspectors what kinds of discrepancies are being missed during aircraft checks.

- An aircraft manufacturer has developed a "task analytic training system" model to address training needs for NDT (Walter, 1990). However, elements of this model are generalizable to visual inspection. This method consists of performing a job task analysis to identify training needs and job instruction training to impart knowledge. This methodology also emphasizes a team approach to developing the training modules. A design team, an approval team, and a team facilitator comprise the personnel creating a module. An iterative procedure involving the work force is utilized. An annual audit assesses the status of each training module.

- Another airline combines orientation training with on-the-job training (OJT). There is a new inspector orientation training called Q.C. transition orientation which lasts for 16 hours over two days. The inspector is then put on a 40 hour OJT at the end of which he/she is Required Inspection Item (R.I.I.) certified. The airline has instituted an inspection research request program. Under this, inspectors who detect a problem with inspection procedures, workcards, documentation, etc. can submit a request to the quality control analyst for a review. Inspectors also have available information on inspection alerts that are generated by engineering. When a new aircraft arrives, Q.C. denotes one of its foremen as a training instructor. The training instructor visits the manufacturer and gets information to set up a training program for inspectors. Quality control then initiates training for the whole department to familiarize them with the new aircraft.

There are, however, drawbacks to the classroom/OJT approach. These have been identified by proposers of simulation as a delivery intervention (e.g., Johnson, 1990; Lesgold, 1990; Kello, 1990) who show that classroom/OJT is not an optimum way to train.

Shepherd and Parker (1990) recognize both the content and the delivery system as crucial to training. Content has been defined for maintenance in terms of knowledge, skills and abilities (KSA) since the Allen report in 1972. Currently, the KSA's for maintenance are being redefined for Part 147 schools, but again there is little on visual inspection. A recent compilation of training research and practice (Patrick, 1992) can be used to define the current status of training and the issues involved. These lead to experimental evaluations of particular training interventions for visual inspection, which in turn provide the basis for enhanced training programs.

8.1 CURRENT METHODS IN TRAINING: AN OVERVIEW OF RESEARCH

Training design deals with the issue of translating training content into a training program. Patrick (1992) identifies training content, training methods and trainee characteristics as the three main components of a well-designed training program. With the advent of computer-based training (CBT) and multi-media approaches, we should add training delivery systems as another component.

8.2 TRAINING CONTENT

Training content pertains to identifying the knowledge and skills needed to perform the set of tasks that define a job. For example, Wirstaad (1988) identifies knowledge categories (e.g., layout knowledge), knowledge objects (documents, etc.) and depth of knowledge as a way of identifying job training requirements.

A systematic analysis of the task is necessary to identify the training content. Patrick (1992) classifies analysis techniques into (a) task-oriented analysis and (b) psychological techniques. Task-oriented techniques use task-oriented data to derive the needs, objective and content of the training program. Examples are Task Analytic Training (e.g., Walter, 1990), hierarchical task analysis, HTA (Drury, et al., 1990), critical incidents technique (Flanagan, 1954) and task inventory (e.g., USAF Task Taxonomy, Christal, 1974). Psychological approaches typically use taxonomies that categorize aspects of the task in terms of human motor/perceptual/cognitive processes. This can help the analyst understand the psychological elements that need to be addressed (e.g., decision making skills, reasoning, etc.) by specific training methods.

8.3 TRAINING METHODS

Training methods deal with techniques that can help transfer the training contents to the trainee in an effective manner. Some of the common/popular methods (Patrick, 1992; Drury and Gramopadhye, 1990) are discussed below.

8.3.1 Pre-Training

Pre-training provides the trainee with information concerning objectives and scope of the training program. Pretests can be used to measure (a) level at which trainees are entering the training program, and (b) cognitive or perceptual abilities that can be later used to gauge training performance/progress. Advanced organizers or overviews, which are designed to provide the trainee with the basics needed to start the training program, have been found to be useful. The elaboration theory of instruction (Reigeluth and Stein, 1983) proposes that training should be imparted in a top-down manner where a general level is taught first before proceeding to specifics. Overviews can fulfill this objective by giving the trainee an introduction to the training program and facilitating assimilation of new training material.

8.3.2 Knowledge of Results

Knowledge of the results is probably the most common and effective method of training. Drury and Kleiner (1990) suggest that training programs start with rapid, frequent feedback which is gradually decreased until "working" level is attained. Additional feedback beyond the end of training will help to keep the inspector calibrated (Drury and Gramopadhye, 1992). Gramopadhye (1992) classifies feedback as performance and process feedback. Performance feedback for inspection typically consists of information on search times, search errors and decision errors. Process feedback, on the other hand, informs the trainee about the search process, i.e., areas not covered, inter-fixation distance, number of fixations. Research (explained in the next sections) supports the beneficial effects of process feedback on inspection performance. Another type of feedback called "cognitive feedback" has emerged from the area of social judgement theory. Cognitive feedback is the information to the trainee of some measure of the output of his or her cognitive processes. It is suggested that cognitive feedback allows the trainee to perceive the error in their judgement as well as why the judgement was in error (Hammond and Summers, 1972; Doherty and Balzer, 1988).

8.3.3 Guidance or Feedforward

Guidance or feedforward provides the trainee with information prior to action, concerning how to carry out part or all of the task. For example, an experienced inspector can tell the novice how he looks for evidences of corrosion in the cargo compartment. Guidance could be physical (for acquisition of motor skills), demonstrations, verbal advice (e.g., prompting), and cueing (telling when and what signal occurs in perceptual detection tasks). Feedforward can also be by informing the inspector what to expect in a certain area that he is going to inspect next. Feedforward should provide the trainee with clear and unambiguous information which can be translated into performance.

8.3.4 Part-Task Training

Part-task training constitutes partitioning or simplifying the whole task into parts and then teaching these parts to the trainee. Part-task methods are classified by the manner in which parts are practiced. Isolated parts training consists of learning each part separately for either a fixed number of trials or to some criterion, and then doing the whole task together. Progressive part training teaches components of the job to criterion, and then successively larger sequences of the components (Drury and Gramopadhye, 1990). Repetitive part training involves practicing one part, then parts one and two, then parts one, two and three and so on (Patrick, 1992). In general, part task training has been found to be beneficial for complex tasks. Drury (1990) reported good results in industrial inspection tasks when using progressive part training methods.

8.3.5 Whole-Task Training

Whole-task training involves the trainee on the task as a whole instead of breaking it into parts. Naylor and Briggs (1963) have postulated that for tasks of relatively high organization, as task complexity is increased, whole-task training should become relatively more efficient than part-task training methods. This is an intuitively appealing principle because as parts of the task become more interdependent the trainee might have a harder time integrating performance on the whole task if trained on part task. Whole-task training also becomes necessary when it is not possible to identify task parts having natural segmentation with relation to the task as well as the skills needed from the trainee.

8.3.6 Adaptive Training

Adaptive training tries to accommodate the characteristics of the trainees. This method comes from a recognition that individual differences exist in the skills, knowledge, abilities and aptitudes that trainees possess. Bartram (1988) has reported a study involving operators at a post office, in which adaptive training methods were used to train the operators to sort mail at an average time per item of 1.8 seconds or less and error rate of less than 1%. Adaptive training involves measurement of the trainee's performance and making changes in the program/method as a function of the trainee's performance.

8.3.7 Active Training

When the trainee has to actively discover information or cues, and make a physical response, learning is enhanced (Czaja and Drury, 1981). A good passive training scheme, where the information is merely presented to the trainee, is often inferior to an equivalent active scheme, where a response is required at each step.

8.4 TRAINING DELIVERY SYSTEMS

Training can be delivered to the trainee through a variety of media. With the increasing use of computers and the advances in multi-media technologies, the choices in delivery systems are varied and the task of selecting one is non-trivial. We can classify training delivery systems under the following broad categories.

8.4.1 On-the-Job Training (OJT)

On-the-job training is a much maligned word in the area of training and much of the literature is full of examples of its inadequacies. While this is true, there is a case to be made for structured OJT supplemented by adequate classroom instructions. This is especially true in cases where realistic simulators cannot be developed or are too expensive.

8.4.2 Traditional or Conventional Training

In this system we have the traditional method of instructors tutoring to trainees using slides, blackboard, and paper and pencil methods. Further embellishments include using video and TV media and even 3-D projections (Rice, 1990) to explain instructions or provide concept training. Training aids can include prototypes or models of objects used in the actual task. All the principles of training design can be used to varied extent in this system. In a review of training systems in the U.S. military, Orlansky and String (1979) found trainee achievement using conventional training and computer-based training about equivalent. However, with the advent of better and more powerful computer systems with higher quality graphics and computer-based training (CBT) methods that utilize more training design principles, the balance might be tilting in favor of CBT systems.

8.4.3 Computer-Based Training (CBT) Systems

Training systems of this type are also called computer-aided instruction systems (CAI). Patrick (1992) identifies four main roles for computers in training: (a) provision of training, (b) development of training, (c) management of training, and (d) research in training. CBT systems typically are used to fulfill the first and third roles, in which the computer is used to present training material to the student and schedule him/her through various training exercises, record progress, administer tests and provide progress summaries to the instructor. Boeing has produced a series of maintenance training CBT lessons on the 767 airplane (Lukins, 1990). They are also investigating the area of instructor-led CBT used in a classroom environment.

8.4.4 Intelligent Tutoring Systems (ITS)

Computer-based tutoring systems attempt to create interactive learning environments in which the learner/trainee can carry out simulated tasks. Typically, an ITS contains (1) an explicit model of the domain, (2) an expert program that solves problems in this domain, (3) a model of the student that explains what the student understands, and (4) a tutoring model that provides instructions (Clancey and Soloway, 1990).

SOPHIE-III (Brown, Burton and deKleer, 1982) is an intelligent simulation training system that supports interactive training by estimating and responding to student needs. SHERLOCK (Lesgold, Lajoie, Bunzo and Eggan, 1992; Lesgold, 1990) is a practice environment for learning troubleshooting a complex device on the F-15 manual avionics test station. It analyzes inferred student models with respect to expert models and emphasizes refinement of mental models. The Integrated Information Management System (IIMS) developed by the U.S. Air Force has imbedded training systems that include such features as multi-level representations for expert, novice and trainee, preview of little used tasks, and troubleshooting simulations (Johnson, 1990). The Environmental Control System (ECS) Tutor is an intelligent simulation that provides appropriate feedback and advice to the student based on observed interaction (Norton, 1992).

8.5 EXPERIMENTAL EVALUATION OF TRAINING INTERVENTIONS

Combining the literature on training for visual inspection with the demonstrated need for new techniques of inspector training, it is apparent that there are major issues in need of testing. Both visual search and decision-making aspects of the task require assistance if we are to achieve and maintain high levels of inspector effectiveness within a visually-complex, but often repetitive, visual inspection task. These issues can be classified as:

- Visual Search Issues

  1. Can training improve defect conspicuity? Specifically, can the visual lobe size be increased, and if so, does this increase generalize across different defects?
  2. How effective is feedback in changing search strategy? Specifically, should feedback be about the inspector's performance, or about the inspector's strategy?
  3. How effective is cueing or feedforward in changing search strategy? Specifically, do inspectors perform better with generalized information or specific recommendations?

- Decision-Making Issues

  1. Perception of multiple defect attributes: is an active training scheme better than the equivalent passive scheme?
  2. Integration of multiple defect attributes: is an active training scheme better than the equivalent passive scheme.

Note that these issues are ones suggested by the literature on industrial inspection, but untested in the airframe inspection context. Note also that issues have been chosen which do not imply the need for particular delivery systems, even though the evaluation of each issue was carried out on a computer-based simulation.

Each of the issues above defines a training intervention which was evaluated using a consistent methodology. Each test was aimed at determining whether a particular intervention had an impact on improved performance. Because these needed controlled conditions, often with many repetitions of similar faults, actual airframes and inspectors were not logically possible. For example, the hundreds of cracks and dents required for the visual lobe training would never be available to an inspector. Thus, a visual inspection simulator was developed, using a SUN workstation computer to reproduce the essential aspects of the visual inspection task.

In this task the inspector searches for multiple defect types and classifies them into different severity categories. The seven possible fault types are missing rivet, damaged rivet, pooched/dished rivet, loose rivet, rivet cracks, dents and corrosion.

The entire inspection task is a series of search areas where each search area is that portion of the task which is shown on one screen. A part of the aircraft fuselage (one search area) is presented to the subject, whose task is to locate the fault in the search area and indicate its discovery by clicking the left mouse button on the fault. The layout of the multi-window simulated inspection task is as shown in **Figure 8.1**. The function of each window is as follows:

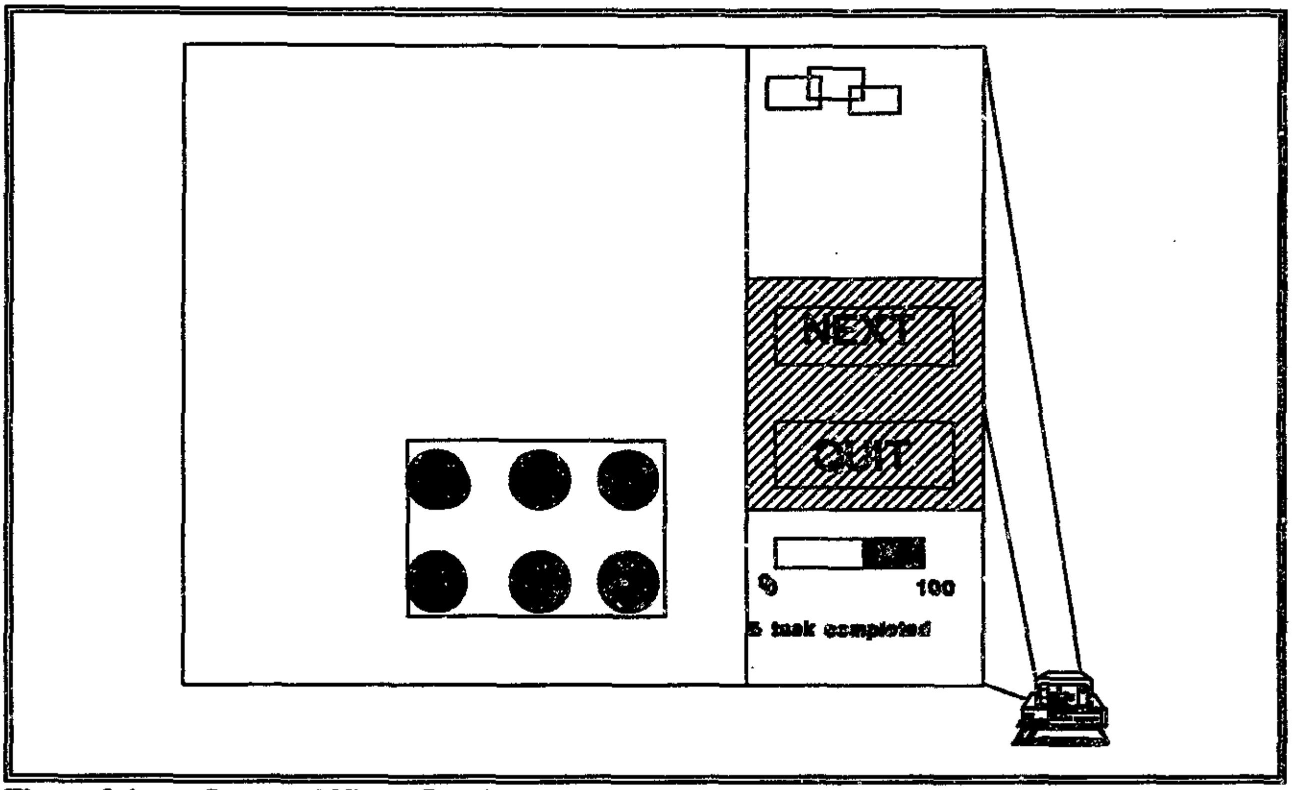


Figure 8.1 Screen of Visual Simulation

1. Inspection Window. The area currently being inspected is shown in the left (large) window. To simulate the use of local lighting, such as a flashlight beam, only a smaller window within this area is fully illuminated. Within this smaller window, or "viewer", faults can be seen and responded to by clicking them using the left mouse button. The entire area of the inspection window can be viewed by successive movements of the viewer.

2. Search Monitor Window. This is a monitoring device which helps the inspector keep track of the window movement in the inspection window. The viewer in the inspection window is represented by a tile in the search area window. As the viewer is moved, so does the tile, which has a different color from its illuminated background area. The darkest shade of the tile is the point of previous fixation so that the sequence is given by the shade of the color--lighter shades indicate earlier fixations in sequence while darker shades indicate later fixations.

3. Macro View Window. This window represents the entire task to be inspected, and provides information to the inspector about his current position with reference to the entire task.

The visual inspection simulation generates an output data file of subject performance consisting of both individual statistics for each search area, and summary statistics averaged over all search areas. Both performance and process measures are collected. The performance measures include: the number of faults located, the time to detect each fault, the stopping time, the number of hits, misses, and false alarms, and the average time spent in each search area's nine zones. The process measures collected include: the percentage of search area covered by the viewer, the number of viewer fixations used to search the area, the interfixation distance, the percentage overlap of the viewer fixations, and the pattern of viewer movements. A set-up program allows fault types to be assigned to different rivets, and information such as feedforward and feedback to be added. Obviously, not all features of the simulation were used in all experiments.

In all of the experiments, engineering student subjects were used, as these represent well the technically-fluent, but inexperienced, labor pool from which the aviation mechanics (who will eventually become inspectors) are drawn.

The five intervention issues were tested in five experiments. All five measured performance by many measures, using statistical tests to determine whether each intervention had the predicted impact on each measure. It is not the intention of this report to provide exhaustive experimental details, but rather to demonstrate the main results, interpret these results in terms of training of airframe inspectors, and outline any still-unresolved issues in terms of future evaluation needs. Full details are available elsewhere (Gramopadhye, 1992), and will eventually be published as a sequence of individual technical papers. This sequence has already begun with the visual lobe training evaluation (Latorella, et al., 1992; Drury and Gramopadhye, 1992).

8.6 VISUAL SEARCH TRAINING EVALUATIONS

In a visual search task, the inspector's eyes move across the inspected area, fixating subareas with the eye stationary and jumping rapidly between fixations. The area within which a target can be detected during a fixation is the visual lobe. The size of this visual lobe is important as it determines how thoroughly an area will be searched in a given time period, and hence directly determines the probability of defect detection. Search strategy is the sequence of fixations used by the inspector and determines the total coverage of the area. An effective and efficient strategy is one which covers the whole area with the minimum overlap between fixations and the minimal back-tracking. The first issue concerns the visual lobe size, while the second and third issues evaluate strategy.

8.6.1 Can Training Improve Defect Conspicuity?

The objectives of this experiment were to determine the relationship between visual lobe and search performance, relate changes in lobe size to search performance, and to evaluate the effectiveness of lobe training. In particular, the experiment measured whether crossover effects exist in visual lobe training. It used two types of rivet faults (cracks and loose rivets) and two types of area faults (corrosion and dents) to determine whether visual lobe training on one fault would generalize to other faults of the same or different classes.

The experiment consisted of a familiarization training followed by four visual search tasks, each consisting of 20 examples of one fault type. Following these tasks, each of the four groups of six subjects undertook different training schemes based on a visual lobe task which presented a fault for 0.3 seconds at one of six positions around a central fixation point. The four groups received the following training:

Rivet Group:	Five trials of 120 visual lobe screens containing loose rivets.
Area Group:	As Group 1 except using dents.
Neutral Group:	As Group 1 except using a fault (cross) which was irrelevant to the search task.
Control Group:	An equivalent time on a word processing task on the same computer.

Following training, the four visual search tasks were repeated for all groups.

To determine whether the visual lobe increased in size during the training, an ANalysis Of VAriance (ANOVA) was conducted for the lobe size for the three groups (1, 2 and 3) receiving lobe training. Over the five training trials significant effects of group ($F(2,15) = 11.05$, $p < 0.0011$) and training trial ($F(4,60) = 13.46$, $p < 0.001$) were found. To test whether the visual lobe training transferred to the visual search task, ANOVAs were performed on the mean search times for each fault type. For all four fault types, the patterns of the ANOVA results were similar. There were no group main effects ($p > 0.15$ in all cases), significant trial effects ($p > 0.05$ in all cases) and significant group x trial interactions ($p < 0.05$ for all cases except Rivet Crack where $p < 0.10$). **Table 8.1** shows the percentage improvements following training (i.e., the group x trial interactions) for each fault type. It can be seen that the two faults trained in the visual lobe training had the largest improvement. For the faults not trained by visual lobe training, the improvement was greater where there was more similarity to the visual lobe fault. Neutral training had a smaller amount of transfer, while no training, i.e., spending equivalent time on another computer task, had no beneficial effect.

Table 8.1 Percentage Improvement in Mean Search Times After Training for the Four Training Groups

	Rivet Faults		Area Faults	
Group	**Loose Rivet**	**Rivet Crack**	**Dent**	**Corrosion**
1. Loose Rivet Training	50.6	41.4	15.0	15.8
2. Dent Training	13.8	13.0	41.6	48.2
3. Neutral Training	5.8	18.5	16.3	26.2
4. No Training	3.9	5.0	-4.9	-12.7

Providing training, even just repeated practice, in rapidly detecting a fault in peripheral vision, did indeed increase the size of the area in which that fault could be detected in a single glimpse, i.e., the visual lobe. This increased visual lobe was not merely a result of increased familiarity with the experimental visual lobe task, as it transfers to a more realistic inspection task, visual search. For each fault type there was a 20-30% increase in lobe size over just five practice trials. There was a close correspondence between the training on actual faults (Groups 1 and 2) and improvement in search times, and even some improvement for training on a neutral fault, i.e., one which did not appear in any search tasks. No training, as expected, produced no effect.

8.6.2 How Effective is Feedback in Changing Search Strategy?

The objectives of this experiment were to evaluate the effectiveness of providing different types of feedback on the inspection search strategy. Performance feedback is traditional, i.e., how many defects were detected, or how long it took to detect all of the defects, but it is possible to give feedback on the process, or strategy, by which the inspector achieved these results. This may or may not help the inspector: evaluation is needed. Within this "cognitive" feedback, there are different ways in which the information can be presented to the inspector. Here we compared process measure feedback (e.g., how much of the area was covered; what percentage of fixations overlapped) to visual feedback of the scan path used by the inspector. The former gives hard numbers, the latter a visual pattern. The literature is silent on which is the better form of cognitive feedback, or whether either is better than traditional performance feedback. Thus, further evaluation is required.

All 24 subjects received familiarization training followed by a visual search task (Trial 1) which consisted of 75 search areas, each with either zero, one or two of the four faults (rivet crack, loose rivet, dent, corrosion). Following this visual search task the subjects were assigned to one of four groups:

Control Group	Three trials of 25 search areas with no feedback
Process Group	Three trials of 25 search areas with feedback or process measures after each trial
Visual Group	Three trials of 25 search areas with on-line visual feedback of search patterns during each trial.
Performance Group	Three trials of 25 search areas with feedback on speed and accuracy after each trial.

After these training interventions, subjects were given a second visual search task (Trial 2) of 75 search areas with no feedback. Comparison of Trial 1 with Trial 2 allowed an evaluation of the relative improvement with either practice only (control group) or the various types of feedback.

Both process measures and performance measures were analyzed for the four groups on the two trials. For all of the process measures, there was a group x trial interaction, showing that certain groups improved more than others. **Table 8.2** shows the percentage changes for each group from Trial 1 to Trial 2.

None of the changes for the control group were significant, showing that more practice did not change search strategy. The process group showed less fixations more widely spaced, and with less overlap, but the area covered decreased. The visual group showed a similar, if smaller, effect but one which did not result in a decreased coverage. Finally, the performance group showed no change in interfixation distance or area covered, but did give a reduced number of fixations more widely spaced. Clearly, cognitive feedback had the major effect on search strategy.

Table 8.2 Percentage Changes in Process Measures after Training for Feedback Experiment

	CONTROL GROUP	PROCESS GROUP	VISUAL GROUP	PERFORMANCE GROUP
Number of Fixations	+6.6	-54.1	-30.9	-27.8
Interfixation Distance	-5.2	+15.3	- 6.8	0
Percentage Overlap	+5.7	-55.1	-50.5	-18.0
Percent of Area Control	+2.1	- 7.0	+ 1.9	- 0.9

Performance measures, in contrast, showed trial effects for search times and stopping times but not for percentage defects detected. **Table 8.3** shows the percentage changes in each measure for each group.

Table 8.3 Percentage Performance Changes by Group for Feedback Experiment

	CONTROL GROUP	PROCESS GROUP	VISUAL GROUP	PERFORMANCE GROUP
Search Time	0	-26.0	-15.2	-37.1
Stopping Time	-4.9	-50.4	-35.4	-51.1
Percent Detected	-1.2	- 8.1	+ 1.6	- 1.0

As with process measures, the control group showed no effect of their practice. The process group was considerably faster at search, but appeared to reduce their stopping time too much, giving a decrease in defects detected. The visual group had smaller reductions in search and stopping times, and no change in percent defects detected. The performance group was able to speed up the most, and reduce their stopping times the most without sacrificing accuracy. Clearly, performance feedback enhances performance.

When feedback, or knowledge of results, was provided, behavior and performance changed in predictable ways. Subjects responded to the feedback given by improving most those aspects of the task included in the feedback. Thus, cognitive feedback helped subjects optimize their search strategy, while performance feedback gave the largest performance gains. All feedback groups changed their strategy in a similar manner, by having less fixations with less overlap. However, the dangers of feedback were illustrated when the process group made dramatic changes in strategy, resulting in slightly reduced coverage. This meant that their speed increase was not achieved with constant detection performance like the other groups, but with poorer detection performance. (A current extension of this experiment is evaluating the combination of feedback types to determine whether provision of both cognitive and performance feedback will yield larger improvements). Finally, it should be noted that more practice without knowledge of results (the control group) gave no improvements in strategy or performance. Only practice with feedback makes perfect.

8.6.3 How Effective is Feedforward in Changing Search Strategy?

Guidance or feedforward (Section 8.3.3) is a powerful tool in helping the trainee concentrate on the appropriate visual cues in the task. In airframe inspection, it comes from two sources: (1) general knowledge of the physics of aircraft structures and the environmental conditions which act to cause faults, and (2) specific guidance (e.g., on workcards, from co-workers) on which faults to expect in which parts of the structure. However, there is a danger with alerting an inspector to one type of defect and/or one area: other defects and areas may be de-emphasized giving poorer performance. The objective of this experiment was to evaluate general specific and combined feedforward in an aircraft inspection task. To measure whether improved performance on the cued defect was being obtained at the expense of other faults, two scenarios were devised, one emphasizing corrosion defects and the other emphasizing rivet defects.

Twenty-four subjects were given the familiarization training and then performed one visual search task under each scenario (Trial 1). Each task consisted of searching 55 search areas for the same four defects. The subjects were divided into the following four groups:

Control Group	No feedforward information
Prescriptive Guidance	Specific, prescriptive information on both scenarios, i.e., which type of defect was most common
Descriptive Guidance	General, descriptive information on both scenarios, i.e., the recent history of the aircraft's use
Combined Guidance	Both types of guidance

Scenario 1 emphasized corrosion defects, either by naming specific areas where corrosion was expected (lower part of fuselage) or by general history (aircraft employed carrying chemicals, and based at coastal airport). Scenario 2 emphasized rivet defects, but without any more detailed information on the location of the rivet defects. After this information, subjects inspected 55 more areas under each scenario (Trial 2).

The results showed no significant group or trial differences, or interaction, for the process measures for Scenario 1 and only a trial effect for percentage area covered for Scenario 2. As this latter change was less than 2%, there were essentially no effects of feedforward on search strategy.

For performance measures, there were group, trial and interaction effects on search time for Scenario 1, and group and trial effects for search and stopping times in Scenario 2. **Table 8.4** shows the percentage changes between Trial 1 and Trial 2.

Table 8.4 Percentage Performance Changes by Group for Feedforward Experiment

	Control Group	Prescriptive Group	Descriptive Group	Combined Groups
Scenario 1				
Search Time	+14.8	-23.4	-23.7	-30.8
Stopping Time	+ 7.3	-10.6	- 6.5	-31.1
Percent Detected	- 4.7	- 9.3	+ 1.6	+ 2.1
Scenario 2				
Search Time	-13.5	-11.3	-13.4	+ 0.7
Stopping Time	-28.8	-15.2	-17.1	-14.4
Percent Detected	0	+ 1.6	+ 2.8	+ 9.8

For Scenario 1, performance improved in terms of speed for all groups except the control group, while accuracy remained constant except for a decrease by the prescriptive group. This decrease was almost entirely due to reduced detection of the rivet defect, i.e., the one which was not cued. Scenario 2 gave peculiar results. Not only did the control group show the largest speed gains, but the combined group was the only condition to post improvements in accuracy. Even this accuracy gain was for the non-cued area faults!

It appears that feedforward information is difficult for subjects to use effectively. Strategy changes were almost all nonsignificant, while performance changes, even though significant, were mixed. Speed may improve, but accuracy may get worse (Scenario 1) or better (Scenario 2) for defects which are specifically called out. Much more needs to be known about the effects of feedforward on aircraft inspection before it can be recommended with any confidence. Even the calling out of specific defects on workcards, long thought to be a pre-requisite to effective inspection (e.g., Drury, et al., 1990) may need to be evaluated more closely.

8.7 DECISION MAKING ISSUES

When a defect has been located (visual search) a decision must be made as to its severity to determine the correct response (ignore, record for later repair, record for immediate repair). This decision is sometimes as simple as judging the free play in a control linkage, but more often it is a complex judgement. For example, a dent must be judged for size, depth and position, or corrosion for location, extent and severity. Inspectors gradually develop a mental picture (schema) corresponding not to any particular defect previously seen, but rather to a prototype of a defect at an action level. The two experiments reported here used a progressive part-training scheme to classify rivets (Section 8.7.1) or corrosion (Section 8.7.2). In each case there was a comparison between active and passive training. In each case a determination was made also as to whether the learning transfer from a decision task to a more complete inspection task involving both search and decision. The experiments differed in that the first concerned the accurate perception of each separate attribute of a defect, while the second examined integration of attributes into an overall schema of the defect.

8.7.1 Perception of Multiple Defect Attributes: Active or Passive Training?

The experimental objective was to compare well-designed active and passive training schemes for the task of correctly classifying individual attributes of a defect. The task was to judge rivets on a panel, where each rivet could have two levels of severity on each of three attributes: edge smoothness, out-of-round, and flatness. Eight combinations (combinations of two levels of each of three attributes) plus the six single-attribute defects gave a total of fourteen different defects.

Twelve subjects were assigned to one of two groups for training:

Passive Group — Subjects saw each defect five times in the center of the screen, followed by progressively more combinations of defect attributes, with the correct classification shown beside each defect.

Active Group — As for passive group, except that subjects had to enter the classification of each defect (with immediate feedback) rather than merely reading the classification.

After training, each subject was presented with fifteen examples of each defect (210 total defects). The defective rivet appeared in the center of the screen, with a classification response required of the subject. Following this decision task, an inspection task was given, again with 210 defects. In the inspection task, the subject first had to locate the defective rivet among the other rivets in a search area, and then give the classification response.

Significant differences between active and passive training for the percentage correct decisions were found for both the decision and inspection tasks, as shown in **Table 8.5**.

Table 8.5 Percent Correct Decisions for Attribute Perceptions

	PASSIVE TRAINING	ACTIVE TRAINING
Decision Task	79.0	83.8
Inspection Task	71.0	85.5

The active training scheme produced a clear advantage in accuracy in both tasks. Indeed, it gave a greater advantage when the decision was embedded in a more complex inspection task involving search and decision. Analysis of the search times in the inspection task showed no difference between the two training schemes, indicating that decision training does not improve search performance, only decision performance once search for the defective rivet has been completed.

Clearly, an active training scheme is preferred, even when compared to an equally well-designed progressive-part passive scheme. The inspector's active involvement in learning is essential.

8.7.2 Integration of Multiple Defect Attributes: Active or Passive Training?

In addition to merely classifying each attribute of a defect correctly, inspectors often need to combine information from more than one attribute when making a judgement. The objective of this experiment was to study the development of such schema, or combinations of attributes, under active and passive training schemes.

The task used areas of corrosion, each of which had three attributes: density, quantity and color. Density and color could be at three levels on the computer display, while quantity could be at two levels, giving eighteen combinations in all. Subjects had to reach an overall judgement of severity as low, moderate or high depending upon the particular combinations of levels of the attributes.

As in the previous experiment, two groups of six subjects each were used:

Passive Group	Subjects saw each of the 18 defects five times in the center of the screen, followed by combinations of defect attributes. The correct classification (L, M or H) was shown beside each defect.
Active Group	As for passive group except that subjects has to enter L, M or H for each defect, with immediate feedback.

The decision task consisted of 120 defects, each shown in the center of the screen, with the subject responding L, M or H after each defect. For the inspection task, each of the 120 defects was embedded in a screen consisting of the same rows and columns of rivets used in the previous experiment. Again, the subject had to first locate the defect and then classify it as L, M or H severity.

There were significant group differences on both decision and inspection tasks for percentage correct responses, and for information transmitted in bits/response. As in the previous experiment, there was no significant group effect on search times in the inspection task. **Table 8.6** shows the mean values.

Table 8.6 Performance by Group, Integration Experiment

	PASSIVE TRAINING	ACTIVE TRAINING
Percent Correct, Decision	77.4	87.8
Percent Correct, Inspection	73.5	86.3
Information Transmitted, Decision	0.76	1.06
Information Transmitted, Inspection	0.74	0.99

Active training was again significantly better on all measures, but here there was no extra improvement on the more complex task. In both decision and inspection, errors were roughly halved when the actively trained group was compared to the passively trained group. Clearly, active participation of the inspector is a key element in learning to integrate information from many attributes.

8.8 TRAINING METHOD FOR AIRCRAFT STRUCTURAL INSPECTION

In this section we outline a general training method that uses the key conclusions of the research and indicates how it can be applied to enhance training on a particular inspection task. The conclusions can be summed up as:

1. Visual lobe training improves mean search times and is generalizable across all faults within a fault set.
2. Feedback of process and performance measures can improve search strategy and reduce search time.
3. Feedforward information helps modify search strategy.
4. Active and progressive part-training help improve classification performance.

Figure 8.2 describes a general methodology for developing training programs for aircraft inspection. This section elaborates on the "training method" part of the methodology, giving as an example, the inspection task during a B-check on a DC-9.

This training methodology uses:

1. A mix of classroom and structured OJT
2. Visual lobe training for specific faults
3. Feedback of process and performance measures
4. Feedforward information
5. Active training for defect classification.

The methodology is explained using a section of a B-check for nose landing gear and wheel well inspection. There are three major components to this inspection:

1. Wheel well, doors, adjacent components
2. Nose gear assembly and installation
3. Nose gear tire and wheel assembly.

Table 8.7 presents a condensed overview of this entire B-check. The entire inspection has been broken down into two parts: structure and defects. The structure explains the component to be inspected, and the defect column lists the non-conformities to look for. This allows us to identify (a) the aircraft knowledge that the inspector should have, and (b) the defects that are being looked for.

8.9 CLASSROOM TRAINING

This should consist of the following components:

1. Information on Area. The trainees should learn the names and locations of all relevant parts of the area (listed under the structure column in **Table 8.7**) to be inspected. Active training methods used should make the trainee name and locate all the relevant parts/areas. This training should

also include functional knowledge about the various components, combined with adequate feedback on performance.

2. Information on Workcard Usage. This part should familiarize the trainees with the workcard. Steps include showing them how to use the workcard, knowledge of various procedures, e.g., checking depth of nicks or digs, releasing of nose steering bypass, etc. Information should be imparted on how to write non-routine repair cards. Again, an active training method is appropriate.

3. Examples of Defects in Each Area. A defect list must be generated from the workcard, giving a listing of all the defects that an inspector using the workcard must look for (Table 8.7). An effort should be made to collet samples, photographs, video tapes of all defects. An active training method where the trainee identifies, and classifies, each defect should be followed. Cognitive feedback should be provided during the training.

4. Visual Lobe Training. Visual lobe training can be provided on a simulator, similar to the one used in the evaluations (Section 8.6) for some of the visual defects like corrosion, visual damage, fluid leaks and worn parts. Now that actual photographs can be scanned readily into computer systems, defects can be placed easily in many places on the scanned visual image of the area to be inspected. Thus, creating a realistic simulator for visual lobe training is possible.

8.10 STRUCTURED ON-THE-JOB TRAINING

A structured on-the-job training methodology imparts a controlled training atmosphere in a work setting. Since it is expensive to produce a realistic simulator of the whole inspection task, we are constrained to have some training done on the job. The OJT method should involve the show-tell-do routine that includes demonstration by the expert inspector on an efficient way to inspect followed by the trainee inspecting the aircraft with subsequent feedback from the experienced inspector. This needs to have both performance and cognitive aspects, i.e., whether the search was successful, and whether it was performed using the most effective strategy. The experienced inspector should also provide feedforward on each area to the trainee in terms of what defects to look for, defect criticality, past history, etc.

In the space available in this report, more depth cannot be included concerning the detailed application of the research findings to aircraft inspection. However, the level chosen for presentation here does demonstrate the principles involved and how they can be applied with minimal hardware requirements.

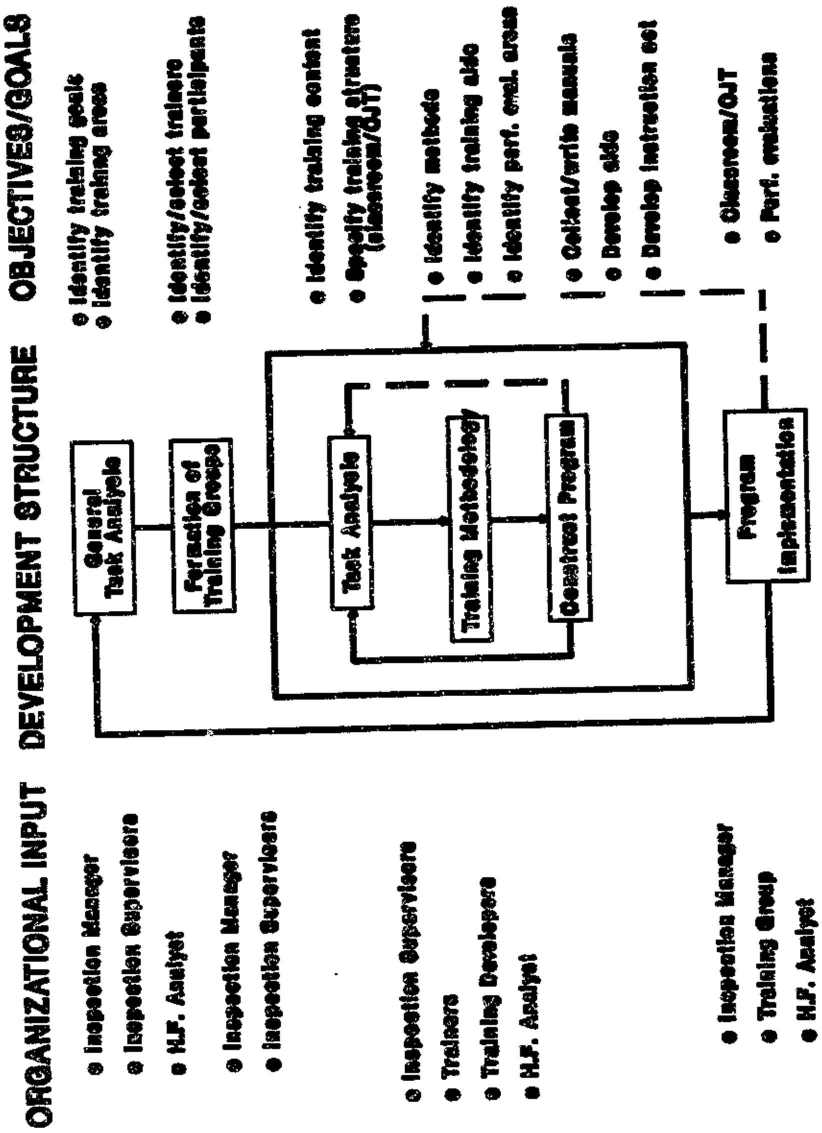


Figure 8.2 Model for Training Program Development in Commercial Aviation

Table 8.7 Nose Landing Gear and Wheel Weli Inspection (B-Check)

WHEEL WELL, DOORS, ADJACENT COMPONENTS		NOSE GEAR ASSEMBLY & INSTALLATION		NOSE GEAR TIRES & WHEEL ASSEMBLY	
Structure	Defects	Structure	Defects	Structure	Defects
1. Wheel well hydraulic tubing conduits	• Condition • Corrosion • Fluid leakage	1. NLG shock stout, bracestrut, torque arm, ground sensing mechanism, cables, actuating cylinder, linkages, springs	• corrosion • visual damage • nicks & dings** • fluid leaks • security	1. Wheel hubs values, tie bolts	• condition • corrosion
2. Wheel well doors linkages springs, stop cables, drive rods and hinges	• condition • visual damage • corrosion • security	2. Landing gear shock strut	• check for normal extension • cleanliness • clean exposed portion of piston with red hydraulic oil & wipe dry	2. Tires	• excessive wear • oil soaking • correct pressure - only after 2 hrs of parking • reinflate with N2 only
3. Downlock markings	• general condition • cleanliness	3. Nose steering mechanism	• condition • leakage • worn cables • release of nose steering bypass • check spring landed to steering position	3. Water deflector assembly	• damage • security of installation
4. NLG alignment spotlight	• check	4. Torque links	• loose bushings & bolts • worn bushings & bolts		
5. NLG taxi light	• cleanliness • filament condition • security of assembly	5. Landing gear lock pins & red warning streamers	• condition • secure attachment of streamers to lock pins • length of streamers should be 24-32" long		
6. NLG doors	• check doors are closed • check doors are secure *procedure given				**<0.005" repair 0.005-0.03 blend, smooth & paint >0.003 notifying eng.
7. Aircraft wheel checking placard *location given	• condition • security				
8. Nose tire pressure placard *location given	• condition • security				
9. Uplock and downlock proximity sensors	• condition-clean if necessary • security				

8.11 CONCLUSIONS

This research program has used observations of the current training of visual inspection, and the principles of effective training to derive experimental evaluations of specific interventions. The experiments had (generally) highly successful outcomes, showing that many of the interventions can indeed be applied to visual inspection training. An example is provided which outlines how these findings can be applied to an existing task. The next challenge is to devise and implement detailed inspection training programs based on these findings and evaluate these new programs on the hangar floor. As a side-benefit of the research, a simulation program is now available to allow rapid evaluation of other training interventions without disruption of on-going inspection activities.

8.12 REFERENCES

Bartram, D. (1988). The design and evaluation of a computer-based training system for foreign destination coding desk operators. In Patrick, J. and Duncan, E. D. (Eds), *Training, human decision making and control.* Amsterdam: Elsevier.

Brown, J. S., Burton, R. R. and DeKleer, J. (1982). Pedagogical, natural language and knowledge engineering techniques in SOPHIE I, II and III. In Sleeman, D. H. and Brown, J. S. (Eds.),

Christal, R. E. (1974). *The United States Air Force occupational research project.* (Tech. Rep. No. AFHRL-Tr-73-75). Texas: Air Force Human Resources Laboratory, Occupation Research Division.

Clancey, W. J. and Soloway, E. (Eds.). (1990). *Artificial intelligence and learning environments.* Cambridge, MA: Bradford/MIT Press.

Czaja, S. J. and Drury, C. G. (1981). Training programs for inspection, *Human Factors, 23, 4,* 473-484.

Desormiere, D. (1990). Impact of new general aircraft on the maintenance environment and work procedures. *Proceedings of the Fifth FAA Meeting on Human Factors in Aircraft Maintenance and Inspection*, 124-134.

Doherty, M. E. and Balzer, W. K. (1988). Cognitive feedback. In Brehmer, B. and Joyce, C. R. B. (Eds.), *Human judgement: The SJT view.* New York: North Holland, 163-195.

Drury, C. G. (1990). Visual search in industrial inspection. In Brogan, D. (Ed.), *Visual Search.* London: Taylor & Francis, 263-276.

Drury, C. G. and Gramopadhye, A. K. (1992). Training for visual inspection: Controlled studies and field implications. *Proceedings of the Seventh FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, 135-146.

Drury, C. G. and Gramopadhye, A. K. (1990). Training for visual inspection. *Final Report of the Third FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, 149-164.

Drury, C. G. and Kleiner, B. M. (1990). Training in industrial environments. *Proceedings of the 23rd Annual Conference of the Human Factors Association of Canada*, 99-108.

Drury, C. G., Prabhu, P. V. and Gramopadhye, A. K. (1990). Task analysis of aircraft inspection activities: methods and findings. *Proceedings of the Human Factors Society 34th Annual Meeting*, 1181-1185.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327-358.

Goldsby, R. P. (1991). Effects of automation in maintenance. *Proceedings of the Fifth FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, 103-123.

Gramopadhye, A. K. (1992). *Training for Visual Inspection*. Unpublished doctoral dissertation, State University of New York at Buffalo.

Hammond, K. R. and Summers, D. (1972). Cognitive Control. *Psychological Review, 79, 1*, 58-67.

Johnson, W. B. (1990). Advanced technology training for aviation maintenance. *Final Report of the Third FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, 115-134.

Johnson, W. B., Norton, J. E. and Utsman, L. G. (1992). New technology for the schoolhouse and flightline maintenance environments. *Proceedings of the Seventh FAA Meeting on Human Factor Issues in Aircraft Maintenance and Inspection*, 93-100.

Kello, J. (1990). The role of simulation in technical and team-skills training. *Proceedings of the FAA Training Technology Symposium*, 133-148.

Kurland, L. and Huggins, A. W. F. (1990). Training trouble-shooting using advanced AI-based training technology. *Proceedings of the FAA Training Technology Symposium*, 133-148.

Latorella, K. A., Gramopadhye, A. K., Prabhu, P. V., Drury, C. G., Smith, M. A. and Shanahan, D. E. (1992). Computer-simulated aircraft inspection tasks for off-line experimentation. *Proceedings of the Human Factors Society 36th Annual Meeting*, 92-96.

Lesgold, A. M. (1990). Intelligent computer aids for practice of complex troubleshooting. *Proceedings of the FAA Training Technology Symposium*, 116-132.

Lesgold, A. M., Lajoie, S. P., Bunzo, M. and Eggan, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In Larkin, J., Chababy, R. and Scheffic, C. (Eds), *Computer assisted instruction and intelligent tutoring systems: Establishing communication and collaboration*. Hillsdale, NJ: Lawrence Erlbaum.

Lukins, J. (1990). Automation and maintenance training. *FAA Training Technology Symposium*, 66-72.

Lutzinger, R. T. (1989). Day-to-day problems in air carrier maintenance and inspection operations. *Final Report on FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, A-17 - A-25.

Naylor, J. C. and Briggs, G. E. (1963). Effects of task complexity and task organization on the relative efficiency of part and whole training methods. *Journal of Experimental Psychology, 65*, 217-224.

Norton, J. E. (1992). Intelligent simulator for maintenance training. *Proceedings of the Seventh FAA Meeting on Human Factor Issues in Aircraft Maintenance and Inspection*, 101-106.

Orlansky, J. and String, J. (1979). *Cost-effectiveness of computer-based instruction in military training* (IDA Paper P-1375). Arlington, VA: Institute for Defense Analyses and Technology Division.

Patrick J. (1992). *Training Research and Practice*, New York: Academic Press.

Payne, H. E. (1990). Army training 21: Concepts of the future. *Proceedings of the FAA Training Technology Symposium*, 174-179.

Reigeluth, C. M. and Stein, F. S. (1983). The elaboration theory of instruction. In Reigeluth, C. M. (Ed.), *Instructional design theories and models. An overview of their current status.* Hillsdale, NJ: Lawrence Erlbaum.

Rice, J. W. (1990). Use of 3-D presentations in maintenance training. *Final Report of the Third FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, 135-139.

Skinner, M. (1990). Aviation maintenance practices at British Airways. *Proceedings of the Fourth FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection*, 143-153.

Shepherd, W. and Parker, J. (1990). Human factors issues in aircraft maintenance and inspection "Training issues." *Final Report of the Third FAA Meeting on Human Factors Issues in Aircraft Maintenance and Inspection.*

Walter, D. (1990). An innovative approach to NDT inspector training at Boeing. *Final Report of the Third FAA Meeting on Human Factor Issues in Aircraft Maintenance and Inspection*, 52-71.

Wirstaad, J. (1988). On knowledge structures for process operators. In Goodstein, L. P., Andersen, H. P. and Olsen, S. E. (Eds.), *Tasks, errors and mental models.* New York: Taylor & Francis, 50-69.

END

FILMED

DATE: 10-93

DTIC