

Phase II Progress Report (1993)

Galaxy Scientific Corporation

Pleasantville, NJ 08323

Office of Aviation Medicine

Federal Aviation Administration

Washington, DC 20591

April 1993

Chapter One

Executive Summary

1.0 SUMMARY

This report details the second phase of the Office of Aviation Medicine (AAM) research program on Human Factors in Aviation Maintenance. This on-going research program enhances human performance in the aviation maintenance system to ensure continuing flight safety and operational efficiency. The research program, as reported in the [Phase I report](#) (Shepherd, et al., 1991) was initially planned to have 4 steps, with feedback mechanisms as shown in [Figure 1.1](#). Phase I focused on preliminary investigation and problem definition of human performance in airline maintenance environments. This Phase II report describes research that centers on the development of hardware and software prototypes with potential to enhance human performance in aircraft maintenance. (In this report the activities subsumed under "inspection" are considered to be part of "aircraft maintenance").

An Ongoing Research & Development Program

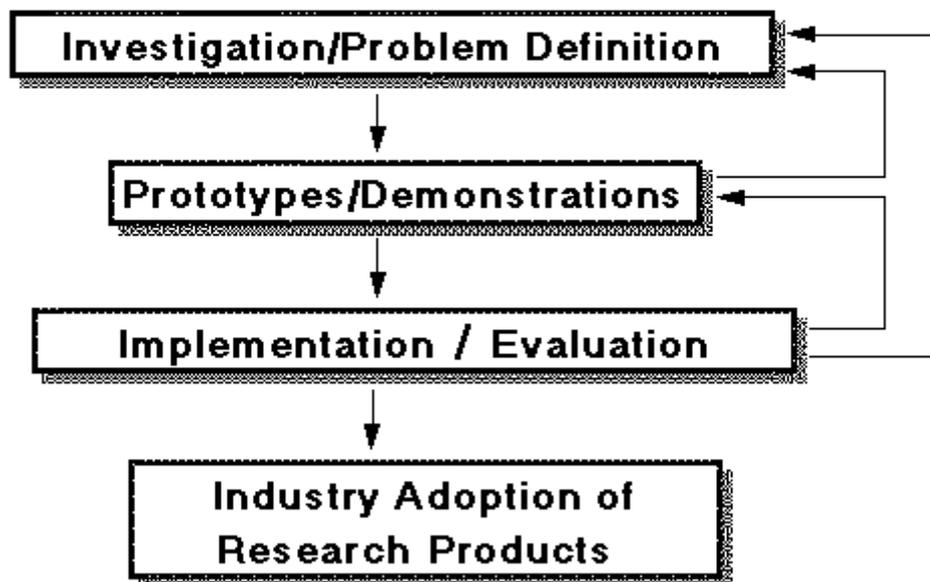


Figure 1.1 The Research Program

The research reported here has been conducted by a multi-disciplinary team of scientists and engineers from industry and academia. The research team has worked in very close cooperation with the international aviation industry, mostly with US airlines and aviation manufacturers.

This report includes seven chapters and can be considered as an edited volume in that each chapter is written to stand alone as the work of each research group.

1.1 CONTINUING RATIONALE

Shepherd et al. (1991) and Shepherd & Johnson (1991), offered an extensive description of the rationale for the research program. These reports described the complexity of the total aviation maintenance system and the role of the human within the system. Increased maintenance workload, caused in part by an increased level of air carrier operations, is one reason to focus on improving aviation maintenance technician (AMT) performance. The challenge of providing continuing air worthiness of the aging fleet while developing knowledge and skills for maintaining new technology aircraft places a burden on airline maintenance organizations.

Phase I research investigated methods for enhancing human performance in aircraft maintenance. There are ample reasons for continuing these investigations. For example, operations will continue to increase. Airlines will fly more hours with the same fleet sizes. Thus, there will be less time for maintenance and greater stress on the fleet. Therefore, enhancing human performance in maintenance continues to be an important priority.

Resources are finite. Airlines, during 1991-92, have not been profitable. Since the [Phase I report](#) was published, major air carriers such as Pan Am and regional carriers such as Midway Airlines have gone out of business. Other carriers have suffered record financial losses and face uncertain futures. Airlines recognize the criticality of cost control in every aspect of their operation. However, cost control cannot jeopardize safety. This research program recognizes that the enhancement of human performance in maintenance is critical to the safety and efficiency of air carrier operations.

1.1.1 Integration of Human Factors Research Efforts

Research to enhance human performance in aircraft maintenance can focus on several dimensions, such as the human, the tools, the work place, work procedures, and management philosophies. The research must be useful to maintenance practitioners as well as to the human factors research community. This report, therefore, has practical as well as scientific value.

1.2 ADVANCED TECHNOLOGY TRAINING (Chapter Two)

Advanced technology training combines artificial intelligence technology with conventional computer-based training. The technology was described extensively in the [Phase I report](#) (Shepherd, et al., 1991) and elsewhere (Johnson & Norton, 1991 and Johnson & Norton, in press).

This chapter describes the continuing effort that has converted a Phase I training prototype to a fully operational advanced technology training system for the Boeing 767-300 environmental control system (ECS). The system is simulation-based in that it permits the user to access and operate all panels, controls, and built-in-test equipment of the [ECS](#). [Figure 1.2](#) shows the human-computer interface for the [ECS](#).

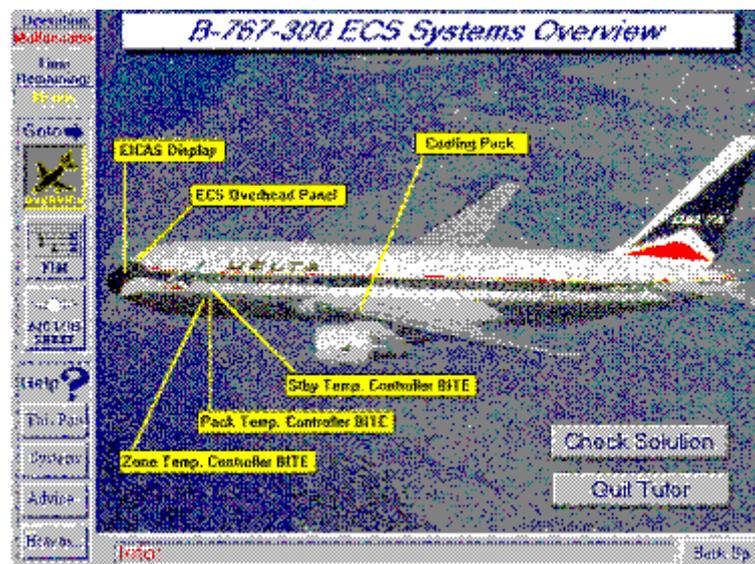


Figure 1.2 Environmental Control System Tutor

The [ECS](#) trainer is unique not only because of the simulation but also because of the robust software used for modeling student performance and providing feedback, explanation, and remediation. These modeling features are described in [Chapter Two](#).

In addition to providing simulation, the [ECS](#) Tutor provides on-line access to the training manual for the [ECS](#). The software makes it easy for the student to use the manual during training. This research is preparing for development of an integrated information system which can provide not only training but also real-time job aiding and maintenance documentation. Research related to the concept of on-line documentation is also described in [Chapter Four](#).

The chapter also describes the process of formative evaluation that took place as the training systems underwent many iterations with software engineers and training professionals at Delta Air Lines and at Clayton State College. A substantive training effectiveness evaluation will be conducted at Delta.

Finally, as advanced technology training systems become more commonplace, it is likely that they can be used for [AMT](#) certification. Therefore, this chapter also reports on the research implications of the pending changes to Federal Aviation Regulation (FAR) Parts 65 and 147.

1.3 ADVANCED TECHNOLOGY MAINTENANCE JOB AIDS (Chapter Three)

This chapter addresses existing approaches to job aiding in maintenance, the drawbacks to such approaches, the prospects for using emerging technologies to develop maintenance job aiding systems, and the impact of emerging technologies on human performance. There were two major themes to the research: 1) many previous attempts at building maintenance job aids consisted of trying to replace human expertise with machine expertise; and, 2) problems with such approaches have led to a reconsideration of the skills and abilities of human operators and ways to capitalize on them.

Accordingly, the chapter calls for a 'cooperative system' approach to designing such systems; a cooperative system is one in which a human and a computer are actively involved in the problem solving process. The chapter presents a study which used this approach in developing a job aid. Some of the results of the study that are relevant to designing maintenance job aids and integrated information systems (Johnson & Norton, 1992 a & b) are also presented. Finally, a research and development plan for building a maintenance job aid for aircraft maintenance is discussed.

1.3.1 Human Performance Implications of Artificial Intelligence Approaches

The bulk of the job aiding systems encountered in a literature review used artificial intelligence and expert systems techniques. While artificial intelligence techniques can provide a computer with powerful problem solving abilities, job aiding systems which rely solely on such techniques often meet with limited operational success. One of the reasons for such limited success is that the computer is supposed to embody the knowledge and abilities of a human expert, when, in fact, such systems are necessarily incomplete. Because builders of expert systems cannot capture all of the human expert's knowledge about a task, such systems often draw erroneous conclusions. Therefore, the operator must have enough expertise to realize that the computer is wrong; the problem is that the operator will not develop such expertise unless he/she is actively involved in the problem solving process. However, these problems do not exclude artificial intelligence techniques from use in operational job aids. Rather, the question is one of emphasis: instead of using artificial intelligence techniques as the foundation for a job aid, they should be used in conjunction with other methods of performance aiding (e.g., representation aiding).

1.3.2 Human Performance with a Cooperative System

A research study (Layton, 1992) which investigated human performance with three forms of a cooperative system provided some interesting insights into how such systems affect human behavior. This system was designed to assist commercial airline pilots and dispatchers in enroute flight planning. (Figure 1.3 depicts a portion of the system displays and controls.) This research has provided some interesting insights into the ways in which job aiding tools affect human performance.

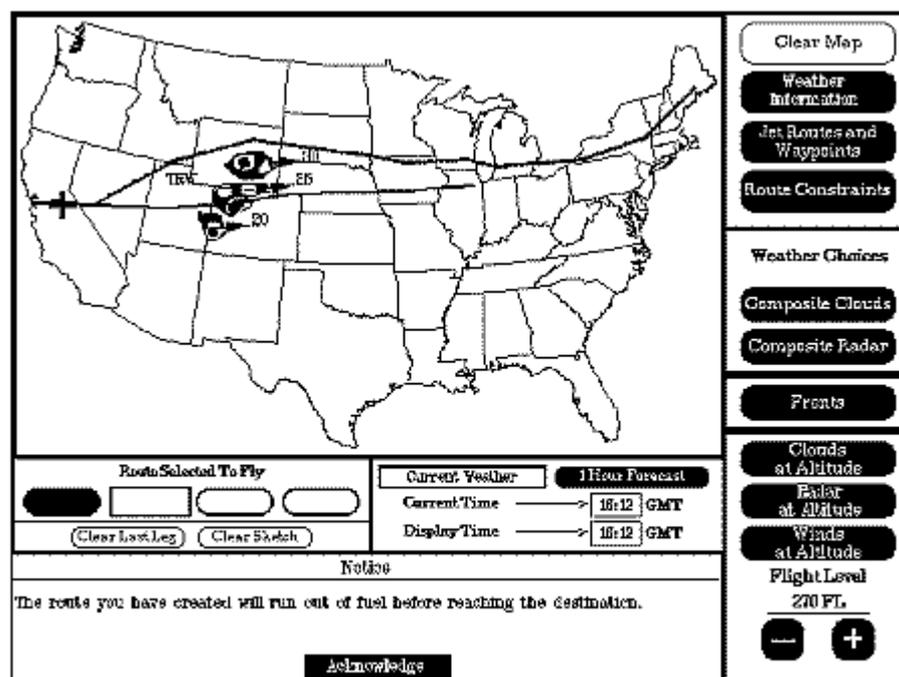


Figure 1.3 Enroute flight planning cooperative system

1.3.3 Research and Development Plan

A three-phase plan for developing an aviation safety inspector job aid using cooperative system techniques was developed. Initial interactions with the Flight Standards Service suggest that the job aid will assist inspectors in researching operator information and documenting inspection activities. The system will link inspection forms so that information that is entered into one form will

automatically be entered into the other forms being used. The system will also provide the ability to search for information using on-line documentation. Such documentation may include the federal aviation regulations, advisory circulars, airworthiness directives, [FAA](#) policies, and operator-specific information. The system will have a cellular modem capability so that it can connect to the Flight Standards mainframe system and will likely use a [CD ROM](#) to store much of the on-line documentation.

1.4 DIGITAL DOCUMENTATION (Chapter Four)

Maintenance personnel are often overwhelmed with the amount of technical documentation necessary to accomplish a given task. The information comes from a variety of sources including company and manufacturer's manuals, and government documents, like advisory circulars or regulations. Currently most maintenance documentation exists as hard copy or microfiche. The task of keeping these databases current is very time consuming and expensive.

This research task, called the Hypermedia Information System, (HIS) shown in [Figure 1.4](#), studies advanced technology software and hardware techniques for information storage and retrieval. The primary products of this research will be techniques for the development and use of large information sources on small portable computer systems.

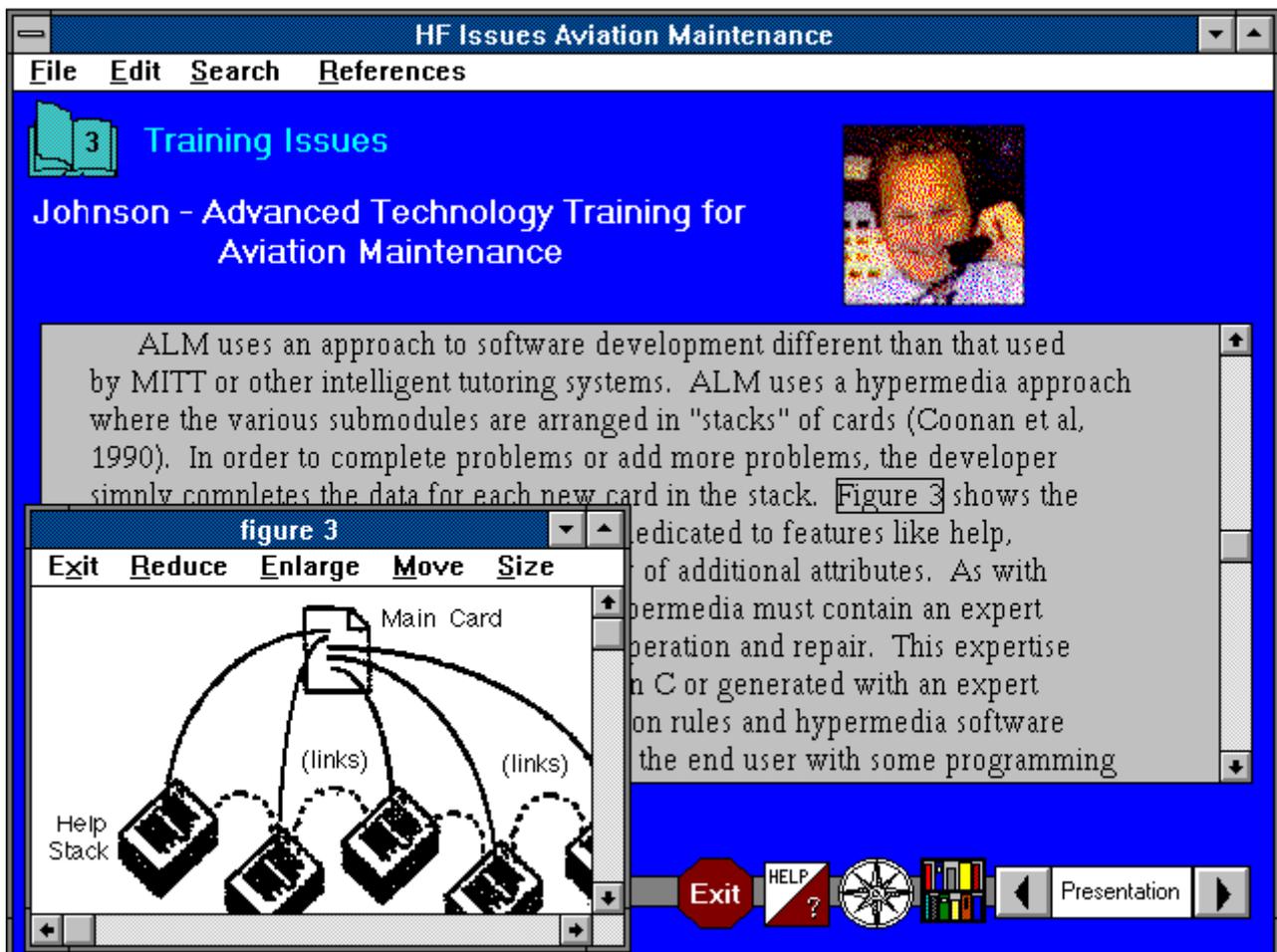


Figure 1.4 Hypermedia Information System

The term "hypermedia" refers to a combination of text, graphics, animation, audio, and video to convey information. Such information bases are designed to be accessed easily, usually in a non-

linear fashion. This hypermedia research will make it possible for a technician to access a manual for all media and information to complete a job. The research fosters co-development of integrated information systems (Johnson & Norton 1992, a & b) that provide training, job-aiding, and on-line documentation.

A by-product of the research is the development of a hypermedia information system for all technical publications from the Aviation Medicine Human Factors in Aviation Maintenance research program. Ultimately this digital source of information will be published on a CD-ROM (Compact Disc-Read Only Memory).

The proceedings of one of the first six conferences on aviation maintenance human factors has already been prepared for distribution as a digital document. The seventh conference will be the first time that the meeting proceedings will be distributed in digital format at the meeting. The software developed through this research effort has facilitated the timely publication of such digital documentation.

1.5 HUMAN RELIABILITY IN AIRCRAFT INSPECTION (Chapter Five)

The research related to improving human reliability in aircraft inspection built upon the solid task analytic foundation derived under Phase I. The chapter describes two studies: one study related to the re-design of workcards for inspection and the other a study of the lighting environment for inspection. Both studies offer practical human factors guidelines applicable to these topics. The chapter also describes a plan to consider human-computer interface issues applicable to computer-based maintenance aids.

This chapter also describes a series of laboratory experiments that evaluate the effects of time pressure on inspection and the improvement of training techniques for visual inspection. The chapter describes a study of the classification of human error in inspection. The classification is particularly valuable in its review of many scientific studies of human error. These studies form the basis for the team's development and presentation of system models of human error in inspection. These models provide the means to understand, predict, report and manage inspection errors.

Inspection is information processing. The chapter reports research on the design of information flow in the inspection environment. The research helps to determine what, when, and how to present information to the inspector. Experimental results are presented regarding optimal methods of information presentation in inspection tasks.

Chapter Five also describes a joint study of inspection practices in the UK and USA. The comparative study observed that management structures of maintenance and inspection are more closely intermeshed in the UK than in the US. Other differences and rationales are reported.

1.6 GUIDELINES FOR HUMAN FACTORS IN MAINTENANCE (Chapter Six)

Human Factors principles are often derived in laboratory studies of procedures, equipment, effects of time, temperature, lighting and other variables. Much of the information derived from these studies is reported for scientists, psychologists, and engineers for academic applications. This task is reviewing the human factors literature from a wide variety of parallel and similar areas to aircraft maintenance. All of the research results from the Aviation Medicine Aircraft Maintenance Human Factors program will be combined with this information base to produce a Human Factors Guide for Aircraft Maintenance. This guide promises to be useful to airline maintenance management system designers, [FAA](#) oversight personnel, and others as they strive to improve human performance in the maintenance system. [Chapter Six](#) offers an example chapter from the Human Factors Guide.

1.7 CREW RESOURCE MANAGEMENT FOR MAINTENANCE: EVALUATION

OF A TRAINING PROGRAM (Chapter Seven)

Phase I (Shepherd, et al., 1991) reported on management-worker communications in the aviation maintenance environment. Phase II research has shifted focus to the effects of crew resource management (CRM) training in an airline maintenance environment. The research has concentrated on communication among maintenance crews. The researcher participated in the evaluation of the effectiveness of a particular airline's CRM training for maintenance personnel and in the post-training performance effects on maintenance managers and technicians.

The CRM course acceptance has been very high. In fact maintenance crews have demonstrated greater acceptance of the CRM principles than have flight crews. The research indicates that relevant attitudes about CRM improved immediately after training. Course attendees have reported that the CRM principles have caused them to be more actively involved in all maintenance decision making.

The CRM evaluation research is valuable in that it has created instruments and criteria to measure post-training maintenance performance. These measures will be helpful to assess the training and cost effectiveness of such human performance enhancement courses.

1.8 CONTINUED COMMUNICATION

The seven workshops that have been conducted to date under the Aviation Medicine research program have facilitated communication between researchers and industry. The immediate application of some of the research activities described above will allow the industry to increase reliability and lower costs. The Office of Aviation Medicine (AAM) intends to continue sponsorship of the workshops throughout the duration of the research program.

The participation and cooperation of the airline industry has been instrumental to the AAM research program. Air carriers, manufacturers, and schools have been extremely cooperative and helpful. This cooperation is gratefully acknowledged.

1.9 REFERENCES

Johnson, W.B. (1992). Challenges and solutions for airway facilities in the 1990's. *Proceedings of the FAA/NASA Conference on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance*. Daytona Beach, FL: Embry Riddle Aeronautical University.

Johnson, W.B. & Norton (1992 a). Integrated information for maintenance training, aiding, and on-line documentation. *Proceedings of the 36th Annual Meeting of the Human Factors Society*. Atlanta, GA: The Human Factors Society.

Johnson, W.B. & Norton, J.E. (In press). Modeling student performance in diagnostic tasks: a decade of evolution. In V. Shute and W. Regian (Eds.), *Cognitive Approaches to Automated Instruction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Johnson, W.B. & Norton, J.E. (1992 b). Integrated systems for training, aiding, and information retrieval. *Proceedings of the East-West Conference on Emerging Technology in Education*. Moscow, Russia, 168-175.

Layton, C. F. (1992). *An investigation of the effects of cognitive tools on human adaptive planning in the domain of enroute flight planning*. Doctoral dissertation, The Ohio State University, Columbus, OH.

Johnson, W.B. & Norton, J.E. (1991). Using intelligent simulation to enhance human performance in aircraft maintenance. *Proceedings of the 1991 International Conference on Aging Aircraft and Structural Airworthiness*. Washington, DC: Federal Aviation Administration and National

Aeronautics and Space Administration.

Shepherd, W.T. & Johnson, W.B. (1991). Aircraft maintenance challenges and human factors solutions. *Proceedings of the ICAO Flight Safety and Human Factors Seminar*. Bangkok, Thailand. 204-220.

Shepherd, W.T., Johnson, W.B., Drury, C.G., Taylor, J.C. & Berninger, D. (1991). [*Human Factors in Aviation Maintenance Phase 1: Progress Report*](#) (DOT/FAA/AM-91/16). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration.

Chapter Two

Advanced Technology Training for Aviation Maintenance

2.0 INTRODUCTION

As technology advances, the job of the aircraft maintenance technician (AMT) becomes increasingly difficult. The [AMT](#) must deal with new technology (digital components, composite materials, etc.) as well as an aging aircraft fleet. Concurrently, the [AMT](#) work force dwindles ([Shepherd, et al., 1991](#)). Therefore, maintenance training must respond to these challenges by increasing the effectiveness of the current work force. This chapter outlines how the Office of Aviation Medicine (AAM) uses advanced technology training to address these issues.

2.1 RESEARCH PHASES

The advanced technology training research began in January of 1990. This earlier phase of the work assessed the status of training technology for maintenance technicians. Based upon this assessment, the AAM built a prototype intelligent tutoring system (ITS) for the Environmental Control System (ECS) ([Shepherd, et al., 1991](#)). See (Polson and Richardson, 1988) for more information on [ITS](#).

The current phase of the research expands this prototype to be an operational tutoring system for the [ECS](#). In order to provide a measure of the effectiveness of the Tutor, this phase also designs an evaluation of the tutor. Finally, as advanced technology training systems become more commonplace, it is likely that they can be used for [AMT](#) certification. Therefore, this phase also reports on the research implications of the mechanic certification rules changes to Federal Aviation Regulation (FAR) Parts 65 and 147.

It is important that the Tutor be an effective training tool. Despite formative evaluation of the Tutor throughout the second phase, a formal evaluation will be conducted during the third phase. This will evaluate the Tutor for user acceptance, training effectiveness, and cost effectiveness. In addition, the third phase of the project will investigate the use of advanced technology training for psychomotor activities.

2.2 ADVANCED TECHNOLOGY TRAINING TUTOR

The Advanced Technology Training research developed an operational tutoring system for the [ECS](#) of the Boeing 767-300, as shown in [Figure 2.1](#). This section describes the features of the Tutor, the design of the Tutor, and the lessons learned while developing the operational Tutor.



Figure 2.1 ECS Overview Screen

2.2.1 Description of the ECS Tutor

In the [ECS](#) Tutor, the student interacts with panels, controls, test equipment, manuals, and displays. These graphics are meant to simulate the "look and feel" of the real [ECS](#) components. A simulation of the [ECS](#) responds to the student's actions by updating the appropriate data values on the Engine Indicating Crew Alerting System (EICAS) display and the Overhead Panel, as shown in [Figure 2.2](#) and [Figure 2.3](#), respectively.

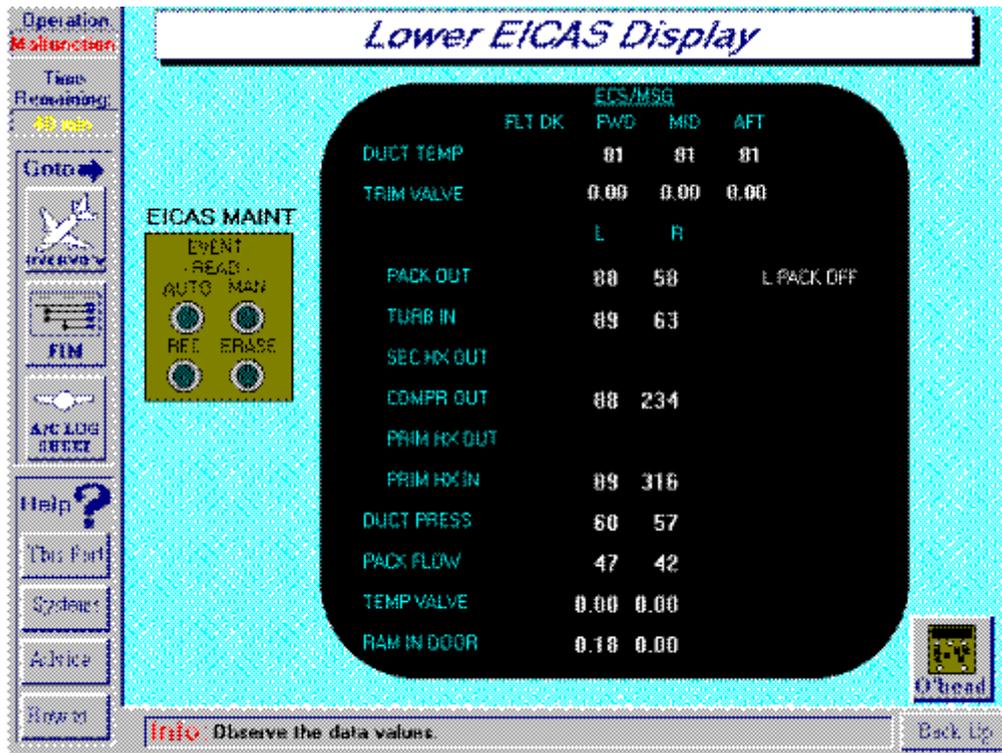


Figure 2.2 EICAS Display



Figure 2.3 Overhead Panel

The system operates in two distinct modes: Normal Operation and Malfunction. During normal operation, the ECS simulation lets the student see the proper operation of the system. During a malfunction, the ECS will exhibit the symptoms associated with the current malfunction. The

student controls whether the system operates in normal mode or malfunction mode.

Regardless of the mode, the student's interaction with the system is similar. In both modes, the student has access to all of the tools needed to operate the [ECS](#). The student interacts with control panels, information displays, built-in test equipment, component information, etc. However, malfunction mode is unique in several respects.

As a malfunction begins, the student sees a description of the problem, with the accompanying fault code. After seeing the description, the student may order replacement parts to be delivered to the plane. If the student does not order a part that they need, when they go to replace the part they must wait for the part to arrive. The Tutor simulates this wait by removing 15 minutes from the time that remains.

During a malfunction, the student has access to the Fault Isolation Manual (FIM), shown in [Figure 2.4](#). The [FIM](#) is the fault tree that the [AMT](#) follows while troubleshooting. It outlines the tests and procedures that should be performed while diagnosing a failure. While troubleshooting from the [FIM](#), the student must perform the tests in the order prescribed by the [FIM](#).

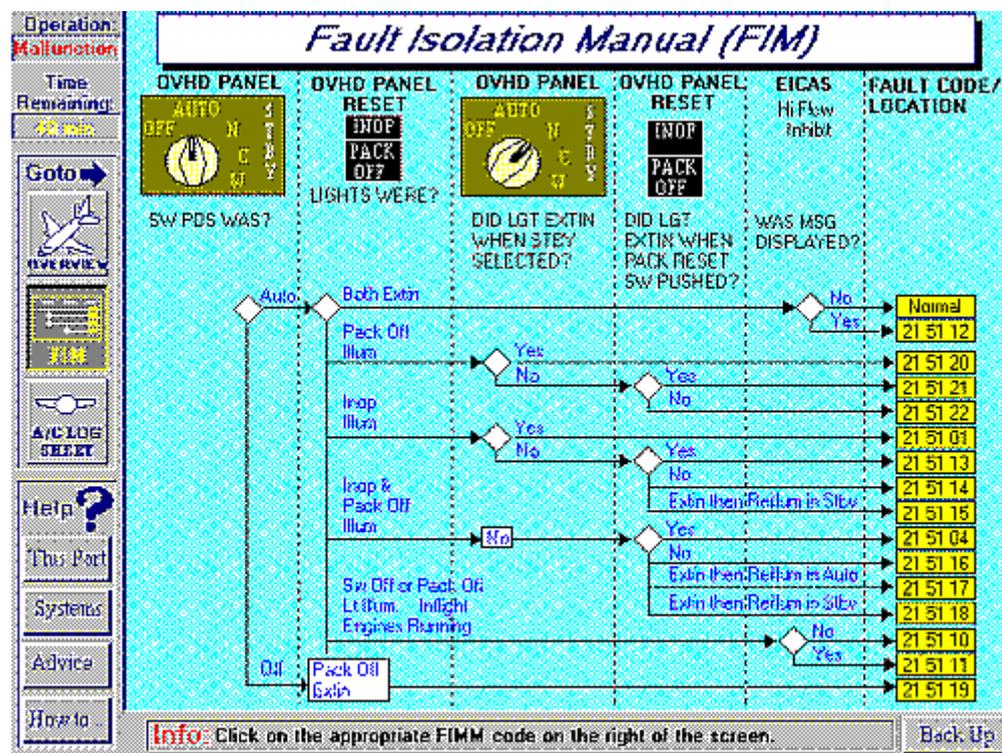


Figure 2.4 Fault Isolation Manual Display

Even though the [FIM](#) is the recommended method of troubleshooting, the [ECS](#) Tutor supports a more flexible way to troubleshoot. The Tutor provides a schematic of the cooling pack from which the student may troubleshoot. The student selects a component (in any order) from the schematic and then chooses whether to Inspect, Test, or Replace the component.

In both the [FIM](#) and the cooling pack schematic, the Tutor records the student's actions. At any time, the student may ask for advice on how to troubleshoot the current malfunction. The tutor compares the student's actions with the actions of an expert. The Tutor suggests the appropriate next step to the student.

When the student replaces a faulty component, the Tutor updates the simulation to reflect proper operation. However, just as on the flight line, the student must verify that the replacement was effective. When the students are confident of success, they can have the Tutor check their solution.

2.2.2 Design of the ECS Tutor

During the design of the [ECS](#) Tutor, several separate design issues were addressed. These issues include interface design, instructional design, and simulation design. This section also describes the tools used to implement this design.

2.2.2.1 Interface Design

The design of the [ECS](#) Tutor ensures that the majority of relevant troubleshooting information is only one mouse click away from the primary display, shown in [Figure 2.5](#). From this Overview display students may access the displays, controls, and components that they need to troubleshoot a malfunction.

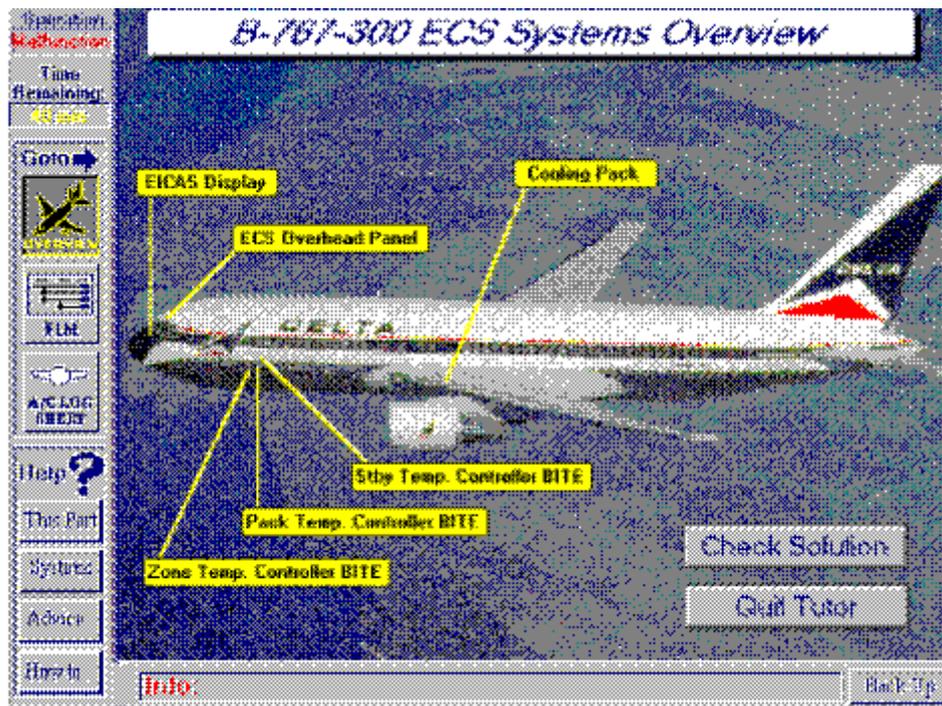


Figure 2.5 ECS Overview Display

Every troubleshooting screen in the Tutor adheres to a standard format, which consists of a grey border on the edge of the screen. The border contains functions that are useful for all troubleshooting screens. The left-hand border consists of three different areas - the Status area, the Navigation (or "Go To") area, and the Help Area. The bottom border contains a Message Area.

The top-most portion of the border is the Status area. This area notifies the student of the current operating mode - "Normal" or "Malfunction". During a Malfunction, this area also indicates the amount of time that remains for troubleshooting. The center portion of the border is the Navigation area. These three buttons allow the student to immediately go to one of three commonly used screens - the Overview Screen, the Fault Isolation Manual, or the Aircraft Log Sheet.

The lower portion of the border is the Help area. This area assists the student in four different ways:

- How to use the current screen
- Information about mechanical, electrical, pneumatic concepts
- How to navigate through the system
- What to do next - advice during troubleshooting

Finally, the border that runs along the bottom of the screen is the Message area. As the student moves the cursor across selectable regions, the Message area shows additional information about that selectable region. For example, the Overview display contains a button labeled "Check Solution". When the cursor enters this button, the Message area will read "Info: Click this button when you think you have solved the problem."

2.2.2.2 Instructional Design

The instructional design of the [ECS](#) Tutor was improved in two ways. First, a lesson orientation provides more context for each of the malfunctions. Also, the "over-the-shoulder" advice helps students who are not making progress toward a solution. The following sections describe each in more detail.

2.2.2.2.1 Lesson Orientation

The instructional design for the [ECS](#) Tutor centers around the idea of a lesson, where a lesson is a logical collection of malfunctions. Since the Fault Isolation Manual (FIM) is driven by symptoms, the [ECS](#) Tutor assembles lessons according to the symptoms for each malfunction (e.g., PACK OFF/INOP lights). As the Tutor presents each lesson, it provides the student with information similar to that shown in [Figure 2.6](#).

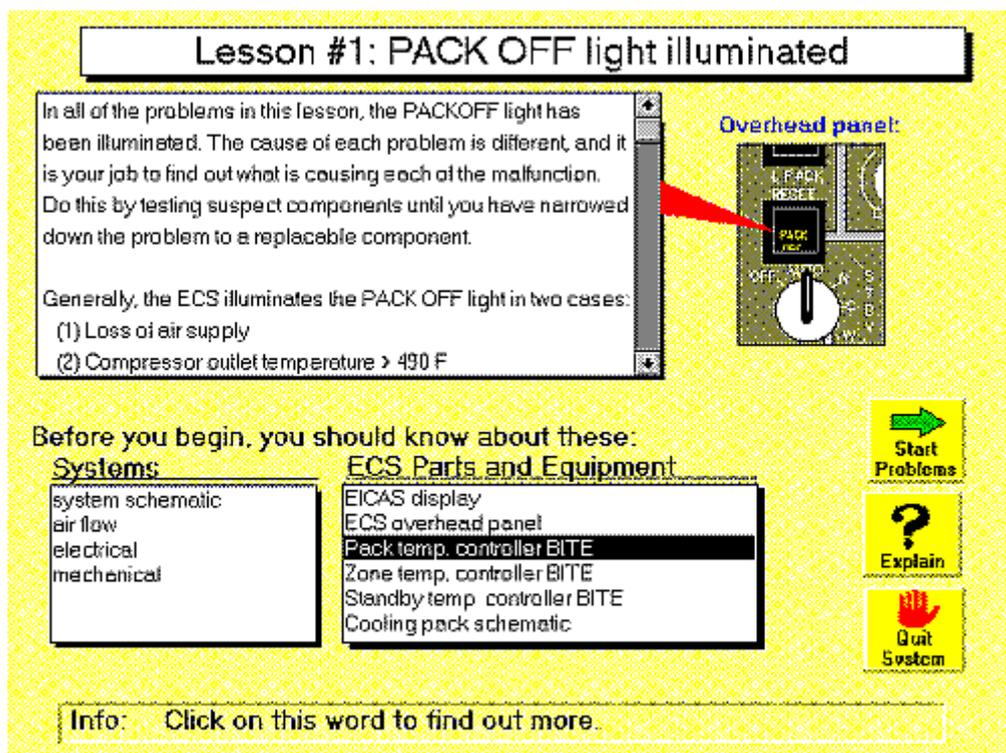


Figure 2.6 Lesson Overview Screen

The Background Information section describes the bases for the indicator lights. This section gives the common causes for the symptom. The Systems section gives a list of systems that relate to the indicator light. The user may choose one of these items to get more detailed information about a specific system. The Objects section provides a list of components that may be useful in solving the current malfunction. The user may choose one of these items to get more information about that component.

After completing a lesson, the Tutor reviews the student's performance from the previous lesson. If

the Tutor detects any deficiencies, it will provide material that will help the student overcome the deficiencies.

2.2.2.2 Over-the-Shoulder Advice

Another addition to the [ECS](#) Tutor makes the advice that the student receives "smarter". In the Prototype, the student received advice only on demand. In the Phase Two Tutor, the system monitors the student's progress toward a solution and offers help to the student if the student does not appear to be making satisfactory progress. The student may choose to see this "over-the-shoulder" advice or ignore it.

2.2.2.3 Simulation Design

The simulation provides realistic responses to a student's actions. For example, the [ECS](#) adjusts output temperature of the cooling pack when the student alters the desired zone's temperature. The new data values appear on the Overhead Panel and the [EICAS](#) display. The way in which the simulation accomplishes this temperature change should be transparent to the student.

In Phase One, the prototype "simulation" consisted of a table of data values for certain predefined conditions. The data shown to the student was updated according to this table. This approach was adequate for the prototype because of the limited scope of the prototype. However, due to the expanded number of potential malfunctions, the Phase Two operational tutor needed a more robust simulation of the [ECS](#) environment.

Phase Two's simulation contains a model of individual components. The simulation acts upon each component's inputs to produce that component's outputs. This approach allows a malfunction to propagate through the system. For example, if a valve fails, it will affect its outputs (such as air flow). The component that is downstream from the valve will receive unsatisfactory air flow and produce an erroneous output. The next component will receive this as its input, and the fault will continue to propagate.

The simulation will act upon inputs to produce outputs, whether those inputs are good or bad. This design allows for the same simulation to be used in both normal operation as well as during malfunctions.

2.2.2.4 Development Tools

The [ECS](#) Tutor uses a variety of software tools. The Tutor uses different tools to create the interface, the simulation, and the graphics, as described below.

The interface was developed using *Asymetrix Toolbook*. *Toolbook* supports quick and easy interface development in the Microsoft Windows environment.

The Prototype simulation also used *Toolbook*. However, as the simulation matured, it migrated from *Toolbook* to the "C" programming language. "C" provided greater speed and more flexibility than *Toolbook*.

The graphics used in the operational tutor combine many different graphics creation techniques. Some graphics were custom-designed using graphics packages, while others were scanned from existing training documentation. The Tutor also used photographs taken with a digital camera and stored as graphics files. As the training display hardware advances, digital photographs will become crisper and more realistic.

2.2.3 Lessons Learned from the Development of the Tutor

As the Tutor progressed from the Prototype to the complete, operational system, formative

evaluations highlighted several different areas that needed to be addressed: advice, instructional motivation, and student confidence with the interface. A description of each area follows.

2.2.3.1 Advice

When the Prototype gave advice to the student, the advice told the student what procedure to perform next. It followed the [FIM](#) exactly. However, it did not provide enough motivation as to why to perform a certain procedure. The current tutor enhances the advice to help answer the question "Why perform this procedure instead of a different procedure?" With this extra information, the student learns the logic behind the [FIM](#), instead of blindly following procedures.

2.2.3.2 Instructional Motivation

The Prototype emphasized troubleshooting, but lacked instructional focus. During this phase, the Tutor adds more emphasis to instructional issues. As described above in [Section 2.2.2.2](#), each lesson contains background information, system information, component information. Much of this information was available for the Prototype, but the student had to search for it. The Tutor now presents the information to the student in a more directed manner. This method reinforces to the student which pieces of information are important under different circumstances.

2.2.3.3 Student Confidence with the Interface

As students used the system, especially for the first time, they were hesitant to click on buttons with the mouse because they weren't sure what would happen. To allay some their indecision, the Tutor now contains the Message area. Anytime the cursor enters a selectable region, a short descriptive message appears in the Message area that describes the function of the selectable region. As training developers, we must be mindful to provide as much reassurance to the student as possible.

2.3 EVALUATION PLAN

If Advanced Technology Training is to become commonplace, it must be accepted by the user population. Also, if it is acceptable to the [AMT](#) it must also prove to be an effective training tool. This section outlines a plan to evaluate the [ECS](#) Tutor for user acceptance, cost effectiveness, training effectiveness.

2.3.1 User Acceptance Testing

User acceptance testing will be conducted in the following different user populations:

- Delta Air Lines instructors
- Delta Air Lines students
- Clayton State College Airframe & Powerplant (A&P) students

This testing will use questionnaires and interviews to gather data. The [ECS](#) Tutor will be modified as necessary to accommodate problems areas that the test identifies.

2.3.2 Training Effectiveness Study

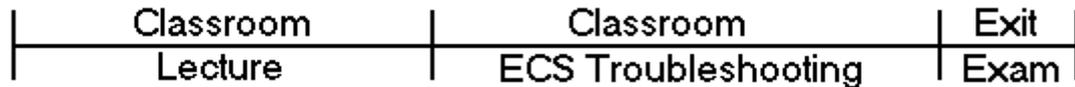
The training effectiveness study will test the following hypothesis:

A combination of conventional, classroom teaching and the [ECS](#) tutor is more effective than conventional, classroom teaching.

This experiment uses both a control group and an experimental group (from the Delta Air Lines

student population). Both groups will receive the normal 4 to 5 hour [ECS](#) classroom training segment. The control group will receive an additional classroom session on [ECS](#) troubleshooting. The experimental group will receive the lesson on [ECS](#) troubleshooting via the [ECS](#) Tutor. After the troubleshooting session, each group will take an exit exam which will measure their [ECS](#) proficiency. The scores on the exit exam will be used as the basis of comparison. [Figure 2.7](#) summarizes the experiment.

Control Group



Experimental Group



Figure 2.7 Evaluation Plan

2.3.3 Cost Effectiveness Study

A cost effectiveness study will measure the effectiveness of the [ECS](#) Tutor along several dimensions. Among these are both development costs and delivery costs.

2.4 PART 147 & PART 65 RULE CHANGES AND HUMAN FACTORS IMPLICATIONS

2.4.1 Part 147 - Aviation Maintenance Schools

From 1989 through 1991 there has been a rule change in the making for FAR Part 147 - Aviation Maintenance Technician Schools. After extensive public hearings, the rule changes are completed and will be published in 1992. The changes in the approved curricula were designed to be more responsive to the needs of today's aviation industry. For example, skill and knowledge requirements for gas turbine engines have been increased while requirements related to fabric covered aircraft have been lessened.

A second significant change in [FAA](#) philosophy regards educational media. The new rule permits the substitution of advanced technology training systems (like computer-based training, CBT) for classroom and laboratory instruction. This change in the rule is likely to foster an explosion of CBT for aviation maintenance training. The [FAA](#) should monitor the quality of such CBT to ensure that content and the delivery methods are appropriately applied.

The changes to Part 147 are not significant in number. However, the recognition of the potential for computer-based training is a very positive step in helping the schools to train for modern aircraft

systems.

2.4.2 Part 65 Certification: Airmen Other than Flight Crewmembers

The public hearings and Aviation Rulemaking Advisory Committees are currently working with the [FAA](#) to consider changes to Part 65. The changes that are being considered address such issues as experience requirements, training and proficiency requirements, and issues related to issuances of certification of personnel for repair facilities outside of the United States.

2.4.3 Job/Task Analysis

All who have been involved in Part 147 and Part 65 agree that there is a dire need for a new Job/Task Analysis (JTA) for the Aviation Maintenance Occupation. Such a [JTA](#) has not been done since 1969. The activity related to this rule change is prompting the appropriate funding for the [AMT](#) JTA.

2.5 SUMMARY

This research shows that Advanced Technology Training can be applied to the Aviation Maintenance community. This particular system concentrated on the [ECS](#) because of the generic nature of the [ECS](#) system (i.e., [ECS](#) principles are similar from aircraft to aircraft). However, this same approach is applicable in all other aviation maintenance areas. Up to this point, the Advanced Technology Training has concentrated on the mental aspects of training (logical troubleshooting). As this work continues, the research will attempt to combine the logical troubleshooting techniques with the psychomotor skills required to maintain the aircraft.

2.6 ACKNOWLEDGEMENTS

This research would not be possible without the cooperation of the aviation industry. Delta Air Lines and Clayton State College have enthusiastically offered their time and ideas to make the final product a success. The authors thank Rick Spavelko, Gary Drake, and Bruce Gindlesperger of Delta Air Lines and Jack Moore of Clayton State College for their cooperation and support.

2.7 REFERENCES

Code of Federal Regulations 65, 147

Polson, M.C., and Richardson, J.J. (Eds.) (1988). *Foundations of intelligent tutoring systems*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Shepherd, W.T., Johnson, W.B., Drury, C.G., Taylor, J.C., Berninger, D. (1991) [Human factors in aviation maintenance phase 1: progress report](#) (DOT/FAA/AM-91/16) Washington, D.C.: Office of Aviation Medicine.

Chapter Three

Emerging Technologies for Maintenance Job Aids

3.0 INTRODUCTION

Maintenance is fast becoming one of the most frequent application areas for job aiding. Maintenance job aids range from automatic preventive maintenance schedulers, to systems that monitor equipment status and recommend maintenance based on trends in equipment behavior, to systems that aid in fault diagnosis and repair. Application domains range from production equipment (e.g., clutch assembly machines), to process equipment (e.g., turbine generators), to high technology specialized equipment (e.g., fighter aircraft). There is a range of methodologies employed, as well, including algorithmic approaches for the preventive maintenance schedulers to expert systems for fault diagnosis and repair. The technologies employed encompass a range from VAX mini computers to desktop microcomputers linked to video disks. This chapter addresses extant approaches to job aiding in maintenance, the prospects for using emerging technologies for such systems, and the impact of emerging technologies on human performance, particularly in aviation maintenance applications. This section also calls for a new design philosophy in building job aids. A study which used this philosophy and compared three different levels of aiding on a task is also discussed. Some of the results of the study and their applicability to maintenance job aids are presented.

This chapter is similar to a previous review of job aids (see [Chapter 5 of Shepherd, et al., 1991](#)), in that many of the systems encountered were concerned with technological developments, rather than performance achievements. Whereas that previous work identified some of the difficulties with introducing advanced technology job aids into an operational environment, this discussion addresses some of the fundamental problems with past approaches to job aids and presents a design philosophy which capitalizes on the skills and abilities of the operator in order to produce a combined human-computer system that attains increased performance.

3.1 SURVEY OF MAINTENANCE JOB AIDS

A survey of academic, industrial, and popular literature revealed a wide variety of approaches to building maintenance job aids (see [Appendix](#)). These differing approaches include both hardware and methodological considerations, ranging from stand-alone, automatic scheduling systems to portable, interactive troubleshooting systems. The hardware aspects are addressed first, followed by a discussion of some of the different methods used.

3.1.1 Hardware Employed

The following systems exemplify different hardware approaches used for maintenance job aids. These systems are presented in order of increasing sophistication.

Folley and Hritz (1987) describe an expert system that assists in troubleshooting clutch assembly machines on a production line. Fault lamps above the machine stations indicate which stations are malfunctioning. A technician takes a maintenance cart to the malfunctioning station. The cart carries a two-button control and a monitor and the technician connects these to a junction box at the station. This junction box links the monitor and control to a remote computer and video disk player. The technician uses the control to move through a menu system to specify the faulty station. The computer then specifies the tests to be performed, along with graphic displays of the equipment, and the technician enters the results of the tests. In this way, the computer guides the technician through

troubleshooting and repairing the malfunctioning equipment.

A similar system developed by the Electric Power Research Institute (EPRI) also uses a video disk player for displaying maintenance information and procedures for gas-turbine power plants. This system uses a dual processor computer system. One processor manages an expert system, while another controls a video disk player. The [EPRI](#) system also uses voice recognition and synthesis for input and output, respectively.

General Motors developed an expert system to assist in vibration analysis of production machinery (cf. "GM unveils `Charley'..."). Named after a retiring technician with many years of experience, `Charley' was intended to help less experienced technicians locate parts that needed repair in production equipment with rotating components. Charley stores a signature file for each properly operating piece of equipment; technicians record the vibration signature of a problematic piece of equipment with a special data recorder and then connect the recorder to a Sun workstation. Charley compares the newly recorded signature with the database and begins diagnosing the problem. Charley guides interactions, may ask the technician for additional information, and explains its troubleshooting strategies. Charley can also be used as a consultant and allow a technician to explore `what if' questions. Finally, Charley is also used to train new technicians. The emphasis of the system is on preventive maintenance, rather than repair of failed equipment.

McDonnell Douglas developed the `Avionics Integrated Maintenance Expert System' (AIMES) for use on F/A-18 fighter aircraft (cf. "McDonnell Douglas flight tests..."). [AIMES](#) is a self-contained on-board box which contains a microprocessor and records flight avionics data on a cassette for later analysis. Production rules detect and isolate avionic failures at the electronic card level. [AIMES](#) generates queries and tests based on data and concludes whether a fault is present. If there is a fault, [AIMES](#) supplies the fault data, the card name, and the reasoning that led to the fault isolation conclusion.

The telecommunications industry is a large user of advanced technology maintenance aids, particularly in network switch and cable analysis (cf. "Expert system from AT&T..."). The `Automated Cable Expertise' system runs automatically each night to detect trouble spots in cables. Upon identifying a problem, it reports the repair history of the area and suggests corrective action.

3.1.2 Methods Employed

The following systems exemplify the range of methodologies employed in maintenance job aids. These systems are presented in order of increasing sophistication.

Berthouex, Lai, and Darjatmoko (1989) discuss a system for determining daily operations for a wastewater treatment plant. This system is billed as an `expert system', although it was developed using standard spreadsheet (Lotus 1-2-3) and database software (d-Base III), rather than one of the many production system shells. (Expert systems have historically been written using production rules (if-then clauses) in one of many languages specifically designed for that purpose, for example OPS5 or LISP. Popularization of the term `expert system' has led to decreasing precision of use of it.)

`Process Diagnosis System' (PDS) was developed by the Westinghouse Research and Development Center and Carnegie Mellon University for maintenance of steam generators. [PDS](#) is a condition monitoring system for preventive maintenance in order to alleviate both breakdown maintenance and unnecessary maintenance. The system is designed to detect deterioration early and predict the duration of safe operation. [PDS](#) also recommends specific preventive maintenance for regularly scheduled down times.

Vanpelt and Ashe (1989) describe the `Plant Radiological Status' (PRS) system for nuclear power plants. The [PRS](#) system presents a three dimensional model of the power station and equipment so that maintenance teams may plan maintenance tasks in advance. The [PRS](#) system facilitates access to and interpretation of radiological conditions by identifying hotspots and contaminated areas, as

well as identifying obstructions and available workspace. The goals of the [PRS](#) system are to reduce maintenance time and radiation exposure.

Several systems for supporting operations and maintenance were reviewed by Bretz (1990). One of the systems was developed by Chubu Electric Power Company and Mitsubishi Heavy Industries, Ltd. in Japan. This comprehensive expert system assists in power plant boiler failure analysis and maintenance planning. The failure diagnosis reports the most probable causes for failure, guidelines for inspection, the items to be investigated, repair methods, and suggested preventive maintenance. The maintenance planning subsystem automatically prepares daily repair schedules, a work estimation plan, and work specifications.

The distinction is sometimes made between 'deep' and 'shallow' knowledge in expert systems. The knowledge typically represented in production systems is considered shallow knowledge because it contains only antecedent-consequent relationships without any information as to why one thing follows from the other. Deep knowledge, on the other hand, captures the functional and causal relationships between the components of the object or system being model. Atwood, Brooks, and Radlinski (1986) call 'causal models,' which use components functions as the basis for their reasoning, the next generation of expert systems. Clancy (1987) describes a system for diagnosing switch mode power supplies which uses a model of the component level of the electronics for its diagnosis. Whereas one can test for signal presence at the module level of the electronics, the component level is concerned with the way in which a signal changes as it passes through the components. Finally, a system developed for Britain's Central Electricity Governing Board uses a model of the cause and effect relationships inherent in turbine generators for diagnosis and maintenance (see "Expert system probes..."). This expert system monitors and analyzes the vibration patterns of the equipment in its analysis.

The most sophisticated system encountered in the survey is the 'Testing Operations Provisioning Administration System' (TOPAS) developed by AT&T. Clancy (1987) describes [TOPAS](#) as a real-time, distributed, multi-tasking expert system for switched circuit maintenance. [TOPAS](#) performs trouble analysis, localization, and referral of network troubles. Clancy claims that [TOPAS](#) "does network maintenance without human intervention or consultation" (p. 103). If this is true, then [TOPAS](#) is not really a job aid, because it performs the job itself.

3.2 THE USE OF ARTIFICIAL INTELLIGENCE IN JOB AIDS

The methods and design philosophies used in building job performance aids vary with the designer (s). While some of the systems surveyed placed the technician in charge of the troubleshooting and maintenance, the majority of the approaches relied on artificial intelligence. The following describes various artificial intelligence approaches and their impact on human performance.

3.2.1 Expert Systems

Expert systems typically have three components: a rule base, a knowledge base, and an inference engine. The rule base contains the problem solving strategies of an expert in the domain for which the system was developed. The rule base is made up of production rules (if-then clauses). The knowledge base contains the history and the current data of the object under consideration (this object may be anything from an aircraft engine to a medical patient). The inference engine is responsible for determining what rules get activated and when the system has solved the problem or is at an impasse. Expert systems are typically written in a programming language specifically designed for such use, such as LISP or OPS5.

Typically, the human expert is not the person who builds the expert system, rather he/she interacts with a 'knowledge engineer' who is responsible for extracting the expert's expertise. One difficulty with expert systems has frequently been referred to as the 'knowledge engineering bottleneck'; it can be difficult to access and program the knowledge of the expert into the expert system. For instance,

the expert may not even be aware of what he/she does to solve a particular problem. Furthermore, it is impossible to guarantee that the rule base contains all of the knowledge of the expert.

3.2.2 Knowledge-Based Systems

Knowledge-based systems place less emphasis on production rules as a way of representing knowledge, and more emphasis on using a large database of information. This database may consist of information such as vibration patterns of equipment, as in Charley discussed above, or it may consist of typical hardware configurations, for instance. The point of knowledge-based systems is that they rely on a large body of readily-available information for the bulk of their processing.

3.2.3 Model-Based Systems

Model-based systems are an attempt to produce more robust problem solving systems by relying on 'deep' representations of a domain. The models depend on a description of the functionality and relationships of the components that make up the domain. Model-based systems are concerned with not only how a component functions, but why it functions that way. Developers of model-based systems believe that these systems will be able to solve novel problems, whereas expert systems can only solve problems with which an expert is familiar.

3.3 HUMAN PERFORMANCE IMPLICATIONS OF ARTIFICIAL INTELLIGENCE APPROACHES

The human performance implications of using an artificial intelligence-based problem solver are many. All of these systems revolve around the 'machine expert' paradigm, in which the computer controls all problem-solving activities. One problem with the machine expert paradigm is that because computers do not have access to the 'world', they must rely on a person to supply all relevant data about the world. Thus, the machine expert directs tests to be run and requests the results of those tests. Based on these data, the computer requests more information or reaches a conclusion, and that conclusion may be erroneous. In the words of one cognitive engineering researcher, the human is reduced to a "data gatherer and solution filter" for the machine.

One problem associated with this lack of environmental access is that the person may have knowledge that the computer does not. Since the computer directs the problem solving, it may never ask for information that may be critical to successfully solving the problem. Furthermore, there is usually no provision for the operator to volunteer such information. The person may even have different goals than the machine or may not know what the machine's goals are when it is attempting to solve a particular problem. Additional difficulties arise when the human operator accidentally enters the wrong data or when he/she misinterprets a request from the computer. Suchman (1987) discusses the problems of human machine communication at length.

Probably the biggest problem associated with expert systems is that they are brittle. As mentioned above, expert systems can only solve problems that the human expert has seen or remembers to discuss with the knowledge engineer. People (either experts or expert system designers) simply cannot anticipate all of the environmental variability encountered in the world. This leads to the tragic irony of such systems: expert systems are most needed when a problem is difficult, and that is precisely when the expert systems fail. The upshot is that the human operator is left to solve a difficult problem without the benefit of having developed expertise through solving other problems, because those were handled by the expert system!

All of these problems and more arose in a study by Roth, Bennett, and Woods (1987), in which the authors observed technicians using an expert system to troubleshoot an electro-mechanical device. One of the major findings of the study was that only those technicians who were actively involved in the problem solving process and performed activities beyond those requested by the expert system

were able to complete the tasks. The technicians who passively performed only those activities requested by the expert system were unable to reach solutions on any but the most trivial tasks.

The above should not be interpreted as a condemnation of all uses of artificial intelligence techniques, however. Indeed, artificial intelligence has greatly advanced our understanding of the capabilities, as well as the limitations, of computational tools. Prudent use of such techniques can greatly enhance the ability of a cognitive engineer to provide operators with powerful problem solving tools.

3.4 EMERGING TECHNOLOGIES

Continued advances in hardware and software technologies will further increase the cognitive engineer's design repertoire. Indeed, there are many emerging technologies that could be profitably used in maintenance job aids. Advances in computer hardware, display hardware, and object modeling all have great potential to improve job aiding capabilities. Each of these is discussed below.

3.4.1 Advances in Computer Hardware

As computer hardware has become smaller and more powerful, there has been a progression to smaller, more portable job aids. Whereas earlier job aids ran on minicomputers, then workstations and personal computers, newer job aids are being designed using laptops. There is no reason to believe that the laptop computer is the smallest, lightest computer that will be developed, however. Indeed, the NCR NotePad has recently been introduced. This computer is pen-based; that is, all input is performed via a pen stylus, rather than through a keyboard or mouse. The NotePad is light enough that it can be easily held in one hand, which greatly facilitates taking it to the maintenance site. The NotePad is relatively quick, it has reasonably large storage capacity, and it has limited handwriting recognition abilities.

An aviation industry working group is currently defining the standards for a 'Portable Maintenance Access Terminal' (PMAT) for use in commercial aviation. As currently conceived, the [PMAT](#) would connect to the 'Onboard Maintenance Systems' of current aircraft and would be used for troubleshooting. Because the emphasis is on portability, it is likely that something similar to the NotePad or a standard laptop computer will be specified.

Another emerging hardware technology is the use of 'built-in test equipment' (BITE) in engineered systems, no doubt due in part to the widespread use of microprocessors. [BITE](#) likely does not eliminate the maintenance technician, however, because it may be difficult to implement such equipment in mechanical systems or in very complex systems. Indeed, [BITE](#) may introduce additional problems for maintenance people because there is a lack of standardization on how [BITE](#) should operate; thus, there may be confusion when dealing with similar, but different, [BITE](#). Further complications may arise due to issues of granularity in [BITE](#); [BITE](#) may simply indicate that a piece of equipment is not functioning properly, without indicating the specific nature of the malfunction or without indicating which component must be repaired or replaced. Another issue is: What happens when the [BITE](#) malfunctions?

3.4.2 Advances in Display Hardware

One of the surveyed systems used a personal computer to control a slide projector for displaying maintenance graphics. Several of the systems used a computer-controlled video disk for such displays. With the advent of digital cameras and compact disc-interactive (CDI) technology, systems with higher fidelity and portability can be achieved. Appropriately designed [CDI](#) systems could store many views of the object(s) being serviced, as well as maintenance procedures and information. Indeed, what graphics were displayed would depend on the fault manifestations.

Furthermore, well-designed [CDI](#) systems would allow the technician to troubleshoot by hypothesizing a failed component and watching how a simulation of the system performed. Similarly, the technician could replace a component in the simulation and see the results. In this manner, the technician could develop expertise more quickly than learning on-the-job (because the technician would have control over what aspects he was learning, rather than relying on whatever malfunction happened to occur).

3.4.3 Advances in Object Modeling

An extension of the three-dimensional model discussed above is virtual reality. Virtual reality has received a lot of attention as a result of the Defense Advanced Research Project Agency's development of the 'Pilot's Associate Program' and consists of replacing an operator's view of the 'real world' with a simulated view of that world. Thus, real world objects are replaced with simulations of those objects. One possible use of virtual reality would be to allow the maintenance technician to 'stand' inside a device, such as an engine, and watch how it functions, both normally and with failed components. The technician could also see the effects of replacing components, similar to the [CDI](#) system above, but with the benefit of observing the effects more directly. As with [CDI](#), the technician need not replace the actual system components, but may replace components in the simulation of that system. The uses of virtual reality appear to be limited only by the job aid designer's imagination.

3.5 HUMAN PERFORMANCE IMPLICATIONS OF EMERGING TECHNOLOGIES

While many past approaches to job performance aids sought to replace human expertise with machine expertise, there is a growing appreciation for the importance of human skill. The machine expert paradigm sought to overcome human information processing 'limitations' with a computer prosthesis. However, even computers are limited resource processors. A more enlightened approach is to view computers as tools to amplify human capabilities, not overcome limitations. In this sense, computers can be seen to be like other tools, such as telescopes or automobiles: they are instruments which provide additional resources for achieving our needs and desires. Woods and Roth (1988) discussed the above issues and addressed many more cognitive engineering issues inherent to developing systems that have powerful computational abilities.

Technology is not a panacea; each new technology brings with it significant drawbacks, as well as benefits. The challenge to designers is to use emerging technologies to build cooperative systems, in which both the human and the computer are actively involved in the problem solving process. Humans can no longer be regarded as passive 'users' of technology, but as competent domain practitioners with knowledge and abilities which are difficult to replace. The following section discusses a study which addressed just such issues.

3.6 A STUDY OF HUMAN PERFORMANCE WITH A COOPERATIVE SYSTEM

A study which addressed some of the human performance issues discussed above was carried out as part of the author's graduate program (Layton, 1992). This study compared three different levels of computer support on the basis of their effects on human performance. Although the domain for which the systems were developed was enroute flight planning, the general principles behind the alternative designs can be applied to developing aviation maintenance aids, as well. The following is a discussion of enroute flight planning, the design concepts behind the three levels of computer support, the method employed for comparing the various systems, the general outcomes of the study, and the implications of those outcomes for developing aircraft maintenance job aids.

3.6.1 Enroute Flight Planning

Enroute flight planning consists of modifying the flight plan of an airborne aircraft in response to changes in the capabilities of the aircraft, to crew or passenger emergencies, to changes in weather conditions, and/or to problems at the destination airport. The study focused on flight plan adaptation in response to changes in weather conditions. From a pilot's perspective, the components important to enroute flight planning include the airplane, possible flight routes, weather conditions, and airline company dispatchers. The pilot is concerned with getting from a given origin to a given destination on time, with a minimum of fuel consumed, while maintaining flight safety. He/she must consider what routes to take (these routes consist of waypoints, or navigational points, and jet routes, the so-called "highways in the sky"), what altitudes to fly, what weather to avoid, and the ever-changing capabilities of the aircraft (e.g., the weight of the plane decreases with fuel consumption; the lighter the plane, the higher it can fly, within limits).

The initial flight plan is rarely followed exactly, due to unforeseen events occurring while enroute. Indeed, minor changes in flight plans are frequently made and major changes are fairly common. These amendments to the original result from the dynamic, unpredictable nature of the 'world' in which the plans are carried out. Weather patterns do not always develop as predicted, resulting in unexpected areas of turbulence, less favorable winds, or storms that must be avoided. Air traffic congestion may delay take-off or restrict the plane to lower-than-planned altitudes. Airport or runway closures can cause major disruptions, not just for one aircraft, but for everyone planning on landing at that airport. Mechanical failures, medical emergencies, or other critical problems may delay take-off or may force an airborne plane to divert to a nearby airport.

Furthermore, there are several constraints on the flight plans that can be developed. Planes must maintain a certain separation distance between each other and between thunderstorm cells, as specified in the Federal Air Regulations. Planes must fly along the jet routes. They are also limited to certain altitudes. Over the continental United States, for example, 33,000 feet is an 'eastbound only' altitude. There are also physical limitations: the plane can't fly if it is out of fuel and it can't land at an airport with runways that are too short. Some of these constraints are actually 'soft', in that they may be violated in some circumstances. If, for instance, there is no eastbound traffic, Air Traffic Control (ATC) may allow a plane to fly west at an 'eastbound only' altitude. Similarly, [ATC](#) may approve a vector that deviates from the jet routes in order to avoid a storm or to save fuel.

3.6.2 System Design Concepts

It is clear that enroute flight planning is a complex activity, but it is not clear how humans deal with these complexities or how one might program a computer to choose the 'optimum' solution to any given problem. For instance, how does one make tradeoffs between fuel conservation, flight safety, and prompt arrival at the destination? Because pilots make such tradeoffs on a routine basis, one goal of the study was to develop a system to support the pilots in making such decisions. There is a heavy emphasis, therefore, on allowing the pilots to explore "what if" types of questions so that they could gain feedback on the impact of a planning decision on flight parameters.

The three levels of computer support corresponded to successively greater flight planning power. Common to all three systems were: 1. a map display which consisted of the continental United States, the aircraft, and flight routes; 2. a representation of a flight log, which included the flight route and altitudes; and, 3. a display of flight parameters. These three items were displayed on two monitors. [Figure 3.1](#) depicts the map displays and controls, and [Figure 3.2](#) depicts the flight log display and controls and the flight parameter display. The pilot could elect to display weather data, waypoints, and jet routes on the map display. The lowest level of enroute flight planning support provided the pilot with the ability to sketch proposed flight plans on the map, in accordance with the waypoint and jet route structure. The latter condition required a pilot to sketch routes one waypoint at a time. Once the pilot completed a proposed flight plan, in terms of geographic location, the computer responded with various flight parameters, such as time of arrival and fuel remaining at the destination. The computer also indicated whether the flight was predicted to encounter any

turbulence and the severity of that turbulence. The computer also proposed the most fuel efficient vertical flight profile for the proposed route. This form of support encouraged the pilots to propose options and see their effects on flight parameters. This form of support is referred to as the 'sketching only' system.

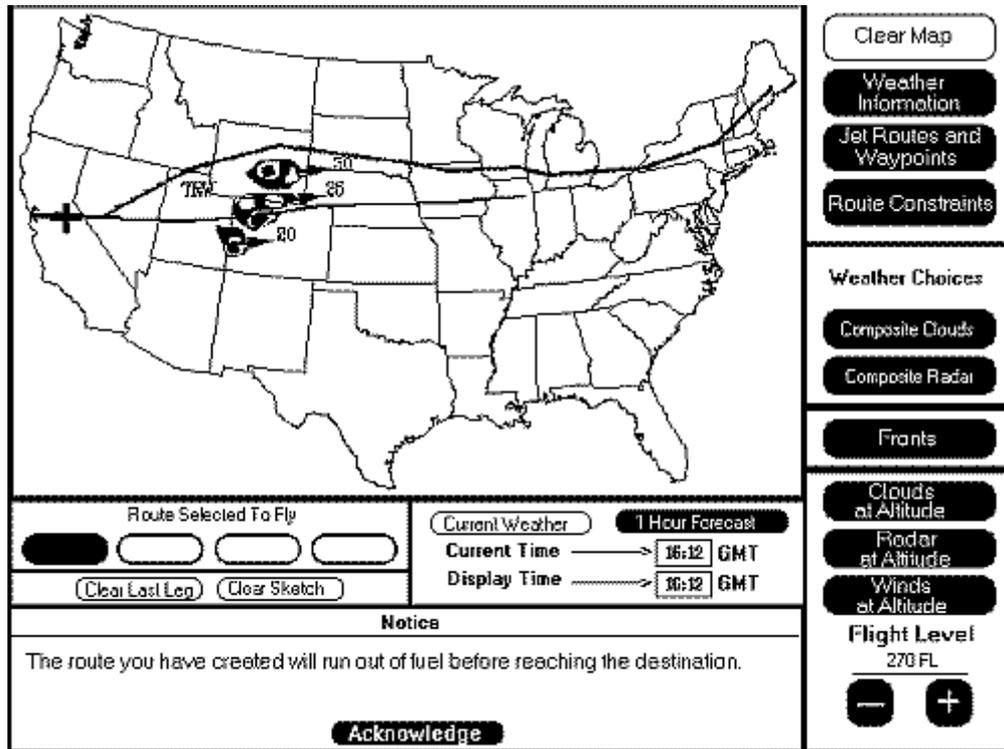


Figure 3.1 Left Monitor Displays and Controls

		DTA	EKR	SNY	OBH	
Save	Clear	Route:	DTA	EKR	SNY	OBH
Copy		Altitude:	FL 330	FL 330	FL 330	FL 330
Calculate		Speed:	Mach .72	Mach .72	Mach .72	Mach .72
Time of Arrival (G.M.T.):		1709	1739	1812	1842	
Fuel Remaining (1000 lbs):		33	26	22	18	
Distance (miles):		216	233	209		
Next Info						
FL 330		moder	moder			
Turbulence						
FL 290		moder	moder			
Wind Components						
FL 270		moder	moder			
Wind Direction/Speed						
FL 250		moder	moder			
FL 230		moder	moder			
Least Fuel Altitude						
Planned Altitude		GRND				
Time of Arrival: 19:45 GMT		Time of Arrival: 21:56 GMT	Time of Arrival: 21:56 GMT	Time of Arrival: 21:56 GMT		
Time Enroute: 3:45		Time Enroute: 5:58	Time Enroute: 5:58	Time Enroute: 5:58		
Fuel Remaining: 11047 lbs		Fuel Remaining: -6225 lbs	Fuel Remaining: -6225 lbs	Fuel Remaining: -6225 lbs		
Total Distance: 1573 nm		Total Distance: 2511 nm	Total Distance: 2511 nm	Total Distance: 2511 nm		

Figure 3.2 Right Monitor Displays and Controls

The next level of computer support incorporated the sketching form of interaction, but also included a method for placing constraints on a desired solution and allowing the computer to propose a solution which satisfied those constraints. For instance, the pilot could place limits on the maximum severity of turbulence and precipitation encountered, and could specify the desired destination. The computer would then perform a search of the data and solution spaces and propose a route that satisfied the pilot's constraints while minimizing fuel consumption. This proposed route would include both the geographic route and the vertical profile, along with its associated flight parameters. This form of flight planning causes the pilot to plan at a more abstract level than the sketching form of interaction, because the pilot is able to think about the characteristics of a desired solution while the computer handles the lower level details of specific routings. Using the sketching tool, the pilot was free to modify the route proposed by the computer and note the impact of such changes on the flight parameters. This second level of planning can be roughly construed to be a form of consultation system because the computer can be asked for its advice on a problem; it is referred to as the 'route constraints and sketching' system.

The highest level of support corresponds to an expert system that automatically solves a problem as soon as it is detected; upon loading the scenario information, the computer would propose a solution which minimized fuel consumption and satisfied the constraints of encountering no turbulence and no precipitation, as well as arriving at the planned destination. As in the previous level of support, the computer would propose both the geographic route and altitude profile, along with the corresponding flight parameters. If desired, the pilot could also request a solution from the computer based on different constraints, and he could sketch his own solutions.

3.6.3 Study Method

Thirty male commercial airline pilots were randomly assigned to one of three treatment conditions, wherein each condition consisted of one of the three forms of computer support described above. There were ten subjects in each condition. Each pilot was trained for approximately one hour on his system prior to solving four enroute flight planning cases. Each case consisted of a planned flight that was disrupted because of a change in weather conditions. The task for the pilot was to decide what to do in each situation. All of the pilots solved the four cases in the same order. It took approximately an hour and a half to solve the four cases.

3.6.4 Study Results

Each of the four cases provided some interesting insights into the influences of computer tools on human behaviors. The overriding results of each of the four cases are discussed below.

3.6.4.1 Case 1 General Results

In the first case, most of the subjects in the 'route constraints and sketching' and the 'automatic route constraints, route constraints, and sketching' conditions chose to fly the computer-suggested route (as expected). However, the 'sketching only' subjects tended to choose routes that were more robust; that is, these subjects put more distance between the aircraft and the storm. These subjects commented that they would like to have more distance from the storm than afforded by a more direct route (such as the one suggested by the computer in the other two treatment conditions). Furthermore, the 'sketching only' subjects were more apt to explore multiple routes and multiple types of routes, than were the subjects in the other two groups. These results suggest that the sketching form of interaction caused the subjects to consider the data more carefully than did the route constraints tool. One reason for this result is that the sketching tool gave the subjects the opportunity to consider the relationships of various route options and the weather at several points and to consider the robustness of those options given the uncertainties associated with weather. The constraints tool, on the other hand, did not encourage such behavior, and, indeed, the subjects using that tool may have been under the impression that the computer was considering the robustness of

routes, when in fact it was not. If the sketching tool encouraged more careful examination of the data than did the constraints tool, and this behavior persisted, one could imagine situations wherein the constraints tool could lead to bad decisions.

3.6.4.2 Case 2 General Results

While Case 1 provided evidence for the benefits of tools that make the operator the sole decision maker, Case 2 provided evidence to the contrary. In Case 2, the 'sketching only' subjects had significant difficulty, as a group, in searching the relatively large data and solution spaces. Many of the routes explored by these subjects passed through strong turbulence. Indeed, four of these ten subjects chose deviations that exacted a high fuel consumption cost, either because they could not find a more efficient route around/through the weather or because they did not examine wind data which would have indicated that their chosen route encountered strong head winds. By contrast, the subjects in the 'route constraints and sketching' and 'automatic route constraints, route constraints, and sketching' groups successfully used the computer to rapidly find a fuel efficient deviation that avoided all of the weather. Furthermore, nearly all of the subjects who chose an inefficient deviation later stated that they preferred the more efficient deviation suggested by the computer to the other groups.

3.6.4.3 Case 3 General Results

As noted in the discussion of Case 1, the 'sketching only' subjects chose rather different solutions than did the 'route constraints and sketching' and the 'automatic route constraints, etc.' subjects. Furthermore, it was hypothesized that the 'sketching only' subjects were more involved in the problem solving process than were the subjects in the other two groups. The third case was designed to address the issues related to what happens when the automatic tools suggest questionable solutions: Does the operator recognize that the solution may not be appropriate? Assuming the operator does recognize that the solution is inappropriate, can he readily come up with a better solution?

In Case 3, the computer suggested two different routes in the 'route constraints and sketching' and 'automatic route constraints, etc.' conditions, depending upon the constraints placed on it. One deviation passed between two large thunderstorm cells of a volatile storm, which is a risky practice, at best; this route was suggested on the basis of no turbulence and no precipitation. The other route avoided the bulk of the weather, at the cost of slightly higher fuel consumption and a small amount of turbulence; this route was suggested on the basis of light chop (or greater) turbulence and light (or heavier) precipitation. The trend in this case was for the 'route constraints and sketching' and the 'automatic route constraints, route constraints, and sketching' subjects to choose the first route more frequently than the 'sketching only' subjects. If these subjects had not examined both routes, then it would suggest that these subjects were simply over-reliant on the computer. However, several of the subjects in the 'route constraints and sketching' and 'automatic route constraints, etc.' groups examined both routes before choosing the more risky route; thus, these subjects chose a risky route despite evidence that it may have been a poor choice and that a better option existed. These subjects nearly unanimously changed their minds when later questioned about their decisions.

With few exceptions, the 'sketching only' subjects planned very conservative deviations that completely avoided the weather. However, the 'sketching only' subjects had considerable difficulty in finding acceptable deviations. In fact, one subject chose a deviation that was predicted to cut into his required landing fuel reserves prior to arrival at the destination. Thus, even though the 'sketching only' subjects may have considered the data very carefully, the problem was sufficiently complex that they would have benefitted from some computer assistance.

3.6.4.4 Case 4 General Results

Case 4 provided some interesting results with regard to individual differences and with regard to the

influence of computer recommendations. The 'sketching only' and 'route constraints and sketching' subjects were nearly evenly divided between a fuel efficient deviation and a robust deviation. When asked about his decision, one of the 'sketching only' subjects made the comment that the decision depended on the person's role in flying the aircraft at the time: if the captain were flying that leg, he would go one way so that he could look at the storm, but if the first officer were flying that leg, he'd go the other way around so that he could see the storm. Obviously this is an extreme example, but it underscores the role of individual differences in decision making.

Unlike the subjects in the other two groups, the 'automatic route constraints, route constraints, and sketching' subjects, were more likely to choose the computer-suggested, economical route, even when they had explored both routes. Combined with the results of Case 3, this result suggests that the computer exerts a strong influence on decision making when it recommends a solution at the onset of a problem.

3.6.5 Study Conclusions

The goal of the research was not to determine which particular version of an enroute flight planning tool resulted in the best human performance. Rather, one goal was to see how human behaviors were influenced by the tools available. Subjects who had multiple tools available to them (the 'route constraints and sketching' subjects and the 'automatic route constraints, route constraints, and sketching' subjects) were able to use them to develop alternative plans. In fact, there were many instances in which the solution recommended by the computer did not meet the needs of the pilots, so the pilots developed their own plans through sketching. Thus, not only is there a need for tools that allow the operator to go beyond a computer's solution, but there is a need to support individual differences, as well.

The subjects who had only the sketching tool available to them closely examined the available data. As a result, these subjects often planned robust deviations that would not need to be altered if there were further changes in the weather. Where these subjects ran into difficulties, however, was in situations in which there were a lot of potential solutions and there was a large amount of data. In such situations, these subjects had trouble finding appropriate solutions. Indeed, some of these subjects made poor decisions because of these difficulties. The subjects who had some form of computer assistance were able to more efficiently search these spaces, but with some costs.

The tool that automatically suggested a solution to the problem as soon as it was detected did not encourage the subjects to closely examine the data. While this fact did not cause problems in some cases, it clearly did lead to bad decisions in others. Furthermore, the automatic tool's influence on decision making went beyond simple over-reliance to the point where it shifted attention from data which were important to making a good decision.

3.6.6 Implications for Maintenance Job Aids

The conclusions outlined above can be readily applied to developing maintenance job aids. For instance, one of the conclusions is that there is a need for tools that allow an operator to go beyond a computer's solution. As discussed above, particularly with regard to Case 3, and as discussed by Roth, Bennett, and Woods (1987) and Suchman (1987), operators frequently have knowledge or information which is not available to the computer, but which is critical to making a good decision. By giving the authority and responsibility for decision making to the operator, and by providing a tool which supports the operators activities (rather than the other way around), the operator is free to explore solutions that may not have been designed into a machine expert.

Another conclusion reached by the above study was that the form of tool that required a person to make a series of decisions (the sketching tool) encouraged the operator to think hard about the problem and to consider the available data at a deeper level, than did the form of tool that encouraged the operator to make a single 'yes' or 'no' decision (the automatic route constraints tool). In this regard, the conclusion supports the notion that designers need to "keep the person in the

loop".

However, another conclusion of the above study was that "keeping the person in the loop" did not provide adequate support in some situations. Indeed, in some of the cases (such as Cases 2 and 3) some of the operators were simply unable to find adequate solutions on their own. These operators could have used some help from a computer in exploring solution possibilities. In such situations this is rarely a reflection of human 'limitations', rather it is an indication of the difficulty of the problem. In maintenance, for instance, diagnosing multiple, interacting faults is a difficult problem. One symptom may be characteristic of several faults, or one fault may mask the presence of another. A tool which helps to focus the diagnostician's attention and eliminate false leads would be very beneficial.

Finally, it is important to realize that each person has a different style of decision making; two people who complete the same training course on a given method for dealing with a problem may use slightly different approaches. Such differences are likely to increase with experience as each person learns methods that consistently work for him/her. Indeed, experts often use several different approaches to solving truly difficult problems because each approach has unique limitations as well as unique benefits. For instance, knowledge of thermodynamics may help localize a fault to a heat exchanger, but knowledge of circuits may lead one to test the power supply to the heat exchanger, as well. Thus, tools need to be flexible to support such individual differences, rather than use a single, lockstep approach, as in the case of 'expert' systems. (Note that although some expert systems do incorporate the observable components of such methods, they do not allow the operator direct access to those methods. Because the knowledge and capabilities of such systems are necessarily incomplete, the systems are 'brittle' in the face of difficult problems, as discussed above in [Section 3.4](#))

3.7 RESEARCH AND DEVELOPMENT PLAN

The above discussion points to the challenge for cognitive engineers involved in designing maintenance job aids: build systems that capitalize on both human strengths and computer strengths so that task performance is improved. As outlined above, because of relatively recent advances in hardware and software, it is possible to use sophisticated computational techniques (e.g., cooperative system techniques) to develop real time, computer-based job aids for a wide range of technical tasks. Furthermore, hardware like the NCR NotePad will make it easier for people without previous computer training to use such job aids. We are working with the Flight Standards Service of the [FAA](#) to develop a Portable Performance Support System (PPSS) to aid Aviation Safety Inspectors in their daily activities. The initial focus of this effort is on the tasks performed by Airworthiness (maintenance) Inspectors, particularly the Ramp Inspections task. Inspectors need access to many of the same types of information that maintenance technicians use. Inspectors must also document their activities and the outcomes of those activities. We are taking a three Phase approach to developing a [PPSS](#): Phase I, already underway, will identify a prospective task and perform an information needs analysis for that task; during Phase II we will design and develop a prototype [PPSS](#) for the task; and we will create a plan for the development of the prototype into a fully functional system during Phase III. Each of these phases is described more fully below.

3.7.1 Phase I: Problem Definition and Information Needs Analysis for Aviation Safety Inspectors

Goal: Identify a typical task for which a computer-based job aid is an appropriate application and conduct appropriate information needs analysis to define the work environment and information needs for aviation safety personnel.

We are working with the personnel of the Fort Lauderdale Flight Standards District Office (FSDO) to help identify an appropriate task for computer-based job aiding. Such a task should be one which is typical for the personnel, but which may require some experience to attain proficiency. This Phase

of the research is an ongoing process continued throughout the life of the project.

The task initially proposed to the Fort Lauderdale [FSDO](#) is one of an Airworthiness Inspector performing Ramp Inspections. Ramp Inspections are used to verify aircraft airworthiness just prior to a planned flight. The inspectors walk around the aircraft, identifying problem areas and documenting those problems on a Program Tracking and Reporting Subsystem (PTRS) form (if an aircraft meets safety standards, that information is noted on the same form). The PTRS form is used to document all activities in which inspectors are involved; such activities include accident/incident investigation, airman certification, flight school certification, etc. Inspectors also use paper-based "job aids", which are essentially checklists, to assist them in their activities. Filling out forms and following checklists are the types of activities for which the NotePad was designed. Therefore, such tasks are amenable to transfer to pen-computer technology. Furthermore, the computer allows multiple forms to be linked together such that entries in one form are automatically propagated to all related forms; this approach would eliminate the duplicate entry of data which currently occurs. Finally, PTRS forms are currently recorded in paper format and given to data entry clerks who must interpret the inspector's handwriting and transfer the data to the [FSDO's](#) local computer-based database (which feeds into the national PTRS database). The PTRS data collected on a [PPSS](#) will be in a format that can be directly transferred into a [FSDO's](#) local PTRS database, thus eliminating the intermediate manual data entry step.

Inspectors must also have access to large amounts of information, such as Federal Aviation Regulations, Inspector's Handbooks, Airworthiness Directives, Advisory Circulars, etc. Whereas inspectors must currently retain hard copies of such information or refer to the [FAA](#) mainframe repository, it can become cumbersome to access and track this information. This suggests that a hypermedia on-line documentation system can be beneficial to the inspectors. This system can run off either the NotePad's internal hard drive or an external [CD ROM](#) device. Such an on-line documentation will facilitate rapid access to up to date information.

Based on initial conversations with the Flight Standards Service, it appears that a [PPSS](#) will:

- Provide inspectors with an integrated, linked form system
- Provide a means to reduce data entry performed by clerks
- Provide on-line documentation, including FARs, handbooks, etc.

3.7.2 Phase II: Design and Development of a Prototype Portable Performance Support System

Goal: Build and demonstrate a prototype Portable Performance Support System to support Aviation Safety Inspectors doing Ramp Inspections.

This phase will involve several iterations of development and demonstration. Rapid prototyping of the [PPSS](#) will permit us to demonstrate the system to inspectors for quick feedback about the design and content of the system. We will work closely with the inspectors during this phase to ensure the accuracy of the information and the usability of the [PPSS](#). While these evaluations will primarily involve the airworthiness inspectors, it is extremely helpful to have other inspectors evaluate the [PPSS](#), because they can provide a fresh perspective and would be likely to identify additional areas for improvement.

3.7.3 Phase III: Create a Plan for the Development of the Prototype System into a Fully Operational System

Goal: Create a plan to convert the prototype [PPSS](#) into a fully operational system for evaluation and integration into the work environment.

This phase will require a formal review of the prototype [PPSS](#) to identify its strengths and weakness. Following this review, a plan will be developed to fully implement the [PPSS](#). The plan will include the design of a study to evaluate the effectiveness of the [PPSS](#) in the work environment.

3.8 SUMMARY

Several past approaches to maintenance job aiding were discussed with respect to their impact on human performance. Such approaches have typically used a 'machine expert' to guide technicians through the maintenance process. However, the 'machine expert' paradigm has met with limited success in operational environments because of problems with unanticipated variability in the environment (or 'brittleness'), extra-machine knowledge, and inflexibility. An alternative philosophy to developing systems was presented, cooperative systems, in which both the human and the computer are actively involved in the problem solving process. This philosophy advocates a change in perspective toward computers as tools to assist people in their work, rather than as prostheses to overcome human 'limitations'. The cooperative problem solving paradigm capitalizes on the strengths of humans and computers in order to improve the performance of both. A study which compared different versions of a job aiding system designed with using this philosophy was presented, along with implications for developing maintenance job aids. Finally, a plan for developing a portable performance support system for aviation safety inspectors was presented.

3.9 REFERENCES

- Aerospace maintenance. (1986, December). *Aerospace America*, p. 46.
- Ahrens, R. B., Marsh, A., & Shannon, P. A. (1984, November). 3B20D computer: Maintenance with a mind of its own. *Record*, pp. 16-19.
- AI to help keep aircraft flying. (1986, June 12). *Machine Design*, p. 4.
- Armor, A. F. (1989, July). Expert systems for power plants: The floodgates are opening. *Power Engineering*, pp. 29-33.
- Artificial intelligence to aid in war on potholes. (1985, December 12). *Engineering News-Record*, p. 215.
- Atwood, M. E., Brooks, R., & Radlinski, E. R. (1986). Causal models: The next generation of expert systems. *Electrical Communication*, 60(2), 180-184.
- Barney, C. (1985, December 23). Expert system makes it easy to fix instruments. *Electronics*, p. 26.
- Benedict, P., Tesser, H., & O'Mara, T. (1990, June). Software diagnoses remote computers automatically. *Automation*, pp. 46-47.
- Bertheouex, P. M., Lai, W., & Darjatmoko, A. (1989). Statistics-based approach to wastewater treatment plant operations. *Journal of Environmental Engineering*, 115, 650-671.
- Bogard, W. T., Palusamy, S. S., & Ciaramitaro, W. (1988, May). Apply automation to diagnostics, predictive maintenance in plants. *Power*, pp. 27-32.
- Bretz, E. A. (1990, July). Expert systems enhance decision-making abilities of O&M personnel. *Electrical World*, pp. 39-48.
- Byrd, T. A., Markland, R. E., & Karwan, K. R. (1991, July-August). Keeping the helicopters flying--using a knowledge-based tank support system to manage maintenance. *Interfaces*, pp. 53-62.
- Callahan, P. H. (1988, January-February). Expert systems for AT&T switched network maintenance. *AT&T Technical Journal*, pp. 93-103.
- Clancy, C. (1987, November). Qualitative reasoning in electronic fault diagnosis. *Electrical Engineering*, pp. 141-145.
- Computer oversees maintenance. (1992). *American Water Works Association Journal*, 84, 107-

108.

Cue, R. W. & Muir, D. E. (1991). Engine performance monitoring and troubleshooting techniques for the CF-18 aircraft. *Journal of Engineering for Gas Turbines and Power*, 113, 11-19.

Culp, C. H. (1989). Expert systems in preventive maintenance and diagnostics. *ASHRAE Journal*, 31, 24-27.

Dallimonti, R. (1987, June 18). Smarter maintenance with expert systems. *Plant Engineering*, pp. 51-56.

de Kleer, J. (1990). Using crude probability estimates to guide diagnosis. *Artificial Intelligence*, 45, 381-391.

de Kleer, J. & Williams, B. C. (1987). Diagnosing multiple faults. *Artificial Intelligence*, 32, 97-130.

Dobson, R., & Wild, W. (1989, May). Plant's computerized maintenance system improves operations. *Power Engineering*, pp. 30-32.

Dohner, C. V., & Acierno, S. J. (1989, August). Expert systems for gas-turbine powerplants passes first tests. *Power*, pp. 63-64.

Doorley, R. (1988, August). Hydraulic troubleshooting using an expert system. *Hydraulics & Pneumatics*, pp. 91-92.

Doorley, R. B. (1989, June 22). Expert systems probe hydraulic faults. *Machine Design*, pp. 89-92.

Expert system from AT&T Bell Laboratories is an 'ACE' at telephone cable analysis. (1983, October). *Record*, p. 1.

Expert system guides tube-failure investigations. (1989, August). *Power*, p. 85.

Expert system probes beneath the surface. (1990, January). *Mechanical Engineering*, p. 112.

Expert systems to hone jet engine maintenance. (1986, April 21). *Design News*, pp. 36-38.

FAA and NASA design program to improve human performance. (1989, May 29). *Aviation Week & Space Technology*, p. 115.

Foley, W. L., & Svinos, J. G. (1989). Expert advisor program for rod pumping. *Journal of Petroleum Technology*, 41, 394-400.

Folley, J. D., & Hritz R. J. (1987, April). Embedded AI expert system troubleshoots automated assembly. *Industrial Engineering*, pp. 32-35.

GM unveils "Charley", an expert machine diagnostic system. (1988, May). *I&CS*, pp. 4-7.

Gunhold, R., & Zettel, J. (1986). System 12 in-factory testing. *Electrical Communication*, 60(2), 128-134.

Hartenstein, A. (1988, January). Computer system controls all maintenance activities. *Public Works*, p. 60.

Hill, S. (1990, February). Ask the expert. *Water & Pollution Control*, pp. 12-13.

Hughes, D. (1988, March 7). Digital develops special applications to meet diverse aerospace needs. *Aviation Week & Space Technology*, pp. 51-53.

Jet fighter uses AI as troubleshooter. (1986, July 21). *Design News*, p. 20.

Keller, B. C. & Knutilla, T. R. (1990, September). U.S. Army builds an AI diagnostic expert system, by soldiers for soldiers. *Industrial Engineering*, pp. 38-41.

King, I. J., Chianese, R. B., & Chow, M. P. (1988, December). Plant diagnostics relies on AI

transmissions from remote site. *Power*, pp. 57-60.

Kinnucan, P. (1985, November). A maintenance expert that never sleeps. *High Technology*, pp. 48-9.

Kolcum, E. H. (1989, January 2). Growing flight, maintenance simulator market attracts many competitors. *Aviation Week & Space Technology*, pp. 91-93.

Layton, C. F. (1992). An investigation of the effects of cognitive tools on human adaptive planning in the domain of enroute flight planning. Doctoral dissertation, The Ohio State University, Columbus, OH.

Maintenance expert in a briefcase. (1986, April). *High Technology*, p. 9.

Majstorovic, V. D. (1990, October). Expert systems for diagnosis and maintenance: State of the art. *Computers in Industry*, p. 43-68.

McDonnell Douglas flight tests AI maintenance data processor. (1986, February 17). *Aviation Week & Space Technology*, p. 69.

McDowell, J. K., & Davis, J. F. (1991). Managing qualitative simulation in knowledge-based chemical diagnosis. *AIChE Journal*, 37, 569-580.

Melhem, H. G., & Wentworth, J. A. (1990, March). FASTBRID: An expert system for bridge fatigue. *Public Roads*, pp. 109-117.

Miller, D. M., Mellichamp, J. M., & Wang, J. (1990, November). An image enhanced, knowledge based expert system for maintenance trouble shooting. *Computers in Industry*, pp. 187-202.

Miller, F. D., Rowland, J. R., & Siegfried, E. M. (1986, January). ACE: An expert system for preventive maintenance operations. *Record*, pp. 20-25.

Moradian, S., Thompson, E. D., & Jenkins, M. A. (1991, May). New idea in on-line diagnostics improves plant performance. *Power*, pp. 49-51.

Nelson, B. C., & Smith, T. J. (1990). User interaction with maintenance information: A performance analysis of hypertext versus hardcopy formats. *Proceedings of the Human Factors Society 34th Annual Meeting*, 229-233.

Nordwall, B. D. (1989, June 19). CTA develops new computer system to speed civil aircraft maintenance. *Aviation Week & Space Technology*, pp. 153-157.

NYNEX cuts costs in 40 offices using expert system. (1990, September). *Industrial Engineering*, pp. 81- 82.

Paula, G. (1990, March). Expert system diagnoses transmission-line faults. *Electrical World*, pp. S41-S42.

Ray, A. K. (1991, June). Equipment fault diagnosis--a neural network approach. *Computers in Industry*, pp. 169-177.

Reason, J. (1987, March). Expert systems promise to cut critical machine downtime. *Power*, pp. 17-24.

Rodriguez, G., & River, P. (1986, July). A practical approach to expert systems for safety and diagnostics. *InTech*, pp. 53-57.

Roth, E. M., Bennett, K. B., & Woods, D. D. (1987). Human interaction with an "intelligent" machine. *International Journal of Man-Machine Studies*, 27, 479-525.

Rowan, D. A. (1988, May). AI enhances on-line fault diagnosis. *InTech*, pp. 52-55.

Rustace, P. (1988, June 9). Knowledge of an expert compressed on computer. *The Engineer*, p. 44.

- Save plant know-how with expert systems. (1987, August). *Electrical World*, pp. 54-55.
- Schaaf, J. R. (1985, September). Computerization of sewer maintenance scheduling. *Public Works*, pp. 128-129.
- Shifrin, C. A. (1985, October 28). Eastern computer system reduces maintenance layovers, staff levels. *Aviation Week & Space Technology*, pp. 40-45.
- Smith, D. J. (1987, May). Diagnostic analysis leads the way in preventive maintenance. *Power Engineering*, pp. 12-19.
- Smith, D. J. (1989, January). Artificial intelligence--today's new design and diagnostic tool. *Power Engineering*, pp. 26-30.
- Smith, D. J. (1989, December). Intelligent computer systems enhance power plant operations. *Power Engineering*, pp. 21-26.
- Stacklin, C. A. (1990, June). Pairing on-line diagnostics with real-time expert systems. *Power*, pp. 55-58.
- Stein, K. J. (1988, March 14). Expert system technology spurs advances in training, maintenance. *Aviation Week & Space Technology*, pp. 229-233.
- Stovicek, D. (1991, February). Cloning knowledge. *Automation*, pp. 46-48.
- Suchman, L. A. (1987). *Plans and situated actions: The problem of human machine communication*.
- Sutton, G. (1986, January). Computers join the maintenance team. *WATER Engineering & Management*, pp. 31-33.
- Thandasseri, M. (1986). Expert systems application for TXE4A exchanges. *Electrical Communication*, 60(2), 154-161.
- Toms, M., & Patrick, J. (1987). Some components of fault-finding. *Human Factors*, 29(5), 587-597.
- Turpin, B. (1986, March 3). Artificial intelligence: Project needs. *Design News*, pp. 104-114.
- Users get expert advice on-site. (1987, March 12). *ENR*, p. 21.
- Uttley, A. (1985, October 17). Computer 'expert' helps find faults. *The Engineer*, p. 76.
- Vanpelt, H. E., & Ashe, K. L. (1989, April). Radiation exposure reduced with computer-aided maintenance. *Power Engineering*, pp. 40-42.
- Woods, D. D. & Roth, E. M. (1988). Cognitive engineering: Human problem solving with tools. *Human Factors*, 30(4), 415-430.
- Yu, C-C., & Lee, C. (1991). Fault diagnosis based on qualitative/quantitative process knowledge. *AIChE Journal*, 37, pp. 617-628.

Chapter 3 Appendix Annotated References

Aerospace maintenance. (1986, December). *Aerospace America*, p. 46.

Article describes artificial intelligence software that continuously monitors systems, isolates faults, and indicates fault presence, on Boeing B1B bomber. The system is projected to save \$260 million in maintenance. Also discusses an F/A-18 on board maintenance processor that creates data files in

flight to be later processed by an expert system. See also "AI to help keep..."

Ahrens, R. B., Marsh, A., & Shannon, P. A. (1984, November). 3B20D computer: Maintenance with a mind of its own. *Record*, pp. 16-19.

Discusses replacement of panel status indicators with microcomputer status indicators for maintenance.

AI to help keep aircraft flying. (1986, June 12). *Machine Design*, p. 4.

'Avionics Integrated Maintenance Expert System' (AIMES) was developed by McDonnell Douglas to monitor circuit cards in flight on F/A-18 Hornet. AIMES can identify which card has a failed component. See also "McDonnell Douglas..."

Armor, A. F. (1989, July). Expert systems for power plants: The floodgates are opening. *Power Engineering*, pp. 29-33.

Discusses the future of expert systems in the power industry, particularly for failure prevention and diagnosis.

Artificial intelligence to aid in war on potholes. (1985, December 12). *Engineering News-Record*, p. 215.

Describes research and development efforts at the University of California on a system for diagnosing and repairing pavement faults.

Atwood, M. E., Brooks, R., & Radlinski, E. R. (1986). Causal models: The next generation of expert systems. *Electrical Communication*, 60(2), 180-184.

A concept paper that distinguishes between 'shallow models' (models that use empirical data to detect previously observed faults) and 'causal models' (models that reason from functional models of system).

Barney, C. (1985, December 23). Expert system makes it easy to fix instruments. *Electronics*, p. 26.

Describes an expert system for diagnosing, repairing, and calibrating electronic instruments. The system has been applied to a signal-switching system. The system is VAX based.

Benedict, P., Tesser, H., & O'Mara, T. (1990, June). Software diagnoses remote computers automatically. *Automation*, pp. 46-47.

Grumman Data Systems and Grumman Systems Support developed a 'Remote Diagnostic System' (RDS) that diagnoses computer malfunctions. The VAX-based system is fully automatic; there is no human involved in the diagnosis process. The [RDS](#) prints a prioritized list of suspect printed-circuit boards with explanations on how the conclusions were reached. [RDS](#) can also serve as a consultant to a human diagnostician. [RDS](#) was designed to perform with the proficiency of an intermediate level diagnostician and to serve as the tool to be used first in diagnosing problems. The [RDS](#) combines rule- and model-based reasoning.

Bertheouex, P. M., Lai, W., & Darjatmoko, A. (1989). Statistics-based approach to wastewater treatment plant operations. *Journal of Environmental Engineering*, 115, 650-671.

Describes an expert system for daily operation of wastewater treatment plant. Uses d-Base III and Lotus 1-2-3.

Bogard, W. T., Palusamy, S. S., & Ciaramitaro, W. (1988, May). Apply automation to diagnostics, predictive maintenance in plants. *Power*, pp. 27-32.

The 'Advanced Diagnostic and Predictive-Maintenance System' is a system for monitoring and diagnosing problems at nuclear and fossil power plants. The system also schedules 'predictive maintenance', wherein maintenance is scheduled based on performance trends. The article describes system modules, with an emphasis on trend monitoring and preventive maintenance.

Bretz, E. A. (1990, July). Expert systems enhance decision-making abilities of O&M personnel. *Electrical World*, pp. 39-48.

Overview of several expert systems. Houston Lighting and Power Co. uses three systems: a materials management system that tracks spare parts and supplies, a maintenance management controls system that consolidates and standardizes methods for requesting and tracking maintenance, and an expert management scheduling system that generates reports and creates schedules.

Westinghouse Electric Corp. developed 'Argus', an alarm response advisor. Argus details alarm causes and required responses. The system collects data on-line, diagnoses problems, and makes recommendations.

Computational Systems, Inc. developed an expert system for vibrational analysis of rotating machinery.

Chubu Electric Power Co. and Toshiba Corp. developed a maintenance support expert system for large turbine generators. The system handles complex and time-consuming tasks. A engineer enters a failure into the system and the system responds with other damages that may result from a suspected root cause, it gives standard repair methods and design specifications, and it displays the most likely failure sources.

Chubu Electric Power Co. and Mitsubishi Heavy Industries, Ltd. developed an expert system for boiler failure analysis and maintenance planning. Failure analysis produces the most probable causes, guidelines for inspection, items to be investigated, repair methods, and suggested preventive maintenance. A maintenance planning subsystem automatically prepares daily repair schedules, work estimation plans, and work specifications.

Byrd, T. A., Markland, R. E., & Karwan, K. R. (1991, July-August). Keeping the helicopters flying--using a knowledge-based tank support system to manage maintenance. *Interfaces*, pp. 53-62.

Discusses a knowledge-based system which generates reports for helicopter maintenance. The system tracks helicopters and notifies maintenance staff of which helicopters are nearing regular inspections or special inspections. Reports specify the time-between-overhaul components that will require maintenance soon and give flying schedules prioritized on mission and maintenance needs. The system replaced a cumbersome manual system.

Callahan, P. H. (1988, January-February). Expert systems for AT&T switched network maintenance. *AT&T Technical Journal*, pp. 93-103.

Describes 'Testing Operations Provisioning Administration System' (TOPAS), a real-time, distributed, multi-tasking expert system for switched circuit maintenance. [TOPAS](#) performs trouble analysis, fault localization, and referral for network switches. [TOPAS](#) is claimed to do maintenance without human intervention or consultation.

Clancy, C. (1987, November). Qualitative reasoning in electronic fault diagnosis. *Electrical Engineering*, pp. 141-145.

Describes an expert system for diagnosing switch mode power supplies by using a functional model.

Computer oversees maintenance. (1992). *American Water Works Association Journal*, 84, 107-108.

Discusses a pc-based preventive maintenance and training system for water mains.

Cue, R. W. & Muir, D. E. (1991). Engine performance monitoring and troubleshooting techniques for the CF-18 aircraft. *Journal of Engineering for Gas Turbines and Power*, 113, 11-19.

Discusses the In-flight Engine Condition Monitoring System (IECMS) as a foundation for 'on-condition' maintenance of fighter aircraft engines. On-condition maintenance actions are undertaken based on actual engine conditions, rather than as preventive maintenance. [IECMS](#) monitors and records engine performance parameters, notifies the pilot when caution should be exercised, and

records maintenance codes when an operating limit has been exceeded. Data are stored on a removable tape cartridge. The article features several examples of data indicating normal and abnormal operating conditions.

Culp, C. H. (1989). Expert systems in preventive maintenance and diagnostics. *ASHRAE Journal*, 31, 24-27.

Sales article on using expert systems for heating, ventilation, and air conditioning maintenance.

Dallimonti, R. (1987, June 18). Smarter maintenance with expert systems. *Plant Engineering*, pp. 51-56.

An introduction to the prospects of using expert systems for maintenance. Surveys systems from Hughes Aircraft Company, Rockwell International, and Campbell Soup Company.

de Kleer, J. (1990). Using crude probability estimates to guide diagnosis. *Artificial Intelligence*, 45, 381-391.

Extension of 'General Diagnostic Engine' discussion in de Kleer and Williams.

de Kleer, J. & Williams, B. C. (1987). Diagnosing multiple faults. *Artificial Intelligence*, 32, 97-130.

An academic discussion of a 'General Diagnostic Engine' for diagnosing multiple faults. Combines model-based prediction with sequential diagnosis to propose measurements to diagnose faults.

Dobson, R., & Wild, W. (1989, May). Plant's computerized maintenance system improves operations. *Power Engineering*, pp. 30-32.

A Lotus Symphony-based system automatically processes maintenance clearances for power plant.

Dohner, C. V., & Acierno, S. J. (1989, August). Expert systems for gas-turbine powerplants passes first tests. *Power*, pp. 63-64.

Citing the limited amount of troubleshooting and diagnostic information in the manufacturer's maintenance manuals, [EPRI](#) developed an expert system for gas-turbine power plants. A portable pc uses voice recognition and synthesis and links to a pc in control room. The control room pc drives a video disk player and a printer.

Doorley, R. (1988, August). Hydraulic troubleshooting using an expert system. *Hydraulics & Pneumatics*, pp. 91-92.

Discusses the 'MindMeld' system for steel mill hydraulic equipment maintenance. MindMeld uses test equipment data and operator information to determine the likely cause of a problem. See also Doorley, (1989).

Doorley, R. B. (1989, June 22). Expert systems probe hydraulic faults. *Machine Design*, pp. 89-92.

More on the 'MindMeld' system for hydraulic equipment troubleshooting in steel mills. The pc-based system focuses on faults that are difficult to locate and which require extensive dismantling of machinery if left unrecognized.

Expert system from AT&T Bell Laboratories is an 'ACE' at telephone cable analysis. (1983, October). *Record*, p. 1.

The 'Automated Cable Expertise' system identifies trouble spots in telephone cable systems. ACE gives repair histories of problematic areas and suggests corrective action. The system is automatic and runs daily.

Expert system guides tube-failure investigations. (1989, August). *Power*, p. 85.

Discusses an expert system for boiler tube failure diagnosis and corrective action (including non-destructive examination, repair, welding, metallurgical tests, references). The system can be used to

determine tube failure mechanisms. The system also has a database for tube history, design, inspection, maintenance and it provides context-sensitive information about repair practices. PC-based, linked to a slide projector. See also Smith, (1989, December).

Expert system probes beneath the surface. (1990, January). *Mechanical Engineering*, p. 112.

Britain's Central Electricity Generating Board developed an expert system for monitoring and analyzing vibration patterns of turbine generators. The expert system uses 'deep knowledge' of cause and effect relationships in turbine generators. The goal in developing the system was to transfer initial analysis from specialist staff to engineering/operations staff.

Expert systems to hone jet engine maintenance. (1986, April 21). *Design News*, pp. 36-38.

Describes an expert system to diagnose engine malfunctions and facilitate preventive maintenance by predicting when parts must be replaced. The system switches maintenance from a scheduled replacement basis to an 'as-needed' replacement basis. It uses qualitative and historical maintenance data. The pc-based system was developed by the General Electric R&D Center in conjunction with GE's Aircraft Engine Business Group for the Air Force.

FAA and NASA design program to improve human performance. (1989, May 29). *Aviation Week & Space Technology*, p. 115.

Outlines joint [FAA](#) and NASA effort to sponsor human factors research.

Foley, W. L., & Svinos, J. G. (1989). Expert advisor program for rod pumping. *Journal of Petroleum Technology*, 41, 394-400.

'EXPROD' is an expert adviser program for rod-pumping diagnostics used by Chevron. The program analyzes field data to identify equipment problems and recommend solutions. EXPROD uses statistical pattern recognition in conjunction with diagnostic rules. Some worker expertise is still required to diagnose problems. EXPROD runs on a microcomputer.

Folley, J. D., & Hritz R. J. (1987, April). Embedded AI expert system troubleshoots automated assembly. *Industrial Engineering*, pp. 32-35. Discusses an expert system to assist technicians in diagnosing a clutch assembly machine. The expert system uses a computer-controlled video disk to indicate what the technician should be doing or looking at.

Fault lamps above stations indicate malfunctioning assembly stations. The technician takes a monitor, a two-button control, and a maintenance cart to the faulty station and plugs into a junction box connected to the computer and video disk player. The technician selects the station or procedure from a menu and the computer specifies tests or actions with graphics. The technician supplies data and the computer specifies the next action.

GM unveils "Charley", an expert machine diagnostic system. (1988, May). *I&CS*, pp. 4-7.

Describes vibration analysis expert system for production machinery with rotating components. Charley helps mechanics: 1. identify parts that need repair; 2. repair or adjust equipment prior to failure; 3. speed up diagnosis; 4. distribute expertise; and 5. avoid fixing functioning equipment. See also Stovicek, (1991).

Gunhold, R., & Zettel, J. (1986). System 12 in-factory testing. *Electrical Communication*, 60(2), 128- 134.

Describes a diagnostic expert system for ITT System 12 printed circuit board assemblies.

Hartenstein, A. (1988, January). Computer system controls all maintenance activities. *Public Works*, p. 60.

The system maintains several types of records and schedules preventive and corrective maintenance. It also issues work orders and monitors progress. It is a database system.

Hill, S. (1990, February). Ask the expert. *Water & Pollution Control*, pp. 12-13.

Concept paper that discusses possibilities of expert systems to design wastewater treatment facilities and control such plants.

Hughes, D. (1988, March 7). Digital develops special applications to meet diverse aerospace needs. *Aviation Week & Space Technology*, pp. 51-53.

Jet fighter uses AI as troubleshooter. (1986, July 21). *Design News*, p. 20.

More on [AIMES](#). See also "McDonnell Douglas..."

Keller, B. C. & Knutilla, T. R. (1990, September). U.S. Army builds an AI diagnostic expert system, by soldiers for soldiers. *Industrial Engineering*, pp. 38-41.

Describes the 'Pulse Radar Intelligent Diagnostic Environment' for troubleshooting the Pulse Acquisition Radar of a Hawk missile system.

King, I. J., Chianese, R. B., & Chow, M. P. (1988, December). Plant diagnostics relies on AI transmissions from remote site. *Power*, pp. 57-60.

Describes a suite of on-line power plant diagnostic systems developed by Westinghouse Electric Corp. Development goals were to maximize availability and efficiency and reduce forced-outage rates of turbine generators. The systems identify worn or damaged components early. 'ChemAID' diagnoses problems in the steam/water cycle; it determines the type, severity, and location of water chemistry problems. ChemAID assists the operator in determining the need for immediate or delayed action. It can also serve as a consultant.

'TurbinAID' diagnoses problems in steam turbines. It diagnoses the condition and thermodynamic performance of turbines and reports current and target performance parameters.

'GenAID' monitors trends for gas-cooled generators.

Kinnucan, P. (1985, November). A maintenance expert that never sleeps. *High Technology*, pp. 48-9.

The 'Intelligent Machine Prognosticator' (IMP) is an expert system for maintenance of an epitaxial reactor (equipment that 'grows' additional silicon crystals on silicon wafers). IMP diagnoses faults and recommends repair procedures. The system was developed because vendor support was difficult to obtain. IMP reportedly reduced repair time by 36%.

Kolcum, E. H. (1989, January 2). Growing flight, maintenance simulator market attracts many competitors. *Aviation Week & Space Technology*, pp. 91-93.

Discusses proposed military expenditures on simulators for maintenance training.

Layton, C. F. (1992). An investigation of the effects of cognitive tools on human adaptive planning in the domain of enroute flight planning. Doctoral dissertation, The Ohio State University, Columbus, OH.

Discusses a study which compared three different levels of computer support for enroute flight planning. The study compared the behaviors of thirty professional airline pilots assigned randomly to each of three treatment conditions (ten subjects per condition, each condition consisted of a different form of computer support). The subjects were trained on system use prior to solving four enroute flight planning scenarios. The focus of the research was not on the principles used to design the enroute flight planning systems, rather it was on what characteristics the designs shared and what characteristics were unique to a particular system. The purpose of the research was to study how these three system designs, as examples of broader classes of planning assistance tools, affected enroute flight planning, as an example of adaptive planning. The goals of this study were to develop a better understanding of the adaptive planning process and to develop recommendations for designing tools to support that process.

Maintenance expert in a briefcase. (1986, April). *High Technology*, p. 9.

`Mentor' is a portable expert system for routine maintenance and diagnosis of air conditioners. Mentor keeps service records of each piece of equipment. The emphasis is on preventive maintenance.

Majstorovic, V. D. (1990, October). Expert systems for diagnosis and maintenance: State of the art. *Computers in Industry*, p. 43-68.

Discussion of typical expert system components and survey of diagnosis and maintenance expert systems. Discusses expert systems for maintenance of flexible manufacturing systems.

McDonnell Douglas flight tests AI maintenance data processor. (1986, February 17). *Aviation Week & Space Technology*, p. 69.

Describes `Avionics Integrated Maintenance Expert System' (AIMES) for F/A-18. [AIMES](#) gathers aircraft data and creates flight files for later analysis. Production rules detect and isolate avionic failures at the electronic card level. Analysis provides fault data, the card name, and the reasoning that led to the fault isolation conclusion. [AIMES](#) is a self-contained on board system with a microprocessor and a data storage cassette. [AIMES](#) includes [BITE](#). See also "AI to help keep..."

McDowell, J. K., & Davis, J. F. (1991). Managing qualitative simulation in knowledge-based chemical diagnosis. *AIChE Journal*, 37, 569-580.

An academic discussion of an approach for dealing with multiple interacting faults with an expert system.

Melhem, H. G., & Wentworth, J. A. (1990, March). FASTBRID: An expert system for bridge fatigue. *Public Roads*, pp. 109-117.

`Fatigue Assessment of Steel Bridges' is a training aid for bridge inspection and planning remedial actions. It is also an advisory system for evaluation of bridges. The advisory system helps organize fatigue inspection, evaluate inspection results, and determine a course of action.

Miller, D. M., Mellichamp, J. M., & Wang, J. (1990, November). An image enhanced, knowledge based expert system for maintenance trouble shooting. *Computers in Industry*, pp. 187-202.

Describes an expert system for diagnosis of the electrical/hydraulic system of an electric utility vehicle. The vehicle system is difficult to diagnose because faults can be masked by behavior of equipment. The expert system displays a limited number of photographs of a vehicle. The expert system is pc based.

Miller, F. D., Rowland, J. R., & Siegfried, E. M. (1986, January). ACE: An expert system for preventive maintenance operations. *Record*, pp. 20-25.

Discusses the `Automated Cable Expertise' (ACE) system for telephone cable analysis. ACE is an automatic report generator which runs daily. See also "Expert system from AT&T Bell Laboratories..."

Moradian, S., Thompson, E. D., & Jenkins, M. A. (1991, May). New idea in on-line diagnostics improves plant performance. *Power*, pp. 49-51.

Discusses Westinghouse expert systems, particularly GenAID. See also King, Chianese, and Chow (1988, December).

Nelson, B. C., & Smith, T. J. (1990). User interaction with maintenance information: A performance analysis of hypertext versus hardcopy formats. *Proceedings of the Human Factors Society 34th Annual Meeting*, 229-233.

Research paper which discusses experimental results of comparing hypertext versions of maintenance manuals with hardcopy versions of those manuals. Subjects were slower with hypertext than hardcopy, but preferred hypertext. Enhanced versions of manuals (hypertext and hardcopy)

improved access to information and comprehension of that information, but subjects did not recognize improvements.

Nordwall, B. D. (1989, June 19). CTA develops new computer system to speed civil aircraft maintenance. *Aviation Week & Space Technology*, pp. 153-157.

Describes the 'Automated Maintenance Management System' (AMMS) and the 'Mobile Enhanced Comprehensive Asset Management System' (MECAMS) developed by CTA, Inc. The [AMMS](#) collects inflight information on F/A-18 aircraft engines and stores it on a floppy disk for later analysis by [MECAMS](#). [MECAMS](#) runs on a laptop computer connected to a minicomputer, which in turn could be connected via satellite or phone link to a database of troubleshooting and logistics information. [MECAMS](#) first identifies periods when engine performance parameters are exceeded, then it assists technicians in troubleshooting and maintaining the engines. CTA indicated that the F/A-18 engine system was a proof of concept and was proposing that the same system could be extended to civil aircraft engines and avionics.

NYNEX cuts costs in 40 offices using expert system. (1990, September). *Industrial Engineering*, pp. 81-82.

'Maintenance Administrator Expert' (MAX) diagnoses problems with residential and small business telephone service. MAX interprets trouble report data in 5-10 sec. as opposed to the 5-10 min. a human requires. MAX dispatches the correct technician and reduces false dispatches. The system reportedly saves \$4-6 million/yr. MAX was developed by NYNEX Science and Technology Center and runs on a Sun 3/260.

Paula, G. (1990, March). Expert system diagnoses transmission-line faults. *Electrical World*, pp. S41-S42.

Describes an automatic expert system that uses a mathematical model of a power transmission system for fault diagnosis. The expert system activates upon failure of the transmission system and produces a prioritized list of possible fault locations and their corresponding failure probabilities.

Ray, A. K. (1991, June). Equipment fault diagnosis--a neural network approach. *Computers in Industry*, pp. 169-177.

Describes a system designed for mechanical equipment diagnosis in the steel industry.

Reason, J. (1987, March). Expert systems promise to cut critical machine downtime. *Power*, pp. 17-24.

Discusses prospects for continuous vibration monitoring systems, leading to automatic diagnostic systems. The article also highlights several expert systems in the power industry. 'Turbomac' is an expert system for diagnosing vibrations in large turbo machinery, particularly power generating facilities.

'GenAid' is an expert system developed by Westinghouse Electric Corp. for diagnosing hydrogen-cooled electric generators. The purpose of GenAid was to avoid catastrophic failure. See also King, Chianese, and Chow (1988, December).

The Central-Hudson Electric & Gas Corp. developed an expert system for scheduling outages. The purpose of the system was to reduce the number of scheduled outages without compromising equipment integrity. The system schedules outages for preventive maintenance at the first sign of trouble.

'Transformer Oil Gas Analyst' is an expert system for detecting and diagnosing signs of impending transformer failure.

General Electric developed an expert system for turbines. The system is portable, links to a video disk display, and uses voice recognition for form fill-in or multiple choice input.

Rodriguez, G., & River, P. (1986, July). A practical approach to expert systems for safety and

diagnostics. *InTech*, pp. 53-57.

Describes an expert system for diagnosis of a 400/200 KV hybrid gas insulated substation of the Laguna Verde Nuclear Power Station in Mexico. Engineers and literature provided information to build fault trees to model loss of current to safety-related control boards. System objectives were the timely diagnosis of abnormal events or transients, and analysis of events leading to, and consequences of, an abnormal situation.

Roth, E. M., Bennett, K. B., & Woods, D. D. (1987). Human interaction with an "intelligent" machine. *International Journal of Man-Machine Studies*, 27, 479-525.

Reports a study investigating technicians using an expert system to troubleshoot an electro-mechanical device. The article documents common problems of 'machine expert' problem solving systems. Only technicians who were actively involved in the problem solving process and who performed actions in addition to those requested by the expert system successfully completed the sample tasks. Technicians who responded passively to expert system requests were unable to solve the problems. This study should be read by all those interested in improving human performance through computational support.

Rowan, D. A. (1988, May). AI enhances on-line fault diagnosis. *InTech*, pp. 52-55.

Describes 'Fault Analysis Consultant' (Falcon) for on-line fault diagnosis in a commercial chemical plant. Falcon reasons from first principles and heuristic knowledge. Falcon went on-line in 1988.

Rustace, P. (1988, June 9). Knowledge of an expert compressed on computer. *The Engineer*, p. 44.

Discusses monitoring expert system for gas turbine-driven compressor sets.

Save plant know-how with expert systems. (1987, August). *Electrical World*, pp. 54-55.

Discusses expert system to resolve power plant control room alarms.

Schaaf, J. R. (1985, September). Computerization of sewer maintenance scheduling. *Public Works*, pp. 128-129.

'Computerization of Sewer Maintenance Operations' (COSMO) schedules routine cleaning operations. COSMO also tracks performance, debris severity, and maintenance history and uses this information to set cleaning priorities and schedule sewer cleaning. Database system.

Shifrin, C. A. (1985, October 28). Eastern computer system reduces maintenance layovers, staff levels. *Aviation Week & Space Technology*, pp. 40-45.

Describes a computer system that includes computerized work cards to be filled in by maintenance personnel. The system will also produce hard copies of tasks and checks. The work cards contain detailed instructions, warnings, and notes. The system aids in capacity planning through tracking line slippage, schedule constraints, and manpower limitations.

The article also discusses a parts tracking and scheduling system.

Smith, D. J. (1987, May). Diagnostic analysis leads the way in preventive maintenance. *Power Engineering*, pp. 12-19.

Describes 'Process Diagnosis System' (PDS) developed by Westinghouse R&D Center and Carnegie Mellon University. [PDS](#) diagnoses problems with steam generators and provides recommendations and procedures for fixing the problems. [PDS](#) is claimed to cut down on 'over-maintenance', but prevent 'breakdown maintenance'. The system monitors the condition of an operational plant and analyzes plant data to detect incipient faults and deterioration. [PDS](#) uses this information to diagnose faults and predict the duration of safe operation without maintenance. [PDS](#) also recommends preventive maintenance tasks to be performed during scheduled down periods.

Smith, D. J. (1989, January). Artificial intelligence--today's new design and diagnostic tool. *Power Engineering*, pp. 26-30.

Overview of AI applications in maintenance. The article describes the [EPRI](#)-developed 'Gas Turbine Expert System' and a Westinghouse-developed system for on-line valve diagnosis

Smith, D. J. (1989, December). Intelligent computer systems enhance power plant operations. *Power Engineering*, pp. 21-26.

Discusses several expert systems in use in the power industry. The 'ESCARTA' system for reducing boiler tube failures has several uses: it permits an engineer to track down a failure mechanism, it provides non-destructive testing procedures, it provides welding procedures, it provides corrective actions for failure repair, and it facilitates training. ESCARTA will show operators the correct procedures for investigating tube failures. It also suggests root causes that could have led to a failure. ESCARTA is pc based.

'Coal Quality Advisor' assesses coal quality. The system helps assess cost and performance aspects of using different coals or coal blends. It is pc based.

'Smart Operator's Aid for Power Plant Optimization' diagnoses causes of heat rate degradation on oil- and gas-fired power plants. The system justifies its diagnosis through logic trees or messages. It also recommends corrective actions.

'TurbinAID', 'GenAID', and 'ChemAID' make up a suite of expert systems for diagnosis of turbine generators. The systems were developed by Westinghouse Electric Corp. See also King, Chianese, and Chow (1988, December).

Stacklin, C. A. (1990, June). Pairing on-line diagnostics with real-time expert systems. *Power*, pp. 55-58.

Concept paper on using expert systems to reduce unscheduled down time. Discusses fault trees, failure modes and effects analyses, and pattern recognition.

Stein, K. J. (1988, March 14). Expert system technology spurs advances in training, maintenance. *Aviation Week & Space Technology*, pp. 229-233.

Discusses emerging expert system technology in maintenance. The Navy Sea System Command's Integrated Diagnostic Support System (IDSS) collects fault-related data and isolates faults. The system continually builds its knowledge base so that it becomes more efficient with use. [IDSS](#) will isolate faults to the microchip level by using fault trees. [IDSS](#) uses a touch screen, a flat panel display, and an interactive maintenance tutorial on a video disk. [IDSS](#) is expected to aid systems designers in building self-diagnostics in new avionics systems.

The article also describes Flex-MATE for use with the USAF modular automatic test equipment.

Stovicek, D. (1991, February). Cloning knowledge. *Automation*, pp. 46-48.

Discusses General Motors' 'Charley' system for vibration analysis of production equipment. Charley is an expert system developed by GM's Advanced Engineering Staff and is named and modeled after a retired vibration analysis expert. Charley contains three modules: 1. a rule-base for vibration analysis, 2. a 'vibration signature' database, which contains the vibration curves of the various pieces of equipment, and, 3. a machine database, which contains historical data on each machine. Charley is used for failure diagnosis, preventive maintenance, and training. Charley can be used to answer 'what if' questions and explains diagnosis strategies. Runs on a Sun computer. See also "GM unveils..."

Suchman, L. A. (1987). *Plans and situated actions: The problem of human machine communication*.

Discusses human action with respect to circumstantial variability and the difficulties in communicating such variability to a machine. Rigid problem solving on the part of the machine and misinterpretations on the part of the human are some of the obstacles to successful human computer interaction discussed.

Sutton, G. (1986, January). Computers join the maintenance team. *WATER Engineering & Management*, pp. 31-33.

`Maintenance Management System' (MMS) for water and wastewater treatment facilities. [MMS](#) tracks organization performance, determines resource utilization and work backlog, and makes personnel and resource utilization projections. It also schedules preventive maintenance. [MMS](#) is a database system.

Thandasseri, M. (1986). Expert systems application for TXE4A exchanges. *Electrical Communication*, 60(2), 154-161.

Describes `Advance Maintenance Facility', an expert system for fault identification. Normally the system controls interactions, but it can be `controlled' by an operator. Output is corrective action and post-repair tests.

Toms, M., & Patrick, J. (1987). Some components of fault-finding. *Human Factors*, 29(5), 587-597.

Research paper on human performance in network fault-finding tasks.

Turpin, B. (1986, March 3). Artificial intelligence: Project needs. *Design News*, pp. 104-114.

Discusses expert system developed by Campbell Soup for diagnosing problems with `cookers'. The system was built to replace a retiring technician with 25 years of experience.

The article also discusses an expert system (Intelligent Machine Prognosticator) developed by Texas Instruments for epitaxial reactor maintenance. See also Kinnucan (1985, November).

Users get expert advice on-site. (1987, March 12). *ENR*, p. 21.

The `Exstra' expert system troubleshoots mechanical equipment failures in compressors, water pumps, and other rotating equipment. Exstra lists possible conclusions with likelihood ratings. It also explains why it asks particular questions. Exstra is VAX based.

Uttley, A. (1985, October 17). Computer `expert' helps find faults. *The Engineer*, p. 76.

Describes expert system shell software.

Vanpelt, H. E., & Ashe, K. L. (1989, April). Radiation exposure reduced with computer-aided maintenance. *Power Engineering*, pp. 40-42.

Describes `Plant Radiological Status' (PRS), a three dimensional computer model of a power generating station and its equipment, developed by Duke Power Co. [PRS](#) is claimed to reduce maintenance time and radiation exposure by supporting planning activities. [PRS](#) identifies maintenance interference problems (e.g., restricted access) and available work space. [PRS](#) also facilitates access and interpretation of radiological conditions in the plant; it identifies hot spots and contaminated areas.

Woods, D. D. & Roth, E. M. (1988). Cognitive engineering: Human problem solving with tools. *Human Factors*, 30(4), 415-430.

Concept paper which describes the fundamental aspects of cognitive engineering. According to the authors, "Cognitive engineering is an applied cognitive science that draws on the knowledge and techniques of cognitive psychology and related disciplines to provide the foundation for principle-driven design of person-machine systems." (p. 415). Like Roth, Bennett, & Woods (1987), this article should be read by all those interested in supporting human performance through computational support.

Yu, C-C., & Lee, C. (1991). Fault diagnosis based on qualitative/quantitative process knowledge. *AIChE Journal*, 37, pp. 617-628.

An academic discussion of an approach for combining qualitative and quantitative reasoning in an

expert system through fuzzy logic. Aside: Although fuzzy logic is a technique that is often used in artificial intelligence because it doesn't carry the overhead associated with Bayesian or other probability theory based methods, it also lacks the mathematical rigor of the latter methods.

Chapter Four

The [FAA](#) Aviation Maintenance Human Factors Hypermedia System

4.0 INTRODUCTION

The Federal Aviation Administration (FAA) Aviation Maintenance Hypermedia Information System (HIS) is part of the Office of Aviation Medicine (AAM) Human Factors in Aviation Maintenance Research (HFAMR) program. The goal of the [HIS](#) project is to create new tools and methods for information access and use and to provide these tools and methods for support of other [HFAMR](#) activities (e.g., training and job aiding systems). These tools and methods provide the vital information access component for any computer-based aviation maintenance integrated information system (Johnson and Norton, 1992).

In its present state the [HIS](#) provides an environment to create and explore large collections of related information. The [HIS](#) provides a simple, yet powerful, way of creating and following associations between related pieces of information. Using the [HIS](#), the user can browse and view information in a variety of ways. This flexible method of information access and utilization is not available to users of conventional text retrieval systems.

Publications and presentations from the [HFAMR](#) program are being placed into the [HIS](#) for the initial domain. This material includes presentations from the first five Human Factors in Aviation Maintenance Conferences, as well as complete material from the [HFAMR Phase I](#) (Shepherd, et al., 1991) and Phase II Progress Reports. The result is an on-line document that employs the latest hypermedia software technology. More importantly, software tools and methods have been constructed that provide the ability to quickly store, locate, and deliver information for a variety of aviation maintenance tasks.

4.1 PRACTICAL ASPECTS OF HYPERMEDIA

The fundamental nature of computer-based hypermedia is to structure information in a fashion that can be quickly and randomly located. Conventional forms of media delivery (e.g., books, television, video, audio tapes, etc.) tend to be linear in nature. The reader of this linear information typically starts at the beginning of a presentation and progresses along a predetermined path (e.g., turning to the next page in a book, or being forced to sit through commercials while watching TV). The reader or viewer of this information often has little choice in determining what information comes next.

Hypermedia, on the other hand, arranges information in a form in which a reader or viewer can "bounce" around between different segments of information (similar to a reader of a mystery book flipping to the end of the story; or a reader of a technical manual first looking up a term in an index, and then going to the correct page in the manual). This idea of associations, or links, between segments of information is a core feature of hypermedia systems. An example of a typical hypermedia system is illustrated in [Figure 4.1](#).

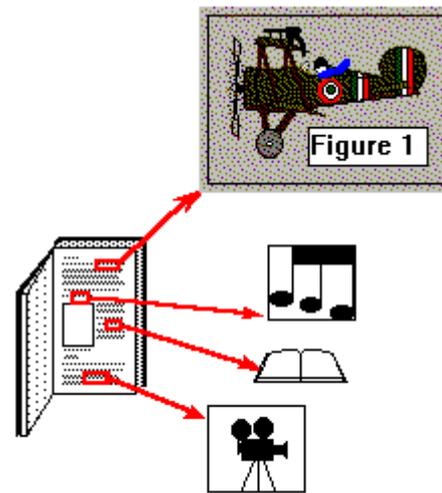


Figure 4.1 A Typical Hypermedia System

For example, a hypermedia version of a car maintenance manual would first present the reader with a diagram of a car. From this diagram, the reader would point to a particular portion of the car (e.g., the engine) and request more information. The hypermedia system would then display a diagram of the engine. The reader would point to an engine component (e.g., the battery) and the hypermedia system would then present a verbal description of the battery and list possible troubleshooting advice dealing with common battery problems.

For the purposes of this paper, the term "hypermedia" is used because the research is not restricting the usage of this system to only text - "hypertext." The scope of this system has been extended to include text, still images, graphical animation, audio, and video - "hypermedia".

One of the advantages of using hypermedia technology is its strong support for easy access to items stored in large collections of information. Typically, structures of large collections of information are complex. This leads to various indexing schemes that are used to aid readers in locating information (e.g., tables of contents and indices for books, Dewey Decimal and Library of Congress indexing for libraries). Even with these schemes however, someone searching for a particular piece of information may still have limited success in locating the desired information. Hypermedia systems extend the indexing of information by providing associations, or links, between particular related pieces of information. The idea is to provide readers with the ability to access a general location in the information base using various indexes and searching schemes. Then the reader can browse the general locations using specific associations to locate the desired information. To support browsing and exploration of information, the presence of anchors (also often called buttons) is employed to signal the reader that related information is available (Duchastel, 1990).

The Intermedia project at Brown University is a very good example of hypermedia technology being implemented for a large information space. Over a thousand pieces, or "nodes," of information are connected by over two thousand links. Intermedia presents the user with a graphical information browser, a set of graphical editors for text, graphics, timelines, animations, and videodisc data, a browser for link information, and the ability to create and traverse links between any two selections in any document in the system (Haan, Kahn, Riley, Coombs, and Meyrowitz, 1992). The Intermedia system illustrates the ease in which information can be placed into hypermedia systems and how that information can immediately benefit other users of the system.

The Artifact-Based Collaboration (ABC) project at the University of North Carolina (Smith and Smith, 1991) is a good example of hypermedia technology being used to support groups of individuals working together to build large, complex structures. [ABC](#) has five components that include a graph server, a set of graph browsers, a set of data application programs, a shared window

conferencing facility, and real-time video and audio. [ABC](#) also demonstrates the ability of hypermedia technology to assimilate new information and to quickly disseminate the information for immediate use.

4.2 THE FAA HYPERMEDIA SYSTEM

The overall goal of the [FAA](#) Hypermedia Information System (HIS) development is to study the prospects and problems of creating an electronic document. The project will also determine how these technologies can be used to improve the delivery of information to support aviation maintenance. Although focused on one particular discipline, this development should produce results that help to guide the development of many kinds of future aviation maintenance support applications, hereafter referred to as [FAA](#) integrated information systems.

4.2.1 Research Phases

The development of the [HIS](#) is divided into two phases that will be conducted over a two year period. The work began in July of 1991. The first eight months (Phase I) were dedicated to the prototype design and development. The results of this phase have been:

1. A functional on-line version of the Third [FAA](#) meeting on Human Factors Issues in Aircraft Maintenance and Inspection. This version is functionally equivalent to the final version to be implemented in 1992.
2. Specifications to authors of future [FAA](#) HF conference papers to aid the incorporation of new information into the [HIS](#).
3. Reusable technology base for indexing and retrieving hypermedia text and graphics.

The second phase is currently under development and is divided into four tasks. The first task focuses on incorporating the remaining four conference proceedings into the [HIS](#) and includes the production of a [CD-ROM](#) version of the [HIS](#). The second task focuses on enhancement of the Hypermedia technology base. The third task will involve support for the transition of Hypertext technology to [FAA](#) integrated information system research. The fourth task will focus on demonstrating and reporting research results to fellow HF team members, the project sponsor, and the HF community.

4.2.2 Features of the Hypermedia Information System

The Hypermedia Information System (HIS) was designed with a variety of features to aid a reader of the document in locating and using the information contained in the system. Most of these features were derived by analyzing the eventual information needs of the readers of the [HIS](#), as well as reviewing state-of-the-art information retrieval and hypermedia systems. These features are discussed in the subsequent sections.

4.2.2.1 Information Browsing

The [HIS](#) allows readers to browse through the information contained in the on-line database. That is, the user is allowed to leisurely wander through the information, selecting items of interest and inspecting various topics that pique the interest of the reader. For example, some readers might browse the information looking only at the pictures that may have some relevancy to the topic at hand. Other readers might choose to view only the titles and authors of individual papers, searching for a topic that may have some application to their current task.

The [HIS](#) supports this browsing operation in two ways. First, the information is loosely structured in the database, allowing the [HIS](#) to present the same information in many different ways. Second, the [HIS](#) provides mechanisms to the readers in which they can easily locate additional information

related to the current topic. For example, while reading a conference paper stored in the [HIS](#), a reader might come across a reference to a related photograph. The [HIS](#) allows the reader to select this reference to view the actual photograph.

This method of linking, or associating, information can easily extend to other objects stored in the [HIS](#). For example, readers can point to references to videos, animations, text, sound, etc. and request to see (or hear) this information. The [HIS](#) is able to retrieve these items and display them to the reader. This linking capability allows the reader to freely move about the information contained in the [HIS](#), literally wandering and exploring at will.

The [HIS](#) supports linking through the use of buttons and icons (see [Figure 4.2](#)). Buttons allow the reader to activate certain links, while icons provide some sort of visual clue to the reader as to what the button will do when it is activated. Usually icons are a picture of some object (e.g., a video camera to represent video). However, text can also serve as an icon (e.g., the words FIGURE 62 represent the actual figure). The [HIS](#) supports both text and graphics as icons to support the reader in determining when a button is present, and what that button will do when it is activated. The [HIS](#) also contains an overview diagram of the information contained in the database ([Figure 4.3](#)). This overview diagram allows a reader to view the entire [HIS](#) database at once. Using the overview diagram, the reader can gain a perspective of how one piece of information is related to other information. Also from the overview, the reader can select any piece of information and go directly to it, without having to browse through the system.

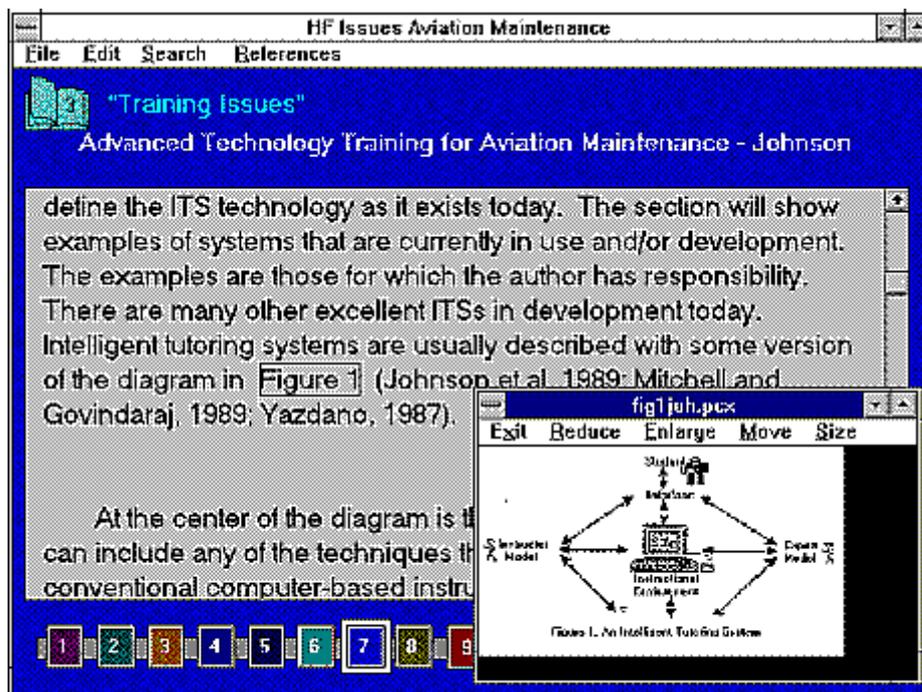


Figure 4.2 Buttons and icons allow readers to access related information

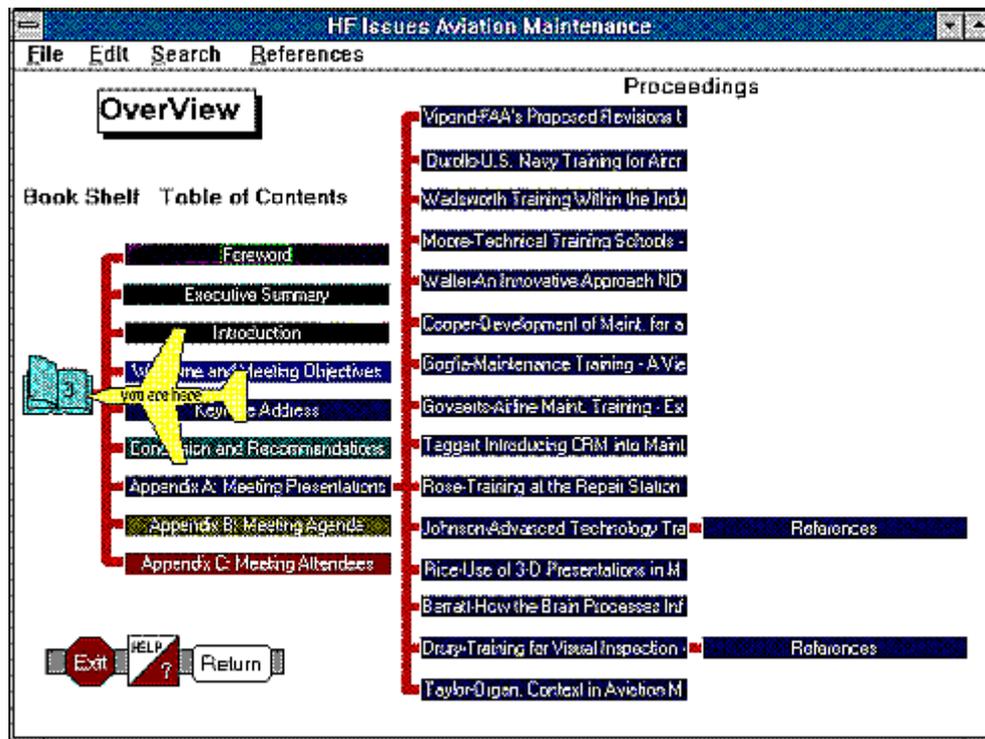


Figure 4.3 The Overview Diagram

4.2.2.2 Information Searching

Often a reader will come to the [HIS](#) with a specific idea of what information is needed. The chances of locating the exact piece of needed information simply by browsing are fairly low, even in relatively small documents. Even if a reader happens to come across one piece of relevant information, there might be additional information that the reader might miss. An even worse scenario is when a reader has a request for information that is not contained in the [HIS](#). The reader might browse indefinitely without finding a relevant piece of information. The reader would not be able to tell if their lack of success was due to their bad luck in browsing or to the fact that the information did not exist at all.

To overcome these problems, the [HIS](#) provides a facility for conducting direct searches for information. The [HIS](#) contains an index into the information that aids readers by allowing them to search the index, rather than the information for a piece of information. The index contains both the search items and the location of those items in the [HIS](#) system. This technique is similar to the uses one might find in the local telephone listings. For example to locate the residence of your friend Timothy Handson, you could use the telephone listings to locate all of the entries for Timothy Handson. Once you have found the correct entry, you are able to get the address of his residence and then go to his home. This directed search method is much more efficient than browsing.

The [HIS](#) system works in a fashion to the telephone listing analogy. The [HIS](#) system contains many indices indicating exactly what each piece of information is, including what type of information it is (i.e., text, video, graphics, sound), and where it is located (i.e., exactly what conference paper, what paragraph, and what sentence). Readers using the [HIS](#) would enter a description of the information that they are looking for, and the [HIS](#) uses these indices to tell the reader where this information is located. The reader can directly "jump" to this information, bypassing the entire browsing process.

In the present [HIS](#) system, the reader initiates a search by entering a query, or a description of the desired information, into the system ([Figure 4.4](#)). This query is made up of individual terms and boolean operators on those terms. For example, search terms could include: Airplane, Human

Factors, Helicopter, Navigation, etc. Any word is a valid search term. The reader also uses boolean operators (AND, OR, NOT) to indicate relationships between multiple search terms. For example, a possible query could be Computer AND Training, indicating that reader is looking for a piece of information that is related to both computers and training.

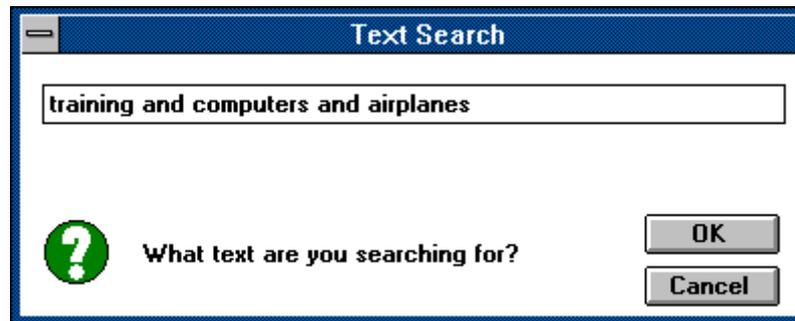


Figure 4.4 Entering Search Terms

The [HIS](#) uses this query to search the indices for related information. In the case of the example query above, the [HIS](#) would first look for information related to computers, then look for information related to training, and would then look for pieces of information that were in both of those wsets. The results of this search are then presented back to the reader (see [Figure 4.5](#)). Often, there are multiple pieces of information that match a readers query (e.g., just as there might be multiple listing for Timothy Handson in the telephone book). In this case, the readers are shown all of these matching pieces of information, and are allowed to browse through those that they feel could be relevant.

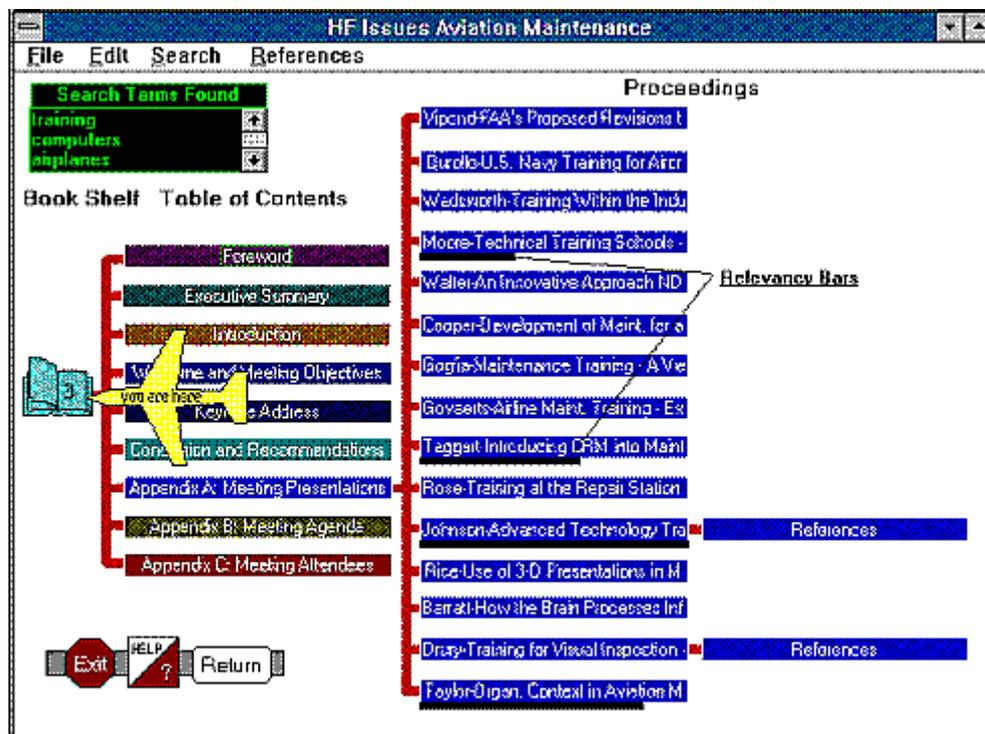


Figure 4.5 Search results are shown on the overview diagram by highlighting matching items

4.2.2.3 Search Aiding

When the [HIS](#) finds multiple matches to a query, these matches are presented to the reader for review. The [HIS](#) also aids the reader in determining which matches are likely to be most relevant to their query by providing relevancy bars along with the query results (see [Figure 4.5](#)). The magnitude of each relevancy bar indicates the likelihood of that document being most relevant; the longer that bar, the more relevant the document is likely to be to the reader.

Another search aiding technique is allowing the reader to refine and edit previous queries. Entering a search query is usually not a one-time process. Sometimes a query is too broad, in that too many pieces of information are retrieved for the reader to handle at a time. Sometimes, a query is too narrow, yielding little or no matches to the query. Readers can refine queries either by adding or deleting search terms. The [HIS](#) retains previous queries to aid the reader in determining which queries to expand.

The [HIS](#) also aids the reader by expanding and refining queries. To be implemented in future versions of the [HIS](#), a term thesaurus will be included to help the reader in locating additional terms that might be relevant to the desired information. These terms can be used by the reader as part of subsequent queries into the information.

4.2.3 System Architecture

The [HIS](#) operates on an IBM PC-compatible desktop platform with 4 Megabytes of system memory, a Microsoft Mouse (or compatible), and Microsoft Windows 3.1. Initial versions of the [HIS](#) can run off of the system's hard disk, but future versions will require a [CD-ROM](#) reader to use the system, due to the large amount of information contained in the system. The overall software architecture is shown in [Figure 4.6](#). The system was constructed using a variety of development tools, including: Microsoft Windows 3.1, Asymetrix Toolbook, Borland C++ Compiler, and Raima Corporation DB_Vista.

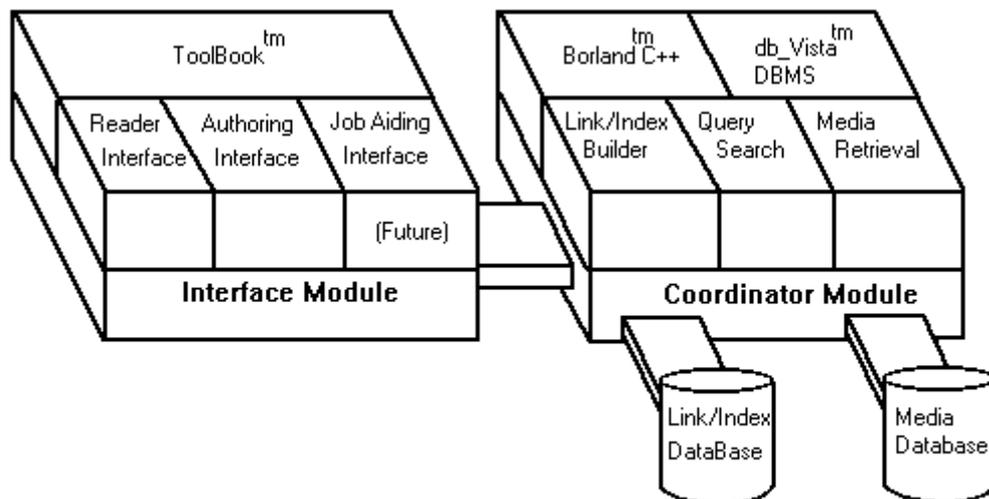


Figure 4.6 The Overall Architecture of HIS

These tools were used to construct two distinct environments. The first environment is the reader interface, which was described in earlier sections. The reader interface allows users of the information to search and browse the information contained in the [HIS](#). The second environment, the author interface, contains tools that allow authors to place information into the [HIS](#). These tools include editors, link builders, indexers, and database manipulation tools. Future versions of the [HIS](#) will allow access to these tools from the reader interface, thus allowing readers to store and index

information in addition to the material already contained in the [HIS](#).

The [HIS](#) contains two distinct databases. The first database, the Media Database, contains the actual information to be retrieved by the user. This material includes graphics, text, and eventually sound, animations, and other multimedia information. The second database, the Link/Index Database, contains information to help the [HIS](#) search for and retrieve information. The Link/Index database contains information regarding file location, search term location, document information, graphic information, and link structures.

4.3 CURRENT RESEARCH STATUS AND FUTURE PLANS

Phase I has produced a working prototype of [HIS](#). It includes an on-line version of the 3rd [FAA](#) Human Factors in Aviation Maintenance meeting. This prototype is available for distribution. Directions for future research include:

1. Provide one operation for indexing the [HIS](#) information. Presently, three distinct operations are required to store this information. By integrating these three steps into a common environment, these three operations can be performed simultaneously.
2. Improve navigational strategies to allow readers to return to information already seen. Similar to the idea of "bookmarks," this strategy will allow readers to quickly mark and return to important information.
3. Improve searching strategies by providing feedback to the users regarding the effectiveness of their searches. Aids such as on-line thesauruses, enhanced search control, and relevancy feedback can be used to improve searching by even the most novice user.
4. Integrate the authoring and reader interfaces. This integration will allow future users of the [HIS](#) to incorporate relevant information into the [HIS](#) system and will allow them to rapidly associate this information with material already stored in the [HIS](#).
5. Incorporate additional media types such as video, animations, and sound.

The technology provided by this research will be used to support the information retrieval needs across a variety of [FAA](#) and airline maintenance software support systems. Already, various training systems have benefitted from incorporating the [HIS](#) technology into these systems. Students are now able to quickly access information when it is needed during a training session. Current plans include the incorporation of the [HIS](#) technology into the various maintenance job aiding systems as well. Anticipated benefits of this marriage between hypermedia and job aiding include quicker access to information, exposure to material that might have been overlooked using conventional information access systems, and subsequent reduction of maintenance personnel error.

4.4 REFERENCES

- Duchastel, P. (1990). Examining cognitive processing in hypermedia usage. *Hypermedia* 2(3), 221-33.
- Frisse, M. E. & Cousins, S. B. (1989). Information retrieval from hypertext: Update on the dynamic medical handbook project. *Hypertext '89 Proceedings*, 199-212.
- Haan, B. J., Kahn, P., Riley, V. A., Coombs, J. H., & Meyrowitz, N. K. (1992, January). IRIS hypermedia services. *Communications of the ACM*, 35(1), 36-52.
- Johnson, W. B. & Norton, J. E. (1992). Integrated information for maintenance training, aiding, and on-line documentation. *Proceedings of the 36th Annual Meeting of the Human Factors Society*. Atlanta, GA: The Human Factors Society.
- Shepherd, W. T., Johnson, W. B., Drury, C. G., Taylor, J. C. & Berninger, D. (1991). [Human factors in aviation maintenance phase I: Progress report](#) (DOT/FAA/AA-91/16). Washington, DC:

U.S. Department of Transportation, Federal Aviation Administration.

Smith, J. B. & Smith, F. D. (1991). ABC: A hypermedia system for artifact-based collaboration. *Hypertext '91 Proceedings*, 179-192.

Spiro, R. (1990). Towards a general theory of hypertext: Nonlinear instruction for development of cognitive flexibility in ill-structured domains. In A. Oliveira (Ed.), *Structures of communication and intelligent helps to hypermedia courseware*. New York: Springer Verlag.

Chapter Five

Human Reliability in Aircraft Inspection

5.0 INTRODUCTION

This section describes the continuing work on aircraft inspection, whose long-term objective is to enhance system reliability through human factors interventions. It builds upon the [Phase I](#) outcomes reported in Shepherd, et al., 1991, and thus does not re-justify human factors applications in this field.

[Phase I](#) provided detailed Task Descriptions and Task Analyses of many aircraft inspection activities observed at major carriers in the U.S.A. During Phase II, visits were made to other inspection sites, with coverage of regional airlines, repair centers, and sites in the U.K. (see [Section 5.3.6](#)). The concentration was on specific aspects of the system, such as Non-Destructive Inspection (NDI), information flow, and training. Although inspection tasks were observed, no additional formal Task Analyses are reported here.

In Phase II, the implications of the data collected earlier have been researched in more detail than was provided in [Shepherd, et al., 1991](#). This has led to a series of studies by the research team, all under the objective of human factors interventions to improve inspection system reliability. These studies can be broadly classified into those with short-term and long-term outcomes. While the former have led to specific, on-going interventions at airline inspection sites, the latter have produced insights and on-going experiments in an off-site setting. One additional activity has been a joint project with the Civil Aeronautics Authority (CAA) in the U.K. to document and evaluate international differences in civil aircraft inspection (Drury and Lock, 1992).

[Chapter 3 of Shepherd, et al., 1991](#) listed a set of short-term and long-term research needs, and this list has provided the guidance for Phase II work. All of these needs were derived from a basic description of the inspection system, and a generic task description of inspection. As these descriptions form the basis of all that follows, an updated system description (from Drury and Lock, 1992) is included here.

5.1 THE INSPECTION SYSTEM: A HUMAN-FACTORS DESCRIPTION

An aircraft structure is designed to be used indefinitely provided that any defects arising over time are repaired correctly. Most structural components do not have a design life, but rely on periodic inspection and repair for their integrity. There are standard systems for ensuring structural safety (e.g., Goranson and Miller, 1989), but the one which most concerns us is that which uses engineering knowledge of defect types and their time histories to specify appropriate inspection intervals. The primary defects are cracks and corrosion (which can interact destructively at times) arising respectively from repeated stretching of the structure from aerodynamic or internal pressure loads, and from weathering or harmful chemicals. Known growth rates of both defect types allow the analyst to choose intervals for inspection at which the defects will be both visible and safe. Typically, more than one such inspection is called for between the visibility level and the safety level to ensure some redundancy in the inspection process. As the inspection system is a human/machine system, continuing airworthiness has been redefined by the design process from a mechanical engineering problem to an ergonomic one. Inspection, like maintenance in general, is regulated by the [FAA](#) in the U.S.A., the [CAA](#) in the U.K., and equivalent bodies in other countries. However, enforcement can only be of following procedures (e.g., hours of training and record-keeping to show that tasks have been completed), not of the effectiveness of each inspector. Inspection is also a

complex socio-technical system (Taylor, 1990), and as such, can be expected to exert stresses on the inspectors and on other organizational players (Drury, 1985).

Maintenance and inspection are scheduled on a regular basis for each aircraft, with the schedule eventually being translated into a set of job cards for the aircraft when it arrives at the maintenance site. Equipment which impedes access is removed (e.g., seats, galleys). The aircraft is cleaned, and access hatches are opened. Next comes a relatively heavy inspection load to determine any problems (cracks, corrosion, loose parts) which will need repair. During inspection, each of these inspection findings is written up as a Non-Routine Repair (NRR) item. After some [NRRs](#) are repaired, an inspector must approve or "buy back" these repairs. Thus, the workload of inspectors is very high when an aircraft arrives (often necessitating overtime working), decreases when initial inspection is complete, and slowly increases towards the end of the service (due to buybacks). Much of the inspection is carried out in the night shift, including routine inspections on the flightline, of aircraft between the last flight of the day and first flight of the next.

At a more detailed level, the task of inspection can be broken into a set of subtasks which follow in logical order. [Table 5.1](#) shows a generic task description based on simpler tasks for industrial inspection tasks (Drury, 1978). For each subtask, [Table 5.1](#) presents an example from both Visual Inspection and Non-Destructive Inspection (NDI). In a typical inspection schedule, well over 90% of the job cards are for Visual Inspection.

TASK DESCRIPTION	VISUAL EXAMPLE	NDT EXAMPLE
1. Initiate	Get workcard. Read and understand area to be covered.	Get workcard and eddy current equipment. Calibrate.
2. Access	Locate area on aircraft. Get into correct position.	Locate area on aircraft. Position self and equipment.
3. Search	Move eyes across area systematically. Stop if any indication.	Move probe over each rivet head. Stop if any indication.
4. Decision Making	Examine indication against remembered standards, e.g., for dishing or corrosion.	Re-probe while closely watching eddy current trace.
5. Respond	Mark defect. Write up repair sheet or if no defect, return to search.	Mark defect. Write up repair sheet or if no defect, return to search.
6. Repair	Drill out and replace rivet.	Drill out rivet. NDT on rivet hold. Drive out for oversize rivet.
7. Buy-back Inspect	Visually inspect marked area.	Visually inspect marked area.

Table 5.1 Generic Task Description of Incoming Inspection, with Examples from Visual and NDT Inspection

Table 5.1 Generic Task Description of Incoming Inspection, with Examples from Visual and [NDT](#) Inspection

With these seven task steps, the complex problems of error control, design of the information environment, and development of training schemes all become more manageable as specific human factors knowledge can be brought to bear on each task step in turn. The current review of projects shows this structure clearly, both in terms of deriving the needs for rapid interventions, and in developing off-line experiments to investigate the sensitivity of human performance to systems variables.

5.2 SHORT-TERM DEMONSTRATION PROJECTS

Although human factors engineering is becoming known to the aviation maintenance community through the [FAA/AAM](#) series of meetings, there is still a need to show straightforward, practical interventions which produce relatively rapid changes. Such demonstration projects can lead to widely disseminated changes, and to a model for how human factors studies can be conducted by

airlines themselves. Three projects were chosen by [FAA/AAM](#), of which two were to be pursued during Phase II, with the choice left to the airlines themselves. The three projects were on redesign of hard-copy workcards (job cards), design of the lighting environment for inspection, and redesign of the human interface of typical [NDI](#) equipment to follow human factors principles. The first two of these to be taken up by the industry were the workcards and lighting projects, so these are described in some detail in the following sections. Other projects, including the [NDI](#) interface design, are to be performed in future years and hence are described briefly.

To our knowledge one aircraft manufacturer and one airline company have started human factors groups in the maintenance/inspection field, but this still leaves many other airlines with a shortage of human factors expertise. Information is available through the proceedings of the [FAA/AAM](#) meetings on Human Factors in Aircraft Maintenance and Inspection, but it is often either human factors specialists telling what could be done, or existing industry personnel showing what has been done without formal human factors knowledge. With this background the short-term demonstration projects have been structured to allow human factors specialists and aircraft industry personnel to work together on projects which neither could conveniently perform alone. To this end, the [FAA/AAM](#) support has provided human factors expertise, while airline partners have provided facilities and personnel with detailed knowledge of inspection of particular aircraft. The airline partners have also agreed to provide travel to and from the work site. For their cooperation, airline partners get their personnel to understand some aspects of human factors, as well as a response to their specific needs. All partners have agreed to allow dissemination of study methodology and results.

As these are on-going projects, with the first two due for completion in May 1992, only the needs and methodology are presented here.

5.2.1 HUMAN FACTORS IN WORKCARD DESIGN

A major air carrier has agreed to become the partner on the workcard design project, working through maintenance facilities. Although the issue of information flow within the inspection/maintenance system is complex (see [Section 5.3.3](#)), and high-technology interventions are possible (Johnson, 1990), many airlines have too large an investment in current hardware to consider alternatives beyond hard-copy workcards as the inspectors' primary information. Airlines often have computer-generated workcards, and wish to continue using some version of the same medium, at least in the near-term. Thus, while we are moving towards new generations of computer-based job information aids, there is still an on-going need to apply human factors techniques to existing workcard generation systems.

The workcard controls the inspection workflow by describing to the inspector the location of the work area, the area(s) to be inspected, and the inspection procedure. It is the primary document that inspectors carry during inspection.

The task analyses of aircraft inspection (Drury, Prabhu and Gramopadhye, 1990) suggested that workcards are the main source of on-line feedforward information. However, even within the relatively homogeneous sample of air carriers, there was considerable variability in the design of these documents. Since the "paper document" is currently the prevalent and preferred means by which the inspector has access to the information that is needed on the job, the availability of quality documentation is of critical importance to inspection performance.

[Table 5.2](#) classifies the various human factors issues which the Task Analysis data showed to be relevant to documentation design. The workcard, which is a paper document, must be evaluated with these issues in mind. The taxonomy also provides a framework with which to design a new workcard which adheres to human factors principles.

1. Information layering	<ul style="list-style-type: none"> • Amount of information • Experts versus novices • Use of cues/indications/checklists 	<ul style="list-style-type: none"> • Levels of information • Accessibility • Flexibility of use
2. Layout of information	<ul style="list-style-type: none"> • Spatial layout • Grouping of information • Chunking 	
3. Presentation of Text	<ul style="list-style-type: none"> • Visual organization • Visual density • Letter case 	<ul style="list-style-type: none"> • Word spacing • Line spacing • Line length
4. Presentation of Graphics	<ul style="list-style-type: none"> • Visual organization • Spatial location w.r.t. text 	<ul style="list-style-type: none"> • Contrast • Labeling
5. Language Constraints	<ul style="list-style-type: none"> • Minimal number of words • Standardized nomenclature • Appropriate abbreviations 	<ul style="list-style-type: none"> • Fixed syntax • Concise wording
6. Physical Implementation	<ul style="list-style-type: none"> • Completeness of information • Consistency across cards 	<ul style="list-style-type: none"> • Physical accessibility • Accuracy of information

Table 5.2 A Taxonomy of Human Factors Issues in Workcard Design

Since the workcard is the means of communication of command information (both directive and feedforward), it is important to understand the effects of workcard design on the use of its information content by the inspectors. Current research in human factors and cognitive science in the areas of information processing, visual perception, learning, document design and computer display design (e.g., Wright, 1991) provide us theoretical, as well as empirical, guidelines that can be used for the design of more effective workcards. The taxonomy is an attempt to organize these guidelines to provide a framework that can direct the documentation design process.

Table 5.3 presents an analysis of the original Task Analysis data of aircraft inspection, classified using the above taxonomy. The points raised are not in any implied order of importance.

Information Layering	<ul style="list-style-type: none"> • None of the workcards provided layered information, i.e., the opportunity to access more detailed information on inspector selected points. • Key points were not clearly differentiated. • Little feedforward on problem areas to be expected on this aircraft at this time. • Few workcards provided checklist of probable/possible defects. • Necessary safety precautions not specified. • Task cards did not specify limits on wear, play, etc. for visual inspection to help the human inspector make more consistent judgements. • Cues indicating defects not listed in workcards, e.g. scuffed paint on fairings (of wing) indicates rubbing.
Layout of Information	<ul style="list-style-type: none"> • Poor legibility in some workcards. • Relevant information spread over multiple pages. • Much information of a legal or general nature occupied prime space at the top of the card. • Column layout, text boxes and other enhancements now available on computer-generated text were not used.
Presentation of Text	<ul style="list-style-type: none"> • Some cards were in all capitals, a violation of human factors principles. • Font design was not considered for legibility in highly variable lighting conditions. • Quality of printing and copying was not uniformly good.
Presentation of Graphics	<ul style="list-style-type: none"> • Workcards designed without considering correct location of graphics. • Some graphics were confusing even to experienced inspectors. • Some graphics for accessing the inspection area were ambiguous. • Most graphics were of poor quality. • Color coding was not used, but may need to be considered if it offers worthwhile performance improvements.
Language Constraints	<ul style="list-style-type: none"> • Procedures were not concisely worded in many workcards. • There seems to be no evidence of any conscious design procedure to use fixed syntax, consistent use of phrases or a standardized nomenclature.
Physical Implementation	<ul style="list-style-type: none"> • Imperfect matching of nomenclature for parts and defects between worksheet and secondary source material. • Inconsistent description of tasks. Some were described very briefly and others in detail. • No tool or aid that ensures that the inspector has covered the entire inspection area. • Drawings on workcard sometimes do not match configuration of the same area on the worksheet. • Illumination specifications not available/not specified on workcard. • Some workcards do not have figures showing relative location of various parts. • Some workcards did not specify equipment/gauges to be used in the inspection. • Completeness and currency of information is not assured, and not, therefore, trusted by all inspectors. • Physical size and shape of workcards is not always well integrated with other tools the inspector must carry and use.

Table 5.3 Classification of Observations from the Task Analyses of Inspection, Classified by the Taxonomy of Table 5.2

5.2.1.1 A Demonstration Program for Workcard Redesign

With our airline partner, a workcard redesign program is being undertaken as a demonstration of how human factors techniques can improve inspection. Existing workcards for a small number of relatively common maintenance events (an A-check and a C-check) are being analyzed with respect to the issues derived in the taxonomy. Good and poor aspects of the workcard design have been noted, both from analysis of the workcard itself and from analysis of its use by inspectors. From this data collection phase will come a series of design requirements which, if met, will ensure good human factors design.

With airline partner representatives, design solutions will be developed to cover both short-term and long-term changes. Short-term interventions for workcards may include, for example:

1. Changing the presentation format and layout to improve ease of use and legibility.
2. Ensuring that visual material is incorporated into the worksheet.
3. Consistent naming of parts, directions, defects, and indications between all documents used by inspectors.
4. Multi-level workcard systems, useable by inspectors with different levels of immediate familiarity with the worksheet content.
5. A better physical integration between the workcard and the inspector's other documents and tools needed at the worksite.
6. Providing a better spatial integration between the workcard and the inspection tasks

around the aircraft.

Each design solution will be implemented and a series of prototype workcards produced. These will be pre-tested by having inspectors use them while providing a verbal protocol of their actions. From this user evaluation will come a refined design.

The final design will be tested against the current design using controlled tests during A-checks and C-checks. Measurements will be taken of inspector verbal protocols, errors/confusions observed, and questionnaire evaluation from both inspectors and supervisors.

The results will be documented as a case-study to show:

- a. How other maintenance/inspection operations can improve their workcards.
- b. How to apply human factors principles to the improvement of other maintenance/inspection functions.

5.2.2 DESIGNING THE VISUAL ENVIRONMENT FOR INSPECTION

A second major carrier is cooperating with the University at Buffalo team to improve the inspector's visual environment. This project is based at the maintenance facilities operated by the carrier at a single airport. There is a single maintenance hangar, with three aircraft bays, and apron areas outside the hangar and by the gates. The main concentration will be on in-hangar activities, but other sites will also be considered. Having a single hangar makes the demonstration project manageable while still providing a representative application of human factors.

Analysis of aircraft inspection activities has shown that visual inspection dominates other inspection activities (Drury, Prabhu, and Gramopadhye, 1990). Since visual inspection is such an important component, accounting for almost 90% of all inspection activities, it is imperative that the task be performed in the most suitable work environment. From the task analysis of various inspection tasks in [Table 5.1](#), it is seen that "visual search" is an important component of the inspection task, and the success of this stage is critical for successful completion of the inspection task. In visual search the inspector must closely examine each area for a list of potential faults. The amount of effort required on the part of the inspector for each area depends upon various factors such as the prior information (from training experience on the workcard) and the suitability of the physical conditions for inspections (lighting, illumination levels, etc.).

Studies in aircraft inspection have shown that poor illumination, glare, and other adverse lighting conditions could be the single most important reason for "eye strain" or visual fatigue. Visual fatigue results in deterioration in the efficiency of human performance during prolonged work. Progressively more effort is required to maintain performance, and eventually performance level decreases despite the extra effort. The purpose of this study is to identify potential sources of improvement in inspection lighting and to suggest modifications so that the task can be performed under improved visual conditions.

From the detailed Task Analyses of numerous inspection activities performed in Phase I, [Table 5.4](#) gives a list of examples of poor human factors design. Each represents an opportunity for intervention to improve the human/system fit and hence, increase job performance with decreased work stress.

1.	Illumination levels beneath the aircraft varied depending on the location. Measured levels during daylight ranged from 25 ft. candles under wings and fuselage areas on the side of the aircraft facing open hangar doors, to 2-5 ft. candles under the wings and fuselage on the opposite side of the aircraft.
2.	Lighting levels beneath the aircraft were adequate only for gross visual inspection.
3.	Inspectors often used a flashlight as an aid during visual inspection. It should be noted, however, that the type of flashlight used was not consistent among all inspectors even within a single aircraft operator, and varied considerably between carriers.
4.	Often it was found that the overhead lighting used for general illumination was covered with dirt and paint, inhibiting full illumination capability.
5.	Aircraft near the hangar doors were exposed to higher illumination. Illumination ranged from 16 ft. candles to 114 ft. candles.
6.	Lighting conditions were different during the day shift and night shift. During the night shift illumination ranged from 29-33 ft. candles.
7.	Supplemental lighting was often not provided under the aircraft. Measured levels in these areas ranged from a day time high of 42 ft. candles to a night time low of 1.2 ft. candles.
8.	The general lighting level inside the aircraft fuselage averaged between 1.5 ft. candles to 3 ft. candles.
9.	Glare from open hangar doors within the inspector's visual field was apparent.
10.	When using supplemental lighting (flashlight, helmet light, portable fluorescent) the illumination on the aircraft surface was increased, but glare was caused when searching inside the aircraft structure.
11.	Portability of supplemental lighting was often poorly designed, and not well integrated into the inspector's kit.

Table 5.4 Observations on Visual Environment of Inspection from Task Analysis Data

In designing lighting systems, the following factors need to be considered:

Recommended Light Levels for Different Tasks

The recommended illumination depends upon the type of task and whether the visual task is of high or low contrast. The Illuminating Engineering Society (IES, 1984) recommends that surface areas requiring visual inspection be provided with 75-100 ft. candles (800-1050 lux) of illumination. Vision can be improved by increasing the lighting level, but only up to a point, because the law of diminishing return operates (e.g., IES Lighting Handbook, New York, 1984). Increased illumination could also 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 an organization.

Selection of Light Sources for Color Rendering

In the selection of artificial light sources one of the most important considerations is color rendering, i.e., the degree to which the perceived colors of an object illuminated by various light sources match the perceived colors of the same object when illuminated by a standard light source. Color rendering could be important, because often "change in color" of true sheet metal is used as a clue to indicate corrosion.

Direct and Indirect Lighting: Glare

The quality of illumination can be improved by reducing glare. Direct glare 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 off a white object such as the workcard can cause glare from surrounding surfaces. Glare can be reduced by resorting to indirect lighting. Of particular concern is that in inspecting partially-hidden areas (e.g., inside access panels) the lighting used to illuminate the defect may cause glare from surrounding surfaces. Carefully designed combinations of general area lighting, portable area task lighting and localized spotlighting need to be produced.

Specialized Lighting

During visual inspection of an aircraft structure the inspector is looking for multiple defects, such as corrosion, ripples, hairline cracks, dents, missing rivets, damaged rivets (e.g., "pooched", "dished" rivets), and rivet cracks.

It is possible that not one single lighting system is suitable for detecting all defects. Therefore, the use of a specialized lighting system for each class of defects may be necessary. However, the use of special light systems has one major drawback. It implies that the area must be examined for each class of defects sequentially rather than simultaneously, which could involve time and expense. A typical example is the difference between general illumination and the grazing illumination provided by special purpose lighting. The diffused nature of general illumination tends to wash out the shadows while the surface grazing light relies upon showing shadows to emphasize objects that project above or below the surface. Task visibility for surface topography is distinctly better with grazing light, whereas color changes or corrosion may be better seen under general illumination. An example of surface topography is the inspection of the fuselage for ripples. Ripples are easy to detect using surface-grazing light but general illumination tends to wash them out. However, strong side lighting may mask important color differences.

Design Requirements for Lighting

Studies of visual search have 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 the size of the defect, defect/background contrast, and lighting intensity (see [Section 5.3.3](#)).

Lighting design has a clear impact upon the final two variables, but it has broader requirements to fulfill as visual inspection involves more than visual search. Lighting should be designed such that the following tasks can all be performed satisfactorily and preferably optimally:

1. Inspection (visual search) of the aircraft fuselage for defects.
2. Reading the workcard/instructions.
3. Movement around the aircraft (using the scaffolding, or equipment, e.g., cherry picker).

In addition, special purpose lighting should not interfere with any other parallel task in progress. In designing the visual environment, 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.

Since inspectors have to move to different areas on the aircraft during a single task and all areas may not be accessible to generalized lighting from a static source, generalized lighting may be augmented from a combination of static portable sources, and then further augmented, if necessary, using flashlights.

It is proposed to use the Task Analyses performed so far and lighting surveys of the inspection work areas to determine the design requirements for lighting in detail. The market will then be surveyed for available solutions (e.g., area lights, flashlights, headlights, stand lights) to choose a small number of promising systems. On-site human factors evaluations of these lighting systems will be performed to determine which, if any, improves visibility of defects or other indications to inspectors, while maintaining portability.

The specific steps to be undertaken for this project are:

1. Site visit and task analysis to determine specific visual requirements and lighting requirements of tasks, and the current visual environment. Luminance and illuminance will be measured throughout the hangar to determine consistency and adequacy. A checklist of visual factors (from Drury, 1990a) will be used to assess the adequacy for the specific tasks performed.
2. Survey market for available solutions to identify promising systems for illumination,

diffusion and specialized lighting.

3. On-site human factors evaluation of selected lighting system to demonstrate advantages. This will include performance evaluation (speed, accuracy) as well as operator acceptability and cost.
4. Produce a set of design recommendations which can be used as the basis for future lighting design.

5.2.3 FUTURE SHORT-TERM PROJECTS

The research will continue to focus on long- as well as short-term projects. These short-term, immediate payoff projects will study such topics as using portable computers for development and evaluation of multi-level job cards. Such studies will involve airline and/or manufacturer participation to ensure that research results and by-products can readily be transitioned into aircraft maintenance work environments.

5.3 LONG-TERM RESEARCH PROJECTS

From Phase I came a wealth of Task Analysis data and descriptions of specific inspection and maintenance organizations at many carriers. In addition, information from the [FAA/OAM](#) meetings, reports and visits to aircraft manufacturers and specialized equipment suppliers, gave a clear description of the inspection and maintenance system to the human factors engineers involved. The [Phase I report](#) (Shepherd, et al., 1991) made a first attempt to merge this data with existing and current concepts in human factors. An obvious need was to perform this integration at a deeper level to guide the long-term human factors needs of the aviation maintenance industry. During Phase II, this step was undertaken; system demands were interpreted in terms of known human capabilities and limitations.

The first fruits of this process were four reports which covered a framework for human reliability in this field, a detailed examination of the information environment, an analysis of the effects of time on inspection (especially the speed/accuracy tradeoff), and a study of the improvement of training for visual inspection. These reports are listed with other publications at the end of this section. The findings of each report are summarized in [Sections 5.3.1](#), [5.3.2](#), [5.3.3](#), [5.3.4](#), and [5.3.5](#), augmented where necessary by off-line experiments. Additionally, a joint venture between the [FAA](#) and the [CAA](#) on inspection is presented as [Section 5.3.6](#).

5.3.1 COMPUTER-BASED INSPECTION EXPERIMENTS

It became apparent that the traditional experimental work in aviation inspection was not always the best way to perform human factors evaluations. Studies of crack detection probabilities (ref) have been large, costly, and complex, but have not addressed many of the human factors issues beyond the psychophysics of [NDI](#) equipment. Factors such as training method, information environment, and time pressure have not been systematically considered. Thus, the need was recognized for a low-cost but realistic simulator for aircraft inspection. Its purpose is not to provide a point estimate of the probability of detection of a given crack, but rather to determine how inspection performance is affected by manipulable human factors such as those above. There is sufficient knowledge of models of human inspectors (e.g., review by Drury, 1991) to be able to determine which aspects of the real task to retain if a simulator is to be "realistic".

Two simulation programs were implemented on a SUN Sparc station 1 workstation, one for an [NDI](#) task (eddy current inspection of rivets) and the other for a visual task (visual inspection of rivets and sheet metal). These programs are discussed below.

5.3.1.1 [NDI \(Eddy-Current\) Inspection Program](#)

The inspection task consists of inspecting rows of fuselage rivets for cracks using an eddy-current probe. The simulator display consists of four windows ([Figure 5.1](#)) as follows:

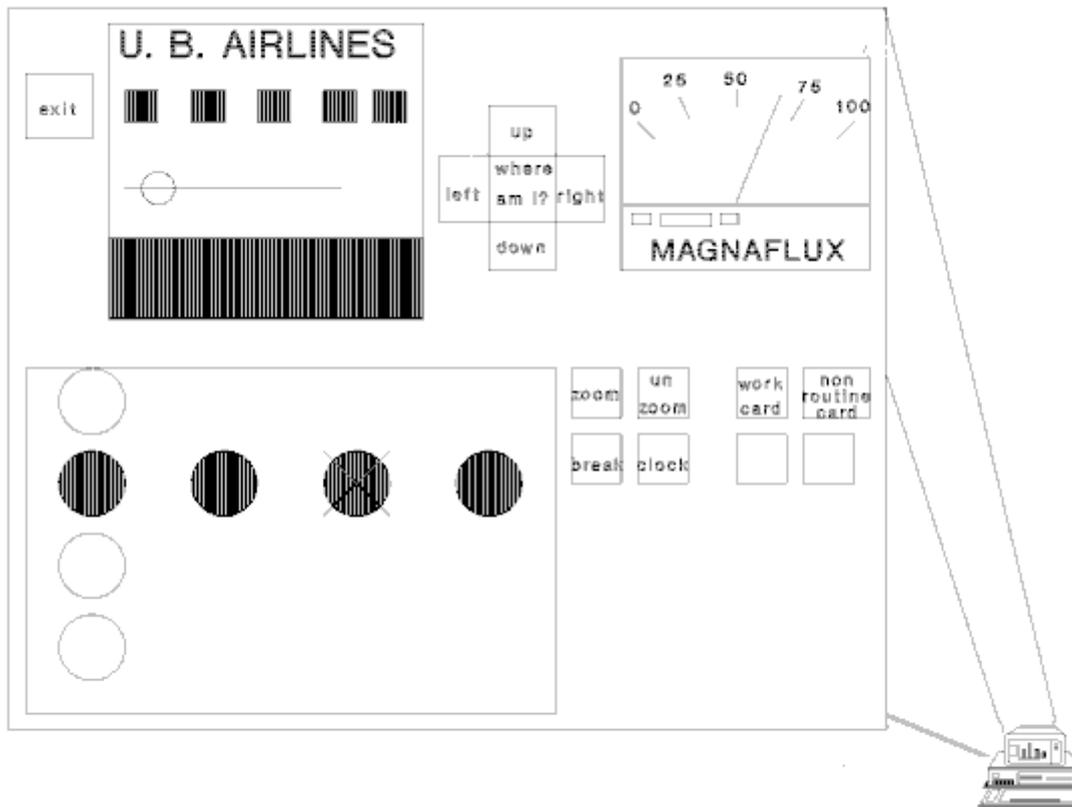


Figure 5.1 NDI Inspection Task Simulation

Inspection Window. This window displays the rivets to be inspected. Six rivets per row are displayed at a time. The simulation program has the capability to display multiple rivet rows at a time. During the training session a circle is placed around each rivet to help the subjects in defining the optimal probe path around the rivet for defect detection. On the upper right hand corner of this window there is an indicator that is green when the subject is in the inspection mode. During this mode, the subject is able to inspect and classify (defective/non-defective) the rivets, but has no access to any of the functions outside the current window. To obtain access to these functions, the subject has to click the left mouse button near any of the rivets. This results in a circular marker being placed around that rivet and the inspection indicator light turns white, indicating the inspection mode is switched off.

Macro-View and Directionals. The macro-view in the upper left window allows the subject to have a view of the total inspection area and its relation to the aircraft fuselage. Thus, for a 400 rivet inspection task, while only six rivets are seen in the inspection window, the entire 400 rivets are marked (on a smaller scale) in the macro-view. A click on the where-am-I button places a circle around the area of the macro-view currently in the inspection window. Thus, the subject is able to determine where he/she is at any point in time with relation to the entire task.

The directionals consists of four square areas marked left, up, right and down (L/U/R/D, clockwise). Clicking the left mouse button on any one of these areas shifts the view (scrolls) in the inspection window in the indicated direction.

Eddy Current Meter. The defect indication is displayed on the meter indicator in the upper right window of the monitor screen. The meter has a fixed scale with divisions marked from 0 to 100, and a moving indicator. A red marker is provided that can be set by the subject at any point on the scale. The deflection of the needle (from its resting position at zero) beyond this set point (default = 60) produces an auditory alarm as well as a red flash of the indicator light at the apex of the meter.

The point of the needle is deflected if any of the following happen:

1. The mouse cursor is moved over a crack on the rivet (the cracks themselves are not visible).
2. The mouse cursor is moved over a grey spot (indicating corrosion, or dent; randomly placed across rivets).
3. The mouse cursor is very close to, or moved over, the rivet head itself.

Subjects are instructed that if the deflection is greater than 60% and they judge it to be from a crack, then the rivet should be marked bad.

Lower Right Window. This area contains functional (dialogue) buttons. Activation of the zoom button allows the subject to take a closer look at the current rivet to be inspected. The zoom is incremental and magnifies the area to twice its original size (within the inspection window) at every click. A mouse click on the unzoom area restores the inspection window to its original condition. Clicking on the "break" area stops all clocks and covers the inspection window to allow the subject to take breaks. Clicking on the "clock" area displays the time elapsed in the task. The other functional buttons includes "display non-routine card," "display workcard," and "turn rivet numbers on/off."

The program also has the facility for recording the subject's assessment of workload using the Pearson Feeling Tone Checklist and the Modified Cooper-Harper Scale. These two scales appear for response at the end of pre-set intervals.

5.3.1.2 Visual Inspection Program

To simulate visual inspection, the SUN Sparc station 1 is used with a program having similar logic and displays to the [NDI](#) program. The major differences are that detection is visual, and that the eddy-current meter is obviously absent. In this task the inspector searches for multiple defect types and classifies them into different severity categories. The various fault types with their descriptions are:

1. Missing Rivet: A rivet missing from the rivet hole.
2. Damaged Rivet: Part or all of the rivet head is damaged resulting in jagged edges.
3. Pooched/Dished Rivets: Rivets with a center which appears raised or sunken.
4. Loose Rivets: Rivets running loose in the rivet holes.
5. Rivet Cracks: Cracks which originate at the edges of the rivets and propagate upwards and outwards.
6. Dents: Sheet metal damage in the aircraft fuselage represented by sunken areas.
7. Corrosion: Damage to sheet metal surface represented by patches of discolored or raised skin.

Depending upon the severity of the defect type, the defects can be classified into critical and noncritical defects.

The layout of the multi-window simulated inspection task is shown in [Figure 5.2](#). The function of each window is as follows:

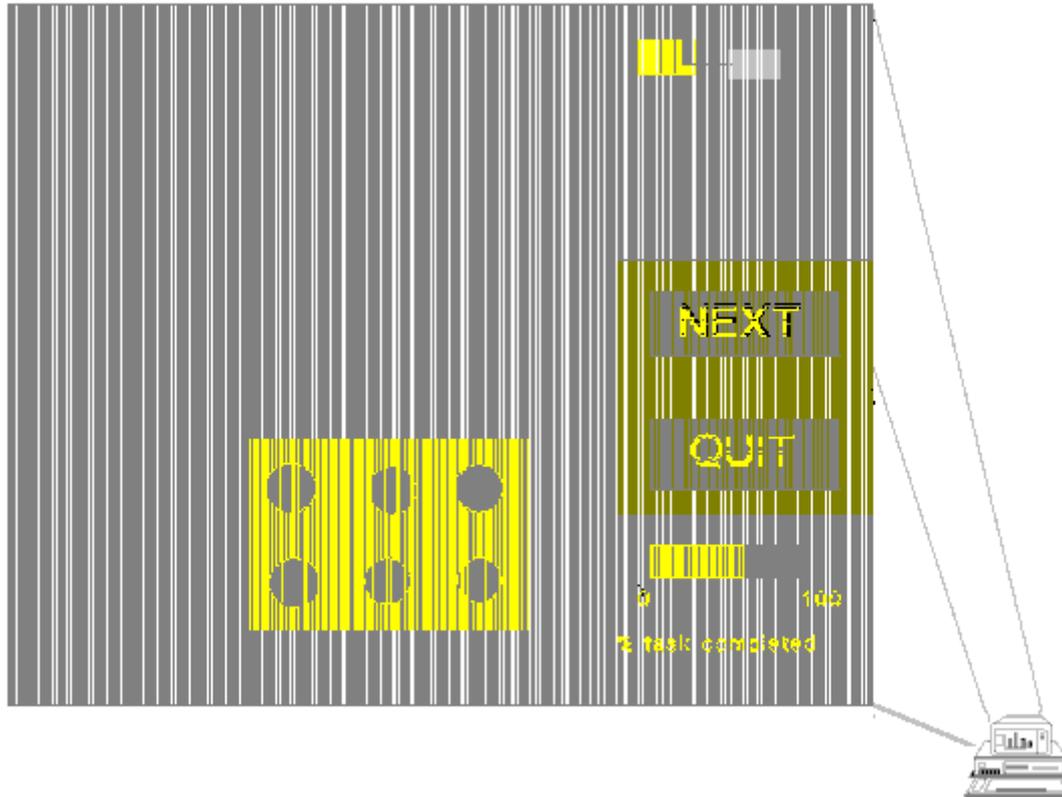


Figure 5.2 Layout of the Simulated Inspection Task

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, faults can be seen and responded to by clicking them using the mouse button. The entire area of the inspection window can be viewed by successive movements of the smaller illuminated window.

Search Monitor Window. This is a monitoring device which helps the inspector keep track of the window movement in the inspection window. It provides the inspector feedback as to the:

1. Point of previous fixation
2. The sequence (pattern) adopted by the inspector
3. The area covered (viewed through the window) up to the current time.

The small illuminated window in the inspection window is represented by a tile in the search monitoring window. As the window is moved, so does the tile. The tile has a different color from its illuminated background area. The background color changes to the color of the tile as the tile passes over it, indicating that the corresponding areas have been fixated (covered by the window). The darkest shade of the tile is the point of previous fixation. The sequence is given by the shade of the color--lighter shades indicate earlier fixations in sequence while darker shades indicate later fixations.

Macro View Window. This window represents the entire task to be inspected, and can be looked upon as the global coordinate referencing system. Thus, it provides information to the inspector as to his current position with reference to the entire task.

5.3.2 A FRAMEWORK FOR HUMAN RELIABILITY IN AIRCRAFT INSPECTION

Maintaining civil aircraft worthiness requires the reliability of a complex, socio-technical system. This system's reliability is dependent on the reliability of its components (i.e., equipment, inspectors, the physical environment), and on how reliably these components interact. Most errors in aircraft inspection and maintenance can ultimately be attributed, at some level, to a human-system mismatch. Operators may cause errors outright, or more likely, human information processing limitations and characteristics may be "catalytic" factors (Rouse and Rouse, 1983), combining with other component characters to evolve "sneak paths" (Rasmussen, 1982) to error situations.

The assessment of human error in complex systems is currently undergoing somewhat of a renaissance (Brown and Groeger, 1990). Classification schemes of errors have expanded from the early "omission/commission" classification (Swain and Guttman, 1983 and Meister, 1971) to more behavior-based classifications (e.g., Norman, 1981; Rasmussen, 1982; Rouse and Rouse, 1983, and Reason, 1990). While error classifications based on task characteristics may provide a convenient descriptive format for errors, error models based on human behavior can define causal mechanisms of errors. Identification of causal mechanisms and catalytic factors is necessary for predicting errors and thereby designing error tolerant systems. The approach taken here is to use a behavior-based and system-based human error classification scheme to identify, predict, prevent or reduce, and report errors in aircraft inspection and maintenance.

This section focuses on describing a methodology to accomplish these goals. [Section 5.3.2](#) provides more detail in defining information flow, and deriving information requirements which will prevent or mitigate the effects of information flow-related errors. Both this section and [Section 5.3.3](#) are responses to [FAA](#) project activities and exist elsewhere as two self-contained separate reports (Latorella and Drury, 1991, and Drury and Prabhu, 1991, respectively). These reports have been considerably abbreviated for their presentation in this report. Both efforts use Rasmussen's (1986) cognitive control levels and Rasmussen and Vicente's (1989) systemic error mechanisms extensively as a conceptual foundation. Both efforts also begin with Drury's (1991) Failure Modes and Effects Analysis (FMEA) of errors in aircraft inspection and maintenance. These concepts and the FMEA are presented only in the first section to avoid redundancy. As a result, some of the material presented in [Section 5.3.3](#) is dependent on the theoretical and data analysis foundation described in this section.

[5.3.2.1 Approaches to Human Error](#)

[5.3.2.1.1 Quantitative Approaches](#)

Early efforts to incorporate human performance in the evaluation of system reliability spawned the field of Human Reliability Assessment (HRA). These methods attempt to assess human reliability with the same techniques used to assess equipment reliability (Meister, 1971). They seek to: (1) develop extensive databases of human reliability data for elemental tasks, (2) provide a method for combining these estimates to generate a measure of human reliability within the system, (3) use this measure of human reliability directly, as the reliability of the human as a system component, in evaluations of total system reliability by Probabilistic Risk Assessment (PRA). Early [HRA](#) methods are criticized for their overly-structured, and hence cumbersome, representations of the human's involvement in systems. [HRA](#) methods are also criticized for their inability to adequately represent the behavioral mechanisms of human errors and hence for their inability to prescribe, rather than merely describe, systems in terms of their propensity for error situations. Quantitative human error assessment techniques include decompositional probabilistic methods (e.g., Fault Trees, Event Trees, Failure Modes and Effects Analysis), classical reliability theory based on Markov modeling, stochastic simulation modeling, and a variety of other techniques (e.g., HCR, TESEO, SLIM-MAUD). These approaches are each described and critiqued in Latorella and Drury (1991). Lock and Strutt (1985) have investigated quantitative human error modeling in the aircraft inspection and maintenance context.

5.3.2.1.2 Qualitative Approaches

Several researchers have arrived at behavior-based classification schemes for human errors. Those of Norman, Hollnagel, Rasmussen, and Rouse and Rouse are described below. Elements of these schemes have been in approaches to managing errors in aircraft maintenance and inspection.

Norman. Norman (1980, 1981) classifies human error into two fundamental categories: slips and mistakes. Slips result from automated behavior when the intention, the goal, is correct but some aspect of the execution is flawed. Mistakes, in contrast, are the result of flawed cognitive processes, such as formation of the wrong goal. Slips are usually minor errors and are often evident and corrected by the perpetrator. Mistakes, however, are more serious errors, and are sometimes opaque to the perpetrator. Mistakes are therefore usually difficult to observe and recover. Slips are partially due to limitations in attention and therefore are more likely to occur in distracting, time-sharing, boring, or stressful situations. Norman identifies six types of slips: capture errors, description errors, data driven errors, associative activation errors, loss of activation errors, and mode errors. Descriptions of these types of slips and examples related to aircraft inspection and maintenance can be found in the original report (Latorella and Drury, 1991). Norman's (1981) classification is intuitively appealing and useful for describing errors. However, the slip/mistake classification is not detailed enough to describe what specific aspects of human information processing generate errors.

Hollnagel. Hollnagel (1989) introduces the conceptual distinction between error phenotypes and error genotypes. Error phenotypes are observable states which are deemed undesirable. Error genotypes are the generative mechanisms of these observable states. Error phenotypes are manifestations of error genotypes expressed in a particular environment. While Hollnagel allows that combining genotypes and phenotypes provides a more complete psychological description of human error, he holds an empiricist's view for the purpose of system design: in order to automate error detection, errors can only be expressed in terms of phenotypes. He therefore proposes a taxonomy to operationalize phenotypes, describe complex phenotypes (combinations of simple phenotypes), and to provide a basis for a computer program which detects error situations. Hollnagel's distinction between error phenotypes and error genotypes is important and is used in the development of this paper's approach to managing aircraft inspection and maintenance errors.

Rasmussen. Rasmussen has contributed to [HRA](#) in two veins: he has developed models of human performance in an effort to identify fundamental causes of human error, and he has related and defined the importance of qualitative human error modeling to system reliability. Rasmussen departs from the more traditional approaches in his conceptualization of human error. He does not rely on the constrained definition of human error presented in most [HRA](#) techniques, rather he states that what is human error is defined by not only the human, but by system and operational tolerances (Rasmussen, 1982). Rasmussen also argues that human errors defined by the outcome of events should not necessarily be attributed to a human having performed incorrectly. For example, should an error resulting from a new situation be attributed to the human? If an error provides feedback about the system without compromising system functioning, should it still be considered something to avoid? Rasmussen also defines stipulations for collecting [HRA](#) error rate probabilities and states the case for qualitative error modeling to aid [HRA](#) in ways that error rates can not, such as prediction and corrections of errors, especially of low probability, high impact "sneak paths". Rasmussen (1982) developed a classification of human error towards this end.

The skill-rule-knowledge (SRK) framework proposed by Rasmussen (1986) classifies human behavior into three categories of ascending complexity: skill-based behavior, rule-based behavior and knowledge-based behavior. Any decision is made at the lowest level possible, with progression to higher levels only when a lower level fails to reach a decision.

Skill-based behavior represents psychomotor behavior without conscious control, consisting of automated routines that are driven by sensory data received as "signals" from the environment (Rasmussen, 1986). Signals represent information that is a quantitative indicator of the temporal and spatial aspects of the environment, and may trigger skill-based behavior by activating the automated

behavioral routines of the human. Skill-based behavior is normally based on feedforward control and proceeds without conscious attention.

From the aircraft inspection viewpoint, the movement of the pencil probe around a rivet or a sliding probe along a stringer (a row of rivets) during, for example, an eddy current inspection or an ultrasonic inspection, represents skill-based sensorimotor performance involving some amount of feedback control. Similarly, the pre-attentive phase of visual search, as well as the extra-foveal process in extended visual search can be considered to be skill-based behaviors that are data driven and based on feedforward control.

Rule-based behavior represents consciously controlled, goal-oriented behavior guided by rules or procedures for action. These rules are stored patterns of behavior that have been empirically derived during previous occasions or communicated as instructions from an external source (Rasmussen, 1986). Information during rule-based performance is perceived as "signs" which represent information that activates or modifies the rules and depicts situations or environmental features along with the conditions to act (Rasmussen, 1986). Rule-based behavior proceeds towards a goal, utilizing feedforward control through rules and without demanding any deeper reasoning on the part of the human.

In aircraft inspection, an experienced inspector interpreting the deflection of the ultra-sonic meter, or the pattern traced on an oscilloscope during eddy current testing, can be assumed to be indulging in a rule-based behavior if the "signs" are familiar. Similarly, the extra-foveal process in search where cues on the periphery guide the next fixation can be considered a rule-based behavior. Rule-based search can also result from information gathered in the foveal component, for example bulging of aircraft skin triggers search for corrosion. Pre-determined search strategies, as a result of past experience, training, or work card instructions, can also lead to a rule-based behavior.

Knowledge-based behavior represents goal-controlled, problem-solving performance in unfamiliar situations. It requires a functional understanding of the system, analysis of the current state, and response of the environment based on conscious, advanced reasoning while utilizing feedback control for error correction (Rasmussen, 1986). During knowledge-based behavior, the human perceives information as "symbols", i.e., concepts about the functional aspects of the environment which refer to an internal representation that can be used by the human for reasoning (Rasmussen, 1986).

In aircraft inspection, knowledge-based behavior can occur in [NDI](#), for example during eddy current testing of rivets, when the inspector sees a curve traced on the oscilloscope screen of a shape never encountered before. In this case the inspector has to use the knowledge of eddy current technology, knowledge about the instrument, knowledge about the aircraft, etc., to interpret whether the signal represents a crack or not. Along similar lines, the use of cues to detect visual defects needs active reasoning (knowledge-based behavior) until the association of the cue to the defect is confirmed, in which case the cue will trigger rule-based behavior.

Rasmussen (1982) provides a framework for classifying causes of human error as a function of situational and task characteristics and the error phenomenon. Basic error mechanisms are derived through the use of a human information processing model, linking human decision-making and responses to internal processes. His model can be used to describe human behavior over the three levels of cognitive control, and can be used to indicate decision aiding devices and training needs at these different levels. He specifically mentions that systems must be designed with interlocks and barriers where it is unreasonable to expect operators not to err and that systems should allow errors to be observed and reversed. A related work (Rasmussen and Vicente, 1989) identifies four systemic error mechanism categories: (1) effects of learning and adaptation, (2) interference among competing control structures, (3) lack of resources, and (4) stochastic variability of individuals. Rasmussen and Vicente (1989) describe examples of errors within these categories and cognitive control levels (see [Table 5.5](#)). Similarly, Drury and Prabhu (1991) used the cognitive control classification to organize error shaping factors (see [Table 5.6](#)).

EFFECTS OF LEARNING AND ADAPTATION:	
• <i>Knowledge-based:</i>	Search for information and hypothesis testings in novel situations may lead to acts which are judged as errors after the fact;
• <i>Rule-based:</i>	The law of least effort may lead to underspecified cues;
• <i>Skill-based:</i>	Optimization of motor skill needs feedback from boundaries of acceptable performance (speed-accuracy tradeoff);
INTERFERENCE AMONG COMPETING CONTROL STRUCTURES:	
• <i>Knowledge-based:</i>	False analogies; interference in means-end hierarchy;
• <i>Rule-based:</i>	Functional fixation; adherence to familiar rules;
• <i>Skill-based:</i>	Capture by frequently used motor schemata;
LACK OF RESOURCES:	
• <i>Knowledge-based:</i>	Limitations of linear reasoning in causal networks; insufficient knowledge, time, force, etc.;
• <i>Rule-based:</i>	Inadequate memory for rules;
• <i>Skill-based:</i>	Lack of speed, precision, force;
STOCHASTIC VARIABILITY:	
• <i>Knowledge-based:</i>	Slips of memory in mental models;
• <i>Rule-based:</i>	Erroneous recall of data or parameters related to rules;
• <i>Skill-based:</i>	Variability of attention; variability of motor parameters, motor noise (variation in force, precision of movements);

Table 5.5 Potential Errors Described by Level of Cognitive Control and Systemic Error Mechanisms (Rasmussen and Vicente, 1989)

KNOWLEDGE-BASED ERROR SHAPING FACTORS	
Information Overload Incomplete Knowledge Delayed Feedback Bounded Rationality Memory Cueing Insufficient Consideration of Process Causal Series Vs. Nets Attentional Limitations Confirmation Bias Vagabonding Overconfidence Memory Slip Selectivity Biased Reviewing Illusory Correlation Lack of Resources Complexity Problems	<ul style="list-style-type: none"> ▪ Excessive demands on memory. ▪ Knowledge lacking on system dynamics, parameters, side effects, etc. ▪ Delays in feedback concerning the execution of decisions. ▪ Limited cognitive capacity relative to problem size leads to satisficing behavior. ▪ Familiar information in LTM is cued by problem content. ▪ Past history disregarded when handling dynamic current events. ▪ Oversimplification of causality, thinking in terms of immediate goals. ▪ Finite resources of attentional processes. ▪ Tendency to maintain current hypothesis facing contradictory evidence. ▪ Moving amongst issues without detailed consideration of any one. ▪ Excessive belief in the correctness of one's own knowledge. ▪ Memory failure due to faulty activation of schemata. ▪ Selectively process information. Attention to wrong features of task can result. ▪ Error in reviewing planned course of action. ▪ Failure to detect correlation or understand the logic of covariation. ▪ Lack of mental capacity for causal reasoning, insufficient knowledge. ▪ Problems in understanding system due to its complex nature.
RULE-BASED ERROR SHAPING FACTORS	
Availability Countersigns Rigidity Encoding Deficiency Inadvisable Rules Wrong Rules Inelegant Rules First Exceptions	<ul style="list-style-type: none"> ▪ Tendency to use intuitive rules or use rules that readily come to mind. ▪ These are input indications that the more general rule is inapplicable. ▪ Mindset (cognitive conservatism) results in refusal to change familiar procedure. ▪ Encode inaccurately or fail to encode the properties of the problem space. ▪ Rules that satisfy immediate goals but can cause errors due to side effects. ▪ Rules that are wrong for the current situation. ▪ Rules that achieve the goal but are inefficient. ▪ Errors caused on first occasion that is an exception to the general rule.
SKILL-BASED ERROR SHAPING FACTORS	
Omissions Perceptual Confusion SATO Motor Schemata Capture Stochastic Variability Reduced Intentionality Repetitions Reversals Interference	<ul style="list-style-type: none"> ▪ Omission of actions/action sequences needed to achieve a specified goal. ▪ A familiar match is accepted instead of the correct match. ▪ Speed accuracy tradeoff. ▪ Focussed attention absence leads to takeover by frequently used schemata. ▪ Variability in control of movements. ▪ Delay between intention and action. ▪ Actions that are unnecessarily repeated. ▪ Unintentionally reversing an action just committed. ▪ Potential problems stemming from concurrent activities.

Table 5.6 Definitions of Error Shaping Factors (Prabhu, Sharit and Drury, 1992)

Rouse and Rouse. Rouse and Rouse (1983) propose a behavioral classification scheme for human errors which borrows heavily from Rasmussen's contributions. They attempt to analyze human error in terms of causes, as well as contributing factors and events. Their scheme organizes human errors around Rasmussen's flow model (1976) of an operator's information processing task. This model

gives the following steps in task performance:

1. Observation of system state
2. Choice of hypothesis
3. Testing of hypothesis
4. Choice of goal
5. Choice of procedure
6. Execution of procedure.

This classification scheme has been used to record and analyze human errors in several contexts: (1) detection, diagnosis, and compensation of engine room failures in a supertanker (van Eckhout and Rouse, 1981), (2) human errors in troubleshooting live aircraft power plants (Johnson and Rouse, 1982), and (3) aircraft pilots in mission flights (Rouse, Rouse and Hammer, 1982). Results of these studies have been applied to the improvement of training programs and the development of checklists and other decision aids.

5.3.2.1.3 Human Error in Aircraft Inspection and Maintenance

Whereas previous research in aircraft inspection and maintenance has utilized various empirical human factors techniques, this effort uses a behavior-based human error modeling approach, housed in a conceptual aircraft inspection and maintenance system model (see [Figure 5.3](#)). The system model provides a framework for error classification and therefore, a basis for improved error management. The following section describes the system model of aircraft inspection and maintenance. The final section details how the model might be useful for managing aircraft inspection and maintenance errors.

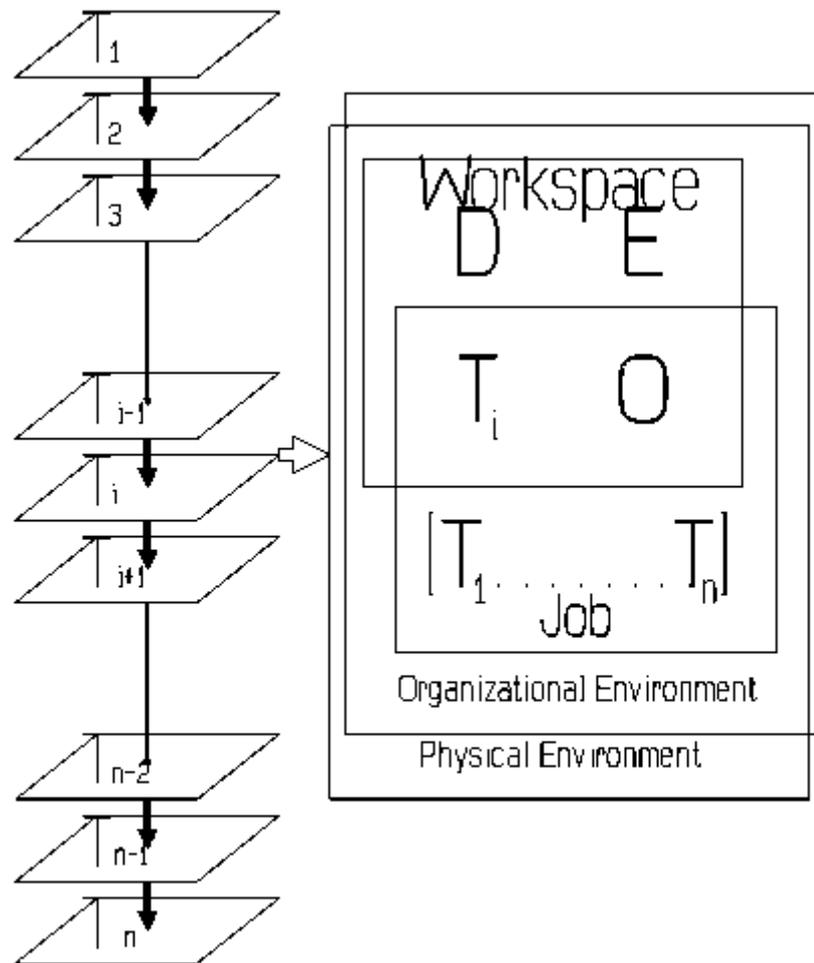


Figure 5.3 System Mode

5.3.2.2 A System Model for Human Error in Maintenance and Inspection

The fact that errors emerge from, and are defined by, the interaction of system characteristics, indicates the necessity of a system approach to the description and control of these errors. Such a system view of aircraft inspection and maintenance includes not only the traditional interaction of the operator and task requirements, but also includes operator interactions with equipment, documentation, and other personnel within the constraints imposed by the environment. The system model (Latorella and Drury, 1991) contains four components: operators (personnel), equipment, documentation, and task requirements. These components are subject to constraints of both the physical environment and the social environment. The job component can also be considered as a subset of the organizational environment in which tasks are defined. Similarly, the workspace component is a subset of the physical environment. This conceptual model is two dimensional as shown in [Figure 5.3](#). The temporal sequence of the individual tasks defines an axis orthogonal to the page. All other system elements interact with the current task component as shown in the plan view. Each individual task is subject to different combinations and degrees of influences from other system components, presented below.

Operators. Aircraft maintenance and inspection **operators (O)** differ between organizations but belong in the same basic categories: inspectors (perhaps distinguished as either visual or [NDT](#)), maintenance, utility, lead inspectors, lead maintenance, inspection foremen, maintenance foremen, production foremen, and engineers. In addition to carrying out sequences of activities, personnel

serve as informational resources to each other. Communication between personnel can be viewed as an information processing task similar to referencing a document. The organizational structure of the system imposes constraints on the amount of, format of, and the personnel likely to engage in, collaborative problem-solving communications. The affective and physical characteristics of people are also important. An individual's affect can influence motivation and hence, performance. Physical characteristics affect perception (e.g., visual acuity), access (e.g., anthropometry), and other tasks.

Equipment. Both visual and [NDT](#) inspection use **equipment (E)**. There is specialized equipment for different types of NDT, including: eddy current, ultrasonic, magnetic resonance, X-Ray, and dye penetrant. Visual inspection requires flashlights, mirrors, and rulers. Use of this equipment requires specialized knowledge of its operating principles, and equally specialized knowledge for the interpretation of its output. Interpretation of visual stimuli or NDT output necessarily requires information processing by the operator, but may also require communication with other personnel. The ability to perceive the information present in the visual stimuli or NDT output may be affected by environmental conditions, such as poor lighting. The ability to operate NDT equipment properly may also be affected by environmental factors. For example, some temperature and humidity combinations make precise movements difficult.

Documents. A variety of **documents (D)** is required for inspection and maintenance. Workcards, which may include graphics and references to more comprehensive standards manuals, specify the task to be performed. Forms (shift turnovers, [NRRs](#)) are used to communicate between personnel and to document procedures, while additional documentation is used for training and retraining purposes. The ability to communicate effectively through documentation is based on many factors. The fields specified on forms dictate the information and the structure of that information. Physical characteristics of forms, documents, and graphics affect the legibility of information and therefore, impact the ability to accurately perceive this information. Issues of comprehension are important for understanding the content of documents. Issues of representation are central to ensuring that graphics are appropriate and useful.

Task. A **task (Ti)** is defined as the actions and elements of one workcard or similar task order. Task characteristics which have been found to influence inspection include: defect probability, physical characteristics of the defect, the number of serial inspections, feedforward and feedback availability, and whether standards are used (Rodgers, 1983). These aspects of the task necessarily interact with personnel, organizational, job, and environmental characteristics. Personal information processing biases may interact with the task structure and present problems such as searching in the wrong area. The definition of a defect is part of the task which is ultimately established by the organization. An indication, which implies a defect, is defined as that magnitude which indicates that, given the cost/benefit tradeoff of repairing versus not repairing, a repair should be performed. The organization also dictates whether feedforward, feedback, and standards are used in inspection. The interaction of task characteristics and job characteristics may produce effects on inspection performance. The probability of defects affects the arousal level of an inspection and the expectation of finding a fault, which is also affected by the length of time an inspector performs a task and by physical factors such as fatigue.

Job. **Jobs (J)** are defined by the collection of tasks that an individual is expected to perform. However, there are many characteristics of the job which can not be described by the characteristics of its individual tasks. Job factors are derivative of the organizational environment and provide constraints for tasks (e.g., shift durations, work/rest cycles, day/night shifts, job rotation policies). These can further impact personnel physical (e.g., fatigue, eyestrain), affective (e.g., motivation, job satisfaction), and information processing (e.g., attention allocation) characteristics.

Workspace. The workspace, a subset of the physical environment, contains the task and the equipment, documentation, and personnel required to perform the task. While illumination is an attribute of the physical environment in general, task lighting (such as a flashlight) is an attribute of the workspace. The degree of physical access afforded by the workspace is an important constraint on performance. Both these issues are currently being researched under continued funding on this

contract (Gramopadhye, Reynolds, and Drury, 1992).

Physical Environment. The physical environment is described by several parameters: temperature, noise level and type of noises, lighting level and light characteristics, and electrical and chemical sources. While some of these factors can either enhance or degrade performance, others indicate potentially hazardous conditions. The level and spectral characteristics of lighting affects the perception of fault indications. Impulse noises interrupt tasks and may result in skipped or unnecessarily repeated procedures. The level and frequency characteristics of noise affect the ability to communicate. Examples of hazardous conditions in the physical environment are exposure to X-rays emitted during X-ray NDT and fuel fumes encountered when inspecting the inside of a fuel tank.

Organizational Environment. The organizational environment, often ignored in the analyses of maintenance systems, has been shown to be influential in the patterns of work (Taylor, 1990) and therefore, possibly in the patterns of errors. Factors which have been identified as important include: the organization of work groups (or conversely, the isolation of workers), reporting structures, payoff structures associated with task performance, trust within one class of personnel, trust between classes of personnel and levels of personnel, selection/placement strategies, and human-machine function allocation of control and responsibility. Organizational constraints are infused into every level of the organization. Regulatory agencies such as the [FAA](#), [JAA](#), and [CAA](#) mandate organizational form to some extent. Each organization has operational strategies and goals. These external and internal goals of the system, and constraints on the system are operationalized into changes in organizational structure, physical environment, task procedures, job descriptions, and personnel (skilled or trained).

Using the System Model. The model in [Figure 5.3](#) is useful for depicting the goals of the system and therefore the functions that should be supported. The goals of the system are defined by the requirements of the personnel component in isolation and in conjunction with other system components. The personnel component is primarily described in terms of information processing characteristics and limitations. These characteristics influence the behavior of individuals and their experience with other system components. The functions associated with the performance of tasks, use of equipment, and communication with co-workers are subject to error and are therefore of primary concern. These functions are then considered within the constraints of environmental factors which may affect error formation and/or propagation. Drury, Prabhu, and Gramopadhye (1990) have compiled a generic function description of the maintenance inspection task requirements as presented in [Section 5.1](#). The desired outcome for each of the task functions (Drury, 1991) which can be considered as the task's goal can be stated and, following Drury (1991), decomposed into the steps taken to accomplish the desired outcome (see [Table 5.7](#)).

Task 1 - INITIATE	1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8	Correct instructions written. Correct equipment procured. Inspector gets instructions. Inspector reads instructions. Inspector understands instructions. Correct equipment available. Inspector gets equipment. Inspector checks/calibrates equipment.
Task 2 - ACCESS	2.1 2.2 2.3	Locate area to inspect. Area to inspect. Access area to inspect.
Task 3 - SEARCH	3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9	Move to next lobe. Enhance lobe (e.g. illuminate, magnify for vision, use dye penetrant, tap for auditory inspection). Examine lobe. Sense indication in lobe. Match indication against list. Remember matched indication. Remember lobe location. Remember access area location. Move to next access area.
Task 4 - DECISION	4.1 4.2 4.3 4.4 4.5 4.6	Interpret indication. Access comparison standard. Access measuring equipment. Decide on if it is a fault. Decide on action. Remember decision/action.
Task 5 - RESPOND	5.1 5.2 5.3	Mark fault on aircraft. Record fault. Write repair action.
Task 6 - REPAIR	6.1	Repair fault.
Task 7 - BUY-BACK	7.1 7.2 7.3 7.4 7.5	Initiate. Access. Search. Decision. Respond.

Table 5.7 Detailed Breakdown of Aircraft Maintenance and Inspection by Task Step

Note that the use of equipment has been included within these task descriptions and therefore would not be considered separately. The most ambiguous situations encountered during aircraft inspection and maintenance typically result in an individual referencing another individual or a document for additional information. These situations are underspecified and are usually unanticipated. It is for

these reasons that understanding the communication errors which may occur at these junctures is important. The type of communication of interest here is only that related to task performance, although other forms of casual communication, not discussed here, may indicate important aspects of the organizational and social structure of the system.

Errors must be described in the situational context in which they occur in order to identify contributing factors. [Table 5.8](#) shows some relevant characteristics of system components with which the individual may interact for the 'initiate' task. Relevant characteristics of each system component can be identified for observed errors. The effect of these factors on performance has been suggested in many studies; however, the manner in which performance is affected, especially by combinations of factors, requires additional empirical investigation.

1.0 PERSONNEL	5.0 JOB
1.1 Physiological	5.1 Physical Factors
1.2 Psychological	5.2 Social and Organizational Factors
1.3 Personality	6.0 ORGANIZATIONAL/SOCIAL
2.0 EQUIPMENT	6.1 Structure
2.1 Hand Tools	6.2 Goals
2.2 Displays	6.3 Trust
2.3 Control	6.4 Motivational Climate/Incentives
3.0 DOCUMENTATION	6.5 Function Allocation/Job Design
3.1 Type of Information Included	6.6 Training/Selection Methods
3.2 Style (Intelligibility)	7.0 PHYSICAL ENVIRONMENT
3.3 Formatting (Visual Clarity)	7.1 Lighting
3.4 Content (Usefulness, Appropriateness, Veridical)	7.2 Noise
3.5 Legibility (Physical)	7.3 Temperature/Ventilation
4.0 TASK	7.4 Chemical Hazards
4.1 Physical Requirements	7.5 Vibration
4.2 Informational Requirements	7.6 Electrical Shock Hazards
4.3 Characteristics	8.0 WORKSPACE
	8.1 Proximity
	8.2 Anthropometrical Constraints

Table 5.8 System Component Influencing Factors

5.3.2.3 Previous Research in Human Error and Aircraft Inspection and Maintenance

There has not been a great deal of research on human error specifically related to inspection and maintenance, less still targeted to the inspection and maintenance of aircraft. Three approaches are discussed below which address this specific research area. Lock and Strutt (1985) employ a fault tree analysis approach to investigating and quantifying human error in aircraft inspection. Drury (1991) developed an error taxonomy of aircraft inspection based on a failure modes and effects analysis. Drury (1991) also has shown a classification scheme for aircraft inspection errors based on Rouse and Rouse's (1983) behavioral framework for investigating errors. These contributions are reviewed below.

Lock and Strutt (1985) begin their reliability analysis of inspection with a microstructural model of the inspection process. They use this model to "develop a flow chart ([Figure 5.4](#)) which describes a typical inspection activity in which visual information is used to trigger further investigation using other senses" (Lock and Strutt, 1985, p. 71). They note that while particularly suited to area checks, the scenario is generally applicable to a wide range of inspection tasks. These authors then analyze the flow chart for error-likely situations and they identify six potential errors in the inspection process:

Schedule Error (E1) Wrong execution of either of the two tasks: "identify next inspection" or "move to location."

Inspection Error (E2) Not seeing a defect when one exists.

Inspection Error (E3) If human induced, due to either "forgetting to cover an area" or "covering the area inadequately". May also be a schedule error.

Error of Engineering Judgement (E4) An error in deciding whether the area in which a defect is found is significant or not.

Errors in the Maintenance Card System Arises because the work cards

(E5) themselves may not be used to note defects on the hangar floor immediately as they are found.

Error in Noting Defect (E6) The error is noted incorrectly or not noted at all.

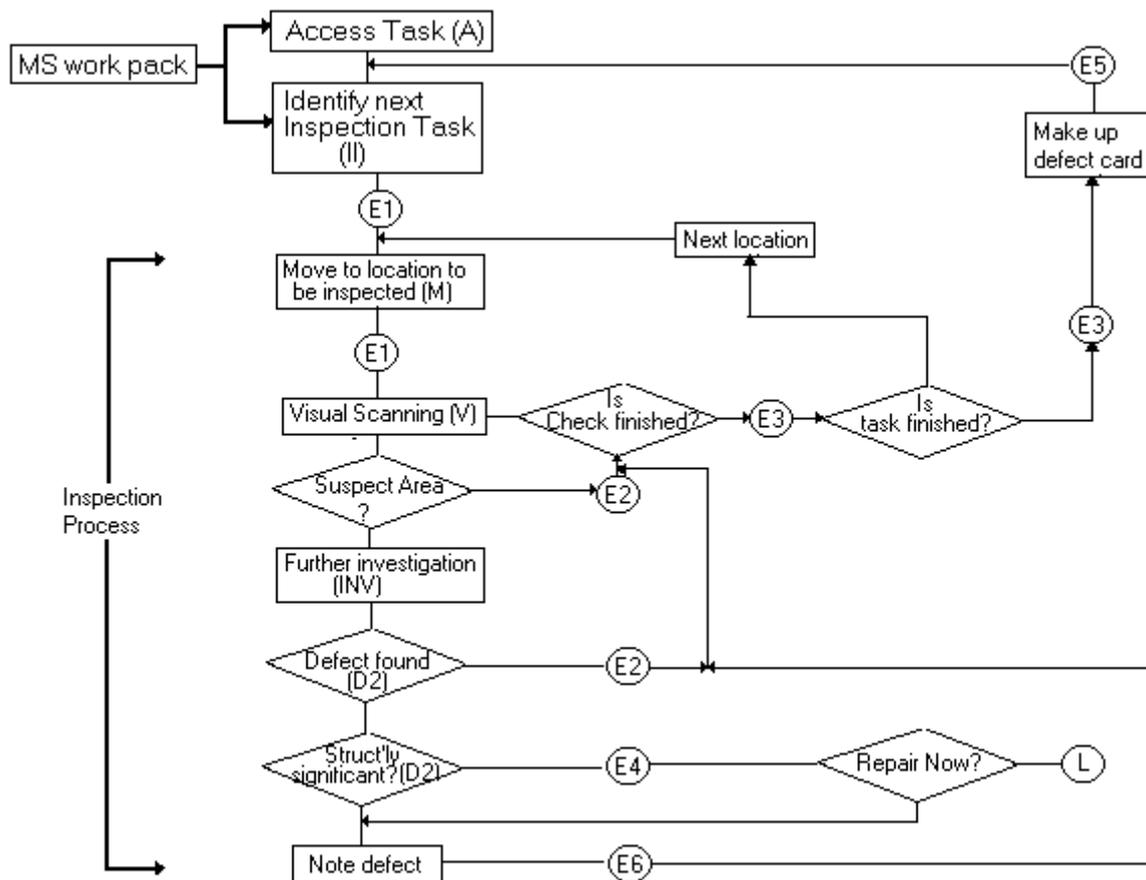


Figure 5.4 Inspection Model Flowchart (Lock and Strutt, 1985)

The authors recognize that these errors may co-occur to form compound errors. Lock and Strutt (1985) take a fault tree analysis approach to the inspection process with "inspection failure" as the top event (Figure 5.5). They note the difficulty of quantifying the probabilities needed for this type of analysis and make the necessary assumptions (i.e., indicating performance-shaping factors relevant to inspection, estimating their relevance at each step, and estimating probabilities of detection at different conditions in the model). Five performance-shaping factors (PSFs) were identified as relevant to aircraft inspection: accessibility of the aircraft area, lighting (general area), access and eyeball enhancement tools, motivation and attitude, and work method (Lock and Strutt, 1985). These PSFs were given relative weights to indicate their importance for each step in the inspection process. The authors propose, but do not actually perform, the fault tree analysis.

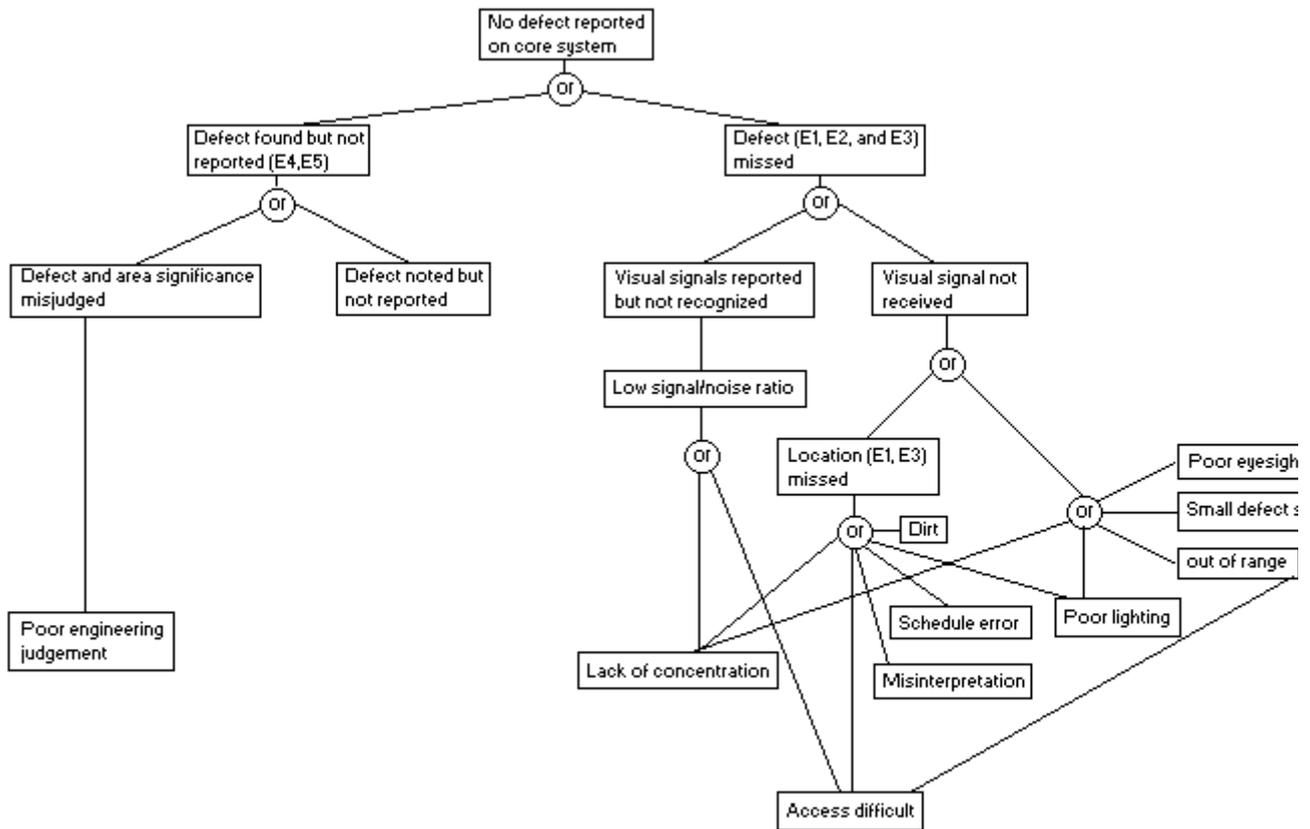


Figure 5.5 Inspection Error Fault Tree (Lock and Strutt, 1985)

Drury (1991) developed an error taxonomy from the failure modes of each task in aircraft inspection. This taxonomy has been developed based on the recognition that a pro-active approach to error control is needed to help identify potential errors. Thus, the taxonomy is aimed at the phenotypes of error (Hollnagel, 1989), that is, observed errors. Using the generic function description of the maintenance and inspection system (Drury, et al., 1990), the goal or outcome of each function was postulated as shown in [Table 5.7](#). These outcomes then form the basis for identifying the failure modes of the task. Towards this end, the tasks within each function were listed and the failure modes for each identified. These included operational error data obtained from observations of aircraft inspectors, and discussions with inspectors, supervisors, and quality control personnel involved in the aircraft maintenance task, over a period of two years (Drury, Prabhu and Gramopadhye, 1990; Drury, 1991). A sample of the error taxonomy (Drury, 1991) is shown in [Table 5.9](#).

TASK		ERROR(S)	
TASK 1 -- INITIATE			
1.1	Correct instructions written.	1.1.1	Incorrect instructions.
		1.1.2	Incomplete instructions.
		1.1.3	No instructions available.
1.2	Correct equipment procured.	1.2.1	Incorrect equipment.
		1.2.2	Equipment not procured.
1.3	Inspector gets instructions.	1.3.1	Fails to get instructions.
1.4	Inspector reads instructions.	1.4.1	Fails to read instructions.
		1.4.2	Partially reads instructions.
1.5	Inspector understands instructions.	1.5.1	Fails to understand instructions.
		1.5.2	Misinterprets instructions.
		1.5.3	Does not act on instructions.
1.6	Correct equipment available.	1.6.1	Correct equipment not available.
		1.6.2	Equipment is incomplete.
		1.6.3	Equipment is not working.
1.7	Inspector gets equipment.	1.7.1	Gets wrong equipment.
		1.7.2	Gets incomplete equipment.
		1.7.3	Gets non-working equipment
1.8	Inspector checks/calibrates equipment.	1.8.1	Fails to check/calibrate.
		1.8.2	Checks/calibrate incorrectly.

Table 5.9 Sample of Aircraft Maintenance and Inspection Errors by Task Step

The error framework developed by Rouse and Rouse (1983) has been used to record and analyze human errors in several contexts: (1) detection, diagnosis and compensation of engine control room failures in a supertanker (van Eckhout and Rouse, 1981), (2) human errors in troubleshooting live aircraft power plants (Johnson and Rouse, 1982), (3) aircraft pilots in mission flights (Rouse, Rouse, and Hammer, 1982). Results of these studies have been applied to the improvement of training programs and the development of checklists and other decision aids. Drury (1991) has shown how this scheme may be used to classify errors occurring in both visual and [NDT](#) inspection tasks (see [Table 5.10](#)).

LEVEL OF PROCESSING	POSSIBLE ERRORS
1. Observation of System State	Fails to read display correctly.
2. Choice of hypothesis	Instrument will not calibrate: inspector assumes battery too low
3. Test of hypothesis	Fails to use knowledge of NiCads to test.
4. Choice of goal	Decides to search for new battery.
5. Choice of procedure	Calibrates for wrong frequency.
6. Execution of procedure	Omits range calibration step.

Table 5.10 Example of Possible Errors for Task Step of Calibrate NDI Equipment (from Drury,1991)

5.3.2.4 An Approach to Aircraft Inspection and Maintenance Error Management

Error management may be considered as a three part objective. Errors which are evident in an operational system (error phenotypes) must be identified and controlled. Secondly, in order to reduce the likelihood of unanticipated error situations, errors must be predicted and systems must be designed to be error tolerant. Thirdly, error reporting systems must provide error and contextual information in a form which is appropriate as feedback to personnel. Operators may then use this information to adjust their error control and prevention strategies or alter environmental characteristics. This section presents strategies for error control and prevention through error-tolerant systems. Finally, the need for a context-sensitive error reporting scheme is discussed. Error phenotypes (Hollnagel, 1989), the specific, observable errors in a system, provide the foundation for error control. Error prevention and the development of design principles for error avoidance rely on genotype identification (Hollnagel, 1989), associated behavioral mechanisms, and their interaction with system characteristics (Rasmussen and Vicente, 1989). Here, error phenotypes are obtained empirically and from a failure-mode-and-effects analysis of task and communication models. These phenotypes are considered in light of their ability to be self-correcting and the type of error which they represent. They are further characterized by the relevant aspects of the system components with which they interact. The resulting list of phenotypes, their error correctability and type, and the pertinent situational factors, allow designers to recognize these errors and design control mechanisms to mitigate their effects. Rasmussen and Vicente's (1989) methodology is used to identify genotypes associated with each phenotype. This methodology yields the mechanisms of error formation within the task context.

This information in conjunction with consideration of influencing situational variables can predict the forms of novel errors and suggest design principles to prevent error formation and/or contain error propagation.

5.3.2.4.1 Error Control and Prevention

Error control is appropriate for the expedient eradication or mitigation of error-situation effects. However, there is much wisdom in the adage "an ounce of prevention is worth a pound of cure:" error prevention is more efficacious than error control. Error prevention requires error prediction and the design of error-tolerant systems.

Error control strategies can be derived by classifying error phenotypes according to components of the system model (see [Figure 5.3](#)) and according to Rasmussen and Vicente's (1989) systemic error mechanisms. This classification framework aids in suggesting intervention strategies appropriate to the error and the system components involved. The system model provides a useful means of classifying observed errors and relating them to specific human factors interventions. There are a number of personnel factors of general importance to controlling errors. Personnel interactions are extremely important aspects of the performance of the inspection and maintenance tasks. These interactions can be immediate but are also accomplished through the use of forms and notes which allow personnel to communicate with fewer temporal and spatial constraints. Communication is information transferred between not only personnel but between personnel and documentation. This extension of the common use of "communication" is logical given that documentation can be considered as a limited, static representation of some individual's (or group's) knowledge. Equipment should be designed to support task requirements and accommodate human information processing characteristics. The job and the individual tasks should be designed such that they can be accomplished at the desired level of performance, for the desired duration of performance, without physical or affective stress. The physical and organizational environments should be designed to enhance task performance and ensure the safety and motivation of personnel.

Various intervention strategies have been suggested for the control and prevention of errors. Rouse (1985) identifies five general interventions and proposes a mathematical model for describing optimal resource allocation among the strategies. These five general categories are also reflected in the more detailed listing of intervention strategies proffered by Drury, et al., (1990). These interventions have been tailored to the aircraft inspection context and were classified as either short-

term or long-term strategies. The intervention strategies from these two sources are described in detail in [Tables 5.11](#), [5.12](#), and [5.13](#). [Table 5.11](#) presents a compilation of the intervention strategies and design guidelines proposed by Rasmussen and Vicente (1989), Drury, et al., (1990), and Rouse (1985). These intervention strategies and guidelines are classified by the level of cognitive control (Rasmussen, 1986) which they affect and the type of systemic error (Rasmussen and Vicente, 1989) they address (see [Table 5.12](#)). Intervention strategies can also be classified by the component(s) of the aircraft inspection and maintenance system they alter. [Table 5.13](#) presents the compiled intervention strategies and design guidelines classified by levels of cognitive control, systemic error and system component. Further refinement of classification within system components (see [Table 5.8](#)) is possible with the aid of a more detailed decomposition of these components (see Latorella and Drury, 1991).

SHORT-TERM INTERVENTIONS (Shepherd, et al., 1991)	
1.	Worksheet design
2.	NDI equipment calibration procedures
3.	NDI equipment interface
4.	NDI equipment labeling of standards
5.	Support stands
6.	Area localization aids
7.	Stands/areas for NDI equipment
8.	Improved lighting
9.	Optical enhancement
10.	Improved NDI templates
11.	Standards available at the workplace
12.	Pattern recognition, job aids
13.	Improved defect recording
14.	Hands-free defect recording
15.	Prevention of serial responding (inadvertent signoff)
16.	Integrated inspection/repair/buy-back - improve written communication
17.	Integrated inspection/repair/buy-back - improve verbal communication
LONG-TERM INTERVENTIONS (Shepherd, et al, 1991 and Rouse, 1985)	
18.	Identification of errors - error reporting
19.	Integrated information systems (feedback, feedforward, directive)
20.	Training
21.	Selection/placement
ERROR REDUCTION RESOURCES (Rouse, 1985)(also notes training and selection)	
22.	Equipment design
23.	Job design
24.	Aiding
RASMUSSEN'S "COPING" GUIDELINES (Rasmussen and Vicente, 1989)	
25.	Make limits of acceptable performance visible while still reversible.
26.	Provide feedback on the effects of actions to cope with time delay.
27.	Make latent conditional constraints on actions visible.
28.	Make cues for action put only convenient signs, but also represent the necessary preconditions for their validity (symbolic).
29.	Supply operators with tools to make experiments and test hypotheses.
30.	Allow monitoring of activities by overview displays.
31.	Cues for action should be integrated patterns based on determining attributes (symbolic representations).
32.	Support memory with externalization of effective mental models.
33.	Present information at level most appropriate for decision making.
34.	Present information embedded in a structure that can serve as an externalized mental model.
35.	Support memory of items, acts, and data which are not integrated into the task.

Table 5.11 Error Management Strategies

SYSTEMIC ERRORS	LEVELS OF COGNITIVE CONTROL		
	SKILL	RULE	
Learning and Adaptation	6, 11, 12, 20, 22, 24, 25	1, 2, 11, 20, 22, 24, 28	1, 13, 16, 22, 24, 28
Interference Among Competing Control Structures	12, 14, 23, 24, 30	3, 11, 20, 22, 23, 24, 31	1, 3, 15, 24, 32
Lack of Resources	33	33, 34	1, 3, 4, 11, 21, 22, 23
Stochastic Variability	2, 3, 5, 6, 7, 8, 9, 10, 12, 14, 16, 17, 20, 21, 22, 23, 24, 35	2, 11, 14, 16, 17, 20, 21, 22, 24, 35	4, 16, 17, 21, 22, 23

Table 5.12 Error Management Strategies by Systemic Error and Level of Cognitive Control

Systemic Errors	Levels of Cognitive Control	SYSTEM ELEMENTS					
		Task	Personnel	Job	Workspace	Equipment	Doc
LEARNING	Skill	6,11,12	20		6	6,22,24,25	6,24
	Rule	11,28	20			11,22,24,28	1,2,3
	Knowledge	13,19,26	16,17,18,20	16,17,26,29		13,19,22,24,26,27,29	1,16
INTERFERENCE	Skill	12,14		23,30		14,24,30	24
	Rule	11	20	23		3,11,22,24,31	24
	Knowledge	19,32	16,17,20	15,16,17,23,32		3,19	1,16
LACK OF RESOURCES	Skill					33	
	Rule					33	33,34
	Knowledge	13,19,34	16,17,20,21	16,17,23,34		3,4,13,19,22,24,33,34	1,16
STOCHASTIC VARIANCE	Skill	6,12,14	16,17,20,21,35	23	5,6,7,8	3,5,6,7,9,10,14,22,24,35	6,24
	Rule	11	16,17,20,21,35			11,22,24,35	2,24
	Knowledge		16,17,20,21,35	16,17		4,22,35	16,3

Table 5.13 Error Management Strategies by Systemic Error, Levels of Cognitive Control, and System Component from Figure 5.3

The above methodology was developed to control errors, i.e., for error phenotypes which are observable errors in the system. An extension of this methodology provides a means by which intervention strategies can be identified to control unanticipated errors once they occur. In this extension, error genotypes, rather than the aforementioned phenotypes, are classified according to the system model, using Rasmussen and Vicente's (1989) systemic error categories and Rasmussen's levels of cognitive control (Skill, Rule, Knowledge). This characterization of error genotypes allows prediction of possible, but so far unanticipated, error phenotypes. Unanticipated errors can be predicted by considering tasks at each level of cognitive control and each error mechanisms' possible perturbation of performance within the context of the specific system components involved. Given an error genotype cell, intervention strategies (which also have been classified by system component, systemic error mechanism, and cognitive control level (see [Table 5.13](#)) can be identified for its control.

5.3.2.4.2 Error Tolerant Design in the Aircraft Inspection and

Maintenance System

An error tolerant system has been defined as a system which ensures that recovery from errors is possible, in the sense that actions are reversible and/or that the system is resilient to inappropriate actions (Rouse, 1985). Reason (1990) suggests that one way of making systems more error tolerant is to identify "those human failures most likely to jeopardize the integrity of the plant and to defend against them by engineered safety devices or procedures" (p. 233). For example, the "30-minute rule" allows nuclear power plant operators 30 minutes of thinking time in an emergency through the use of automatic systems which can return a plant to a safe state without human intervention. Reason also notes that, where these safety devices are themselves subject to human errors, independent, redundant systems should be provided (p. 233). The design of error tolerant system procedures and devices can be guided by the error control and prediction framework previously described by incorporating interventions in plant and operating procedure design.

5.3.2.4.3 An Approach to Reporting Aircraft Inspection and Maintenance

Errors

Currently, error reports are primarily used for documenting error situations for administrative purposes by internal or external regulatory agencies. There are many different regulatory mechanisms for reporting errors to the [FAA](#). In addition, the Air Transport Association (ATA) has proposed modifications to those. All of these reporting systems have the following common features:

1. They are event driven. The system only captures data when a difficulty arises or a defect is found.
2. Aircraft type and structure serve as the classification parameters for reporting.
3. Expert judgements of error criticality are used to further classify data and determine its urgency.
4. To some extent in all systems, the feedback of digested data to users is not well-engineered. Thus, for the end-user level, the data collection effort is largely for naught.
5. They can result in changes in maintenance and inspection procedures; for example, by issuing Airworthiness Directives (ADs).

Error reports in maintenance and inspection produced for administrative purposes are typically concerned with establishing accountability for an error and its consequences rather than understanding the causal factors and situational context of the error. This type of information is not appropriate for use as performance feedback to inspectors or maintenance personnel, nor is it helpful information for error tolerant system design. Error reporting schemes are developed from within an

organization and therefore vary greatly among organizations. The framework of these error reporting schemes is event driven and developed iteratively, thus additions are made only with the occurrence of a new error situation. To a large extent, the information recorded about a situation is constrained by the format of the error reporting scheme. For example, in one error reporting scheme, the reviewer is required to attribute the error to some form of human error unless the situation can be described as an "act of God" (Drury, 1991). Analysis of the data collected by such a scheme will invariably find the human at fault, rather than working conditions, equipment, procedures, or other external factors. This biased representation has serious implications for error prevention, especially considering that equipment design and job aiding have been found to be more efficacious than selection or training approaches in error prevention (Rouse, 1985). To alleviate the difficulties of inconsistency, and provide an appropriate and useful structure for error data collection, an error reporting scheme should be developed from a general theory of the task and the factors which shape how the task is performed. Principally, the behavioral characteristics of the operator, but ideally also organizational environment, job definition, workspace design, and the operators' physical, intellectual and affective characteristics should be considered. Effective error categorization systems are not only descriptive but are prescriptive, providing information for specific intervention strategies (i.e., Langan-Fox and Empson, 1985 and Kinney, et al., 1977).

As Rasmussen, Duncan, and Leplat (1987) note, it is necessary to shift the focus of analysis from the task to the interaction of the task and the operator for classifying errors. Furthermore, taxonomies of human error must encompass the analysis of not only the task characteristics but also the information processing mechanisms associated with the subtasks. It is apparent that other situational characteristics (i.e., environmental conditions) are also useful for the sensitive classification of errors (Stager and Hameluck, 1990). Correlations of errors with situational factors, with remedies attempted, and with the effects of these remedies, may provide important feedback for identifying error situations, assessing error criticality, and determining error consequence-minimizing solutions. Both error control and error prevention would benefit from an error reporting system which captures the causal factors and situational context of an error situation.

Both the taxonomic approach of Drury and Prabhu (1991) and the taxonomy for error management strategies developed here can be used as a basis for formulating error reporting schemes. Upon occurrence, errors can be classified by level of cognitive control, type of systemic error, and by causal or catalytic elements of the system. As previously mentioned, the categories of system elements can be refined as illustrated in [Table 5.8](#) to provide a more descriptive error characterization. Identification of these parameters will likely involve detailed investigation of the error situation, including extensive operator interviewing. This data store can be analyzed for trends in error sequences, effects of different intervention strategies on error-type frequency, and for the efficacy of intervention strategies over all types of errors. Identification of error sequences and the effects and interactions of system elements provides important feedback information for performance and feedforward information for training, equipment, and job design. A prototype error reporting system based on the above considerations has been proposed as a short-term project with an airline partner.

5.3.3 A FRAMEWORK FOR INFORMATION ENVIRONMENT DESIGN FOR AIRCRAFT INSPECTION

Inspection is information processing. Other aspects of the inspector's task, such as physical access to the work and body posture during work, are subordinate to this central task. If information processing is the essence of inspection, we must examine the sources of information used (and not used) by the inspector: how information is received, processed and generated. Hence, the inspector's information environment is a critical part of the inspection system.

Any system involving a human is typically closed loop (e.g., Sheridan and Ferrell, 1974). Obvious examples are in flying an aircraft or driving a car, but the concept applies equally to inspection tasks. As shown in [Figure 5.6](#), the human in the task receives some instruction, or command input to use systems terminology. The operator and any associated machinery transform this command input

into a system output. To ensure stable performance, the system output is fed back to the input side of the system, where it is compared against the command input. If there is any difference (command minus output) the system responds so as to reduce this difference to zero.

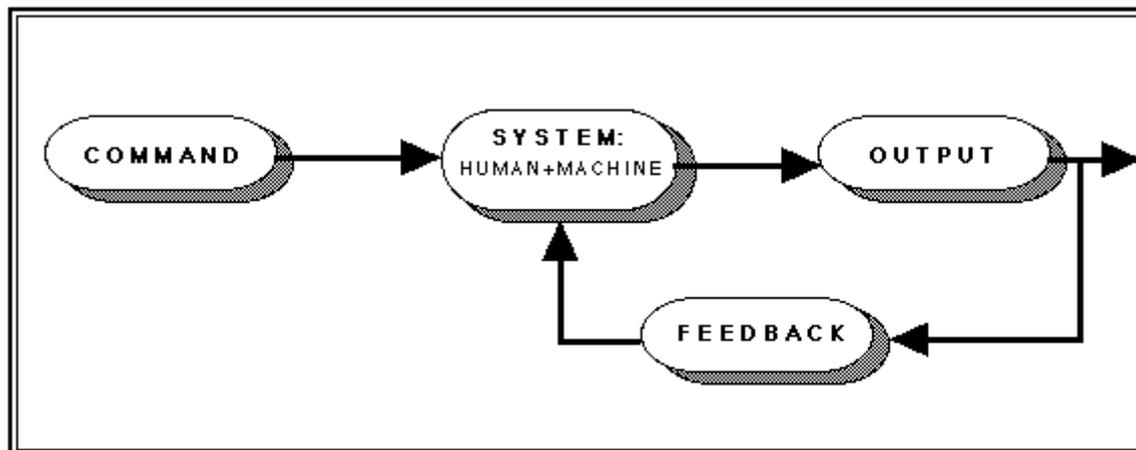


Figure 5.6 Closed-Loop Control

From the model in [Figure 5.6](#), it is obvious that two types of information can be distinguished. The input is command information, while the output is feedback information. Both have been shown to be amenable to manipulation to improve system performance. Not obvious from [Figure 5.6](#) is that the command input may be complex, and includes both what needs to be accomplished and help in the accomplishment. Thus, input may give both directive and feedforward information. A work card may contain "detailed inspection of upper lap joint" in a specified area (directive) and "check particularly for corrosion between stations 2800 and 2840" (feedforward). Thus, there are really three potential parts to the information environment: directive information, feedforward information, and feedback information.

Three of the strongest influences found in case studies of inspection performance are time pressure on the inspector, feedforward of information to the inspector, and feedback of detailed performance measures. We restrict ourselves to examining the various aspects of feedforward and feedback information in the context of aircraft inspection; the time pressure aspect is dealt with under speed accuracy tradeoff in [Section 5.3.4](#).

In the subsequent sections we present a model of the information flow in aircraft inspection. This model serves as the basis for understanding the information environment that the inspector is a part of. We then present two approaches to analyze the information requirements of the inspection task: (a) skill-rule-knowledge (S-R-K) based approach, and (b) error taxonomic approach. Finally, a study to investigate the effect of feedback information is described.

5.3.3.1 A Model of Information Flow in Aircraft Inspection

To perform optimally in the system, the inspector has to have access to the relevant information and the information environment has to provide this information. We have to reconcile the, perhaps conflicting, issues of:

- What information to present.
- When to present this information.
- How to present this information.

In designing the flow of information, the designer must take into account human processing of information and the cognitive abilities of humans. It is important to develop a model of the information environment in order to analyze the current system and propose design changes based on

identified problems. Towards this end we propose a feedforward/feedback information model of aircraft inspection (see [Figure 5.7](#)). This model represents both the physical work flow and the information flow. It also highlights the cognitive aspects of the inspection task and its interaction with the information environment.

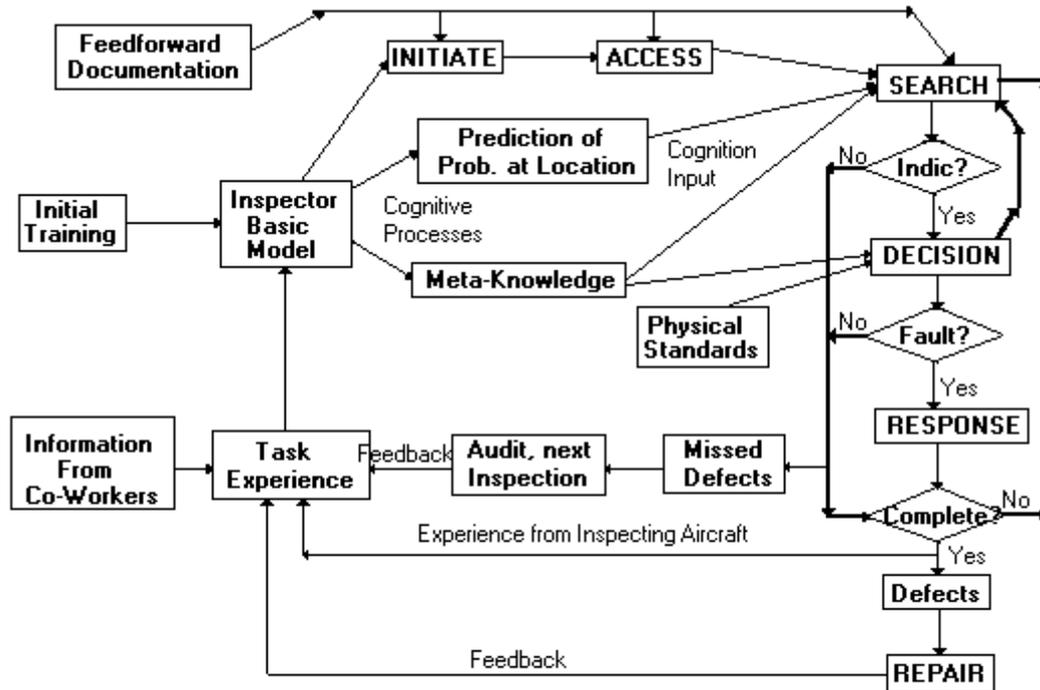


Figure 5.7 Model of Information Flow in Aircraft Maintenance and Inspection (Drury and Prabhu, 1991)

This model allows us to target the components of feedforward (training, documents, etc.) and feedback (missed defects, defect rate, etc.) that have to be analyzed for efficient design of the information environment.

5.3.3.1.1 Feedforward Information

From the model ([Figure 5.7](#)), feedforward information to the inspector is seen to come from the following sources:

1. Initial Training
2. Manufacturer/[FAA](#)/Airline Operator documents.
3. On-the-Job experience on a particular aircraft.
4. Information gathered from co-workers.
5. Command information in the form of standards.
6. Utilization of understanding about the fault causation mechanism in an aircraft.

Initial Training. Taylor (1990) notes aircraft orientation training for new mechanics, at large sites. However, smaller sites had no formal training programs in place. No formal inspection training programs were observed or reported at any of the airlines. Typically, inspectors hold an A and P license and have maintenance experience. Taylor (1990) found that the current hangar maintenance organization has a bi-modal experience distribution of 30 plus years and three or fewer years. The inspection group is expected to have a similar distribution with three to five years added to the lower

value.

The current state of training places much emphasis on both the procedural aspects of the task (e.g., how to set up for an X-ray inspection of an aileron), and on the diagnosis of the causes of problems from symptoms (e.g., troubleshooting an elevator control circuit). However, the inspectors we have studied in our task analysis work have been less well trained in the cognitive aspects of visual inspection itself. How do you search an array of rivets -- by columns, by rows, or by blocks? How do you judge whether corrosion is severe enough to be reported?

Most of the training is on the job where an experienced inspector puts the novice through his paces and shows him the various aspects of inspection. This is highly realistic but uncontrolled and there is a high likelihood for development of inconsistent inspection practices. Our experience in training inspectors in manufacturing industries (Kleiner, 1983) has shown that a more controlled training environment produces better inspectors. If training is entirely on-the-job, then two of the main determinants of the training program--what the trainee sees and what feedback is given--are a matter of chance; i.e., of which particular defects are present in the particular aircraft inspected.

We need to develop training procedures for the search and decision making components of aircraft inspection by using human factors techniques that include cueing, feedback, active training, and progressive part-training as suggested by Drury and Gramopadhye (1990) and detailed in [Section 5.3.5](#). It has been found that off-line controlled training successfully transfers to the more complex on-the-job environment. The trainee is prepared to make maximum use of what is seen on the job, rather than confining the learning process to trial and error. Because of the controlled and concentrated training experience, trainees can progress faster to the same level as experienced inspectors.

Documentation. There is an immense amount of potentially useful information available both in paper (hard copies) and paperless (computer, microfiches) form. We list below some of the important documents that form the information environment. Note that this is not a complete listing of all available documents.

The documents are generated by a triad consisting of the Federal Aviation Administration (FAA), aircraft manufacturers, and aircraft operators. There is a complex, multi-dimensional interaction in the flow of data between these three. Manufacturers require feedback from operators to determine acceptability and reliability of a product and its components. Airlines require product support information from the manufacturer. The [FAA](#) requires data from both the airlines and the manufacturers concerning product reliability and safety issues. The Air Transport Association (ATA) coordinates the flow of data among the three triad members (Shepherd, 1990).

We have to understand the problems created by the mismatches between the needs of the inspector (who is looking for information) and the design of the documents (that present data). There is a critical need for usable knowledge, which gets translated to utilized information on the job. From a document design viewpoint we have to focus on creating usable documents. Information flow design and system design should ensure the availability of documents at the right place at the right time. The demonstration project presented in [Section 5.2.1](#) is an example of applying document design techniques to one type of document, the workcard.

Experience on a Specific Aircraft Type. Aircraft at a maintenance facility are serviced over various lengths of time depending on the type of service. The transfer of an aircraft to a different facility (other than the one it normally goes to) is very rare and occurs in case of contingencies or in case of heavy workload at the regular facility. Similarly, movement of personnel between different facilities is very low. Thus, most maintenance and inspection personnel accumulate experience on a particular type of aircraft. The effect of such job specialization on the occasional inspection of a different aircraft type has not been studied.

Knowledge about the aircraft is accumulated over a period of time through on-the-job work. Experienced inspectors gradually develop an understanding of the cause-effect relationship of defects and also know what to look for and where. Thus, there is a store of distributed knowledge or

expertise residing in the inspection organization. Individual inspectors normally have access to this distributed knowledge through informal contacts with fellow inspectors, which leads us to the next section.

Information from Co-Workers. The relevant relationships in heavy maintenance have been identified by Taylor (1990), to include:

1. Superiors with subordinates
2. Members of same group with one another
3. Members of different work groups
4. People inside enterprise interacting with people outside that system.

Airline inspectors typically work independently and occasionally in teams of two. The frequency of formal meetings amongst inspectors varies from airline to airline. In one airline, weekly safety meetings are held where any communications from management are conveyed to the inspectors. In another case there is a daily meeting at the beginning of the shift where the day's work and assignments are discussed. Drury, et al., (1990), during the task analysis of inspection in the airline industry, found few formal meetings of mechanics or inspectors despite frequent informal contact among inspectors, and less frequent contact between inspectors and mechanics.

Contact between inspectors, in different shifts, was observed at some sites where shifts overlapped by an hour or so. The mechanics and inspectors contact each other for buy-back or for approval of a repair. This contact for advice/instruction is the only formal information exchange between the inspector and the mechanic. There appears to be no formally organized forum that can channel the distributed knowledge for more efficient access by individuals who need this information.

Mechanics who find faults during scheduled maintenance notify the inspectors. Thus, an informal system of communication exists. However, there are various ways in which such a system can break down. An experienced inspector might know, for example, that the line maintenance people have in the past improperly used magnetic screws around the landing light as a contingency measure. Thus, he/she would examine the screws around the landing light in view of this knowledge. A new inspector may not have had access to this issue (which is not mentioned in a workcard or any documentation elsewhere) and could fail to catch such a fault. Similarly, an inspector who documents a fault and the inspector who approves the repair done on this fault may not be the same and thus, any inspection error in this case goes unnoticed by the inspector because of a lack of a formal feedback system.

Command Information with Comparative Standards. There seem to be almost no standards that are accessible to inspectors for defects like corrosion, cracks, dished/pooched rivets, wear, component play, etc. A small subset of standards does exist with the manufacturer, [FAA](#), etc., but these have not been organized into a scheme for utilizing comparative standards on the job. The closest inspectors come to a standard in visual inspection is to use adjacent areas to make a comparison, which is not a reliable method (Drury, 1991).

During a decision making process, both the internal and external retrieval of information is necessary. The degree with which external and internal retrieval of information is required could be a major determinant of the strategies adopted during decision making.

As an example, during visual search for corrosion around rivets or in a door frame the inspector comes across an indication. The inspector has to make a judgement call whether this indication should be marked as a defect or let go. If the corrosion is evident without a doubt, then the decision process is simple and the task is almost like a pure search task. On the other hand, when the evidence for corrosion in the indication is not conclusive, the inspector has to:

1. Retrieve internal information about instances of corrosion to make a match (recall patterns).
2. Approach peers or supervisors for help on judgement.
3. Refer to comparison standards available at the work point.

It has been found that the higher the information load and the more likely the chance of error, the more an operator is forced to remember or recall information of relevance. Also, external information retrieval (from other inspectors) is a function of the operator's perception of criticality of this particular decision and availability of inspectors within a reasonable vicinity. For example, the inspector perched up on the horizontal stabilizer of a DC-9 is less likely to go down and call a supervisor to come up and have a look at an indication, particularly if he perceives that a wrong decision on his part may not be critical.

It has been known for many years that if comparison standards are available at the work point, more accurate inspection will result. Yet in many cases such standards are not available to the aircraft inspector. For example, if the maximum allowable depth of a wear mark is given as 0.010 inches, there is neither a convenient way to measure this, nor a readily available standard for comparison. Other examples are play in bearings and cable runs, areas of corrosion, and looseness of rivets. All are considered to be "judgement calls" by the inspector, but simple job aids, perhaps as part of the worksheet, or standard inspection tools, would remove a source of uncertainty. Leaving standards to unaided human memory may be expeditious, but it is also unreliable.

Utilization of Understanding about the Fault Causation Mechanism in Aircraft. Inspection of aircraft is largely composed of pure search activities followed by decision-making tasks whose output is of the form of "acceptable/non-acceptable". However, some areas of inspection involve utilization of cues, knowledge of how faults are caused, and knowledge of how the behavior of one particular aircraft component indicates behavior of related components. Examples are:

- dirt streaks around a rivet on the fuselage indicate a loose rivet,
- bulging of the paint on the aircraft skin indicates underlying corrosion,
- scraped paint at the fairings indicates underlying fairings are rubbing,
- play at the flap vanes points to worn out bearings or tracks,
- flat spots on the wheel indicate a possible problem with the anti-skid system,
- powdery material on the skin indicates probable corrosion.

Use of such indirect evidence is a powerful technique to enhance detection and discovery of a fault, particularly where faults are not directly accessible to pure visual/auditory/tactile search.

There is a necessity to gather the knowledge required for this indirect fault indication from experienced inspectors who understand the utilization of such cues. There is also a need to identify the mappings between defects and fault causation mechanisms for a wide variety of such defects. The decision-making activity can then be converted to a rule-based, procedural type of task. Rules thus formed can be used in an effective training scheme to help inspectors increase the efficiency of the search and decision making process.

This approach can be extended further to form an inspection data base which can be continually revised and updated to reflect the distributed knowledge that exists not only in a specific airline but across all airlines. Such a global knowledge-base would thus receive its input from experienced inspectors all over the aviation industry, thus consistently benefitting all users. It is also conceivable that an expert system could be developed that makes use of such a data base and supports decision-making tasks. Such a system would support queries like:

- "I am in the tail compartment. Current inspection area is aft of APU compartment bulkhead, list keypoints."
- "Inspection area is APU shroud, list past history of cracks."
- "Indications at rivet on lap joint at stringer S-34 between body station 890 and 900 points to corrosion, show graphics of likely corrosion in this area."
- "There is excessive play at the flap vanes, what are the problems indicated by this.", etc.

5.3.3.1.2 Feedback Information

Feedback information in aircraft inspection can be used either on the job or in training. Use of feedback on the job has been found to reduce the number of false alarms as well as reduce missed defects. Training schemes implementing feedback have been used to improve learning rates, to develop schemes, and for the efficient transfer of training skills to on-the-job performance.

On-The-Job Feedback. There seems to be no systematic and obvious system in place that provides feedback to the inspector. For example, feedback during access can be given by a well designed workcard system incorporating unique landmarks in the figures (Drury, 1990b). Feedback in search/decision making comes when the inspector talks to a supervisor or a fellow inspector to confirm a borderline case, although this occurs rarely. Also rare is the feedback that could come from the repairer or the buy-back inspector who both have potential data on the fault.

Feedback also seems to depend on the type of defect. Airlines have a system to classify the various defects found during inspection/maintenance. There are specific rules by the [FAA](#) for this classification. Normally, defects get classified in three broad categories: A, B, and C. Type "A" defects are the most critical ones and have to be immediately corrected. Type "B" defects are corrected immediately, or the maintenance action is deferred to a pre-specified time based on current and projected workload. The "C" defects are generally deferred to the next inspection. Thus, there exists a possibility of feedback in the case of "A" defects, and some "B" defects, because of the time frame within which maintenance action is taken. This would normally occur through buy-back inspection. However, even this opportunity would be lost if the buy-back inspector is different from the one who wrote the non-routine defect item.

There is very little feedback on any defect that the inspector misses. This feedback can only occur through audits and quality control inspections, but these systems do not ensure consistent feedback to all inspectors on a regular basis.

At this point we have to also recognize that, although it is very desirable to provide feedback, there are bound to be instances where this would be economically infeasible, and in some cases impossible, due to the nature of the task. For example, providing regular feedback on missed defects is not viable, as it would involve re-inspection similar to auditing on a regular basis. Similarly, having a system that calls for feedback on every defect may be too expensive due to time factors and logistics. In such cases, alternate schemes like periodic re-training or off-line feedback could be utilized to re-calibrate inspectors.

Feedback in Training. As explained in the earlier section, the feedback in aircraft inspection is relatively scarce, and on the occasions that the inspector gets feedback (e.g., an audit), it is delayed in time. Delayed feedback makes learning by practice alone difficult (Woods, 1989).

The use of knowledge of results (feedback) in training is well documented. The trainee needs rapid, accurate feedback in order to correctly classify a defect or to know whether a search pattern was effective. However, when training is completed, feedback is rare. The training program should start with rapid, frequent feedback, and gradually delay this until the "working" level is reached. More feedback beyond the end of the training program will help to keep the inspector calibrated (Drury and Kleiner, 1990).

We see that there is a great deal of research support to indicate that use of feedback in initial training is beneficial. From the airline inspection context this points to the necessity of developing a training methodology that incorporates performance feedback. Drury and Gramopadhye (1990) have demonstrated a training scheme for gamma ray inspection of a nozzle guide vane area of a JT9D engine. This includes part naming and defect naming (cueing and active response), search, and decision training. Feedback is used judiciously in this training scheme to help the trainee to build a schema.

5.3.3.2 Analysis of Information Requirements: An S-R-K Approach

So far it has been established that (a) errors in aircraft inspection are costly, and therefore must be

minimized, (b) human performance limitations can, and do, result in inspection errors, and (c) provision of information in the correct form (physical and cognitive aspects) is critical to reducing human errors.

For effective use of feedforward and feedback information, the information requirements of human inspection have to be identified. Furthermore, the information needs of experts and novices may be very different. Thus, we can posit that studying the behavior of the human inspector interacting with the system (while performing the inspection) will help identify possible information support points, as well as provide guidance to the type of information (either feedforward or feedback) that is needed at these points. The skill-rule-knowledge based hierarchy of Rasmussen (1983) presented in [Section 5.3.2.1.2](#) affords us a robust framework within which this analysis can be carried out, and will be mapped onto both visual inspection and [NDI](#).

5.3.3.2.1 Visual Inspection

Search and decision making form the critical components of visual inspection. The search component can be further decomposed into pre-attentive search, and a detailed search consisting of foveal (pure search or search plus decision making) and extra-foveal processes. Similarly, [NDI](#) can be decomposed into three broad stages: calibration, probe movement, and display interpretation. Identification of the behaviors associated with each of these subtasks results in a many to many mapping as seen in [Table 5.14](#) (Visual Search) and [Table 5.15](#)([NDI](#)). These mappings have been identified for an expert inspector. An interesting aspect of these mappings is the existence of relatively few knowledge-based behaviors exhibited by the expert inspector. This seems logical since there is less problem-solving or active reasoning in aircraft inspection and more detection, identification, and classification.

VISUAL INSPECTION PROCESSES	BEHAVIOR CATEGORIES		
	SKILL-BASED	RULE-BASED	KNOWLEDGE-BASED
PRE-ATTENTIVE SEARCH	Scan and Detect		
FOVEAL (PURE SEARCH)	Fixate and Detect		
FOVEAL DECISION		Identify and Classify	
EXTRA-FOVEAL SEARCH	Trigger move to next area		
DECISION-MAKING (OUTSIDE OF SEARCH)		Move to next area, Rules of what to look for	Reason and Decide

Table 5.14 Mapping a Visual Inspection Task to Cognitive Behaviour for an Expert Instructor

NDI PROCESSES	BEHAVIOR CATEGORIES		
	SKILL-BASED	RULE-BASED	KNOWLEDGE-BASED
CALIBRATION	Probe Movement Over Test Specimen	Calibration Procedures	
PROBE MOVEMENT	Tracking Along Desired Path	Supportive Mode Identifying Boundary Conditions	
DISPLAY INTERPRETATION		Interpreting Familiar Signal	Interpreting Unfamiliar, Unanticipated Signals

Table 5.15 Mapping an NDI Process to Cognitive Behavior for an Expert Inspector

The [SRK](#) framework also aids understanding of how behavior will be qualitatively modified as the inspector goes from a novice to an expert. Thus, although both the novice and the expert exhibit, say, rule-based behavior, the behavior of the expert will be qualitatively different from the novice (Sanderson and Harwood, 1988). In [Table 5.16](#) we have mapped a specific visual inspection task (inspection of rivets) to the [SRK](#) framework, to represent the performance of an expert inspector. We can expect that some of the defects identified at the skill-based and rule-based levels by the expert will be identified at the rule-based and knowledge-based levels by the novice, indicating a rightward shift on [Table 5.16](#), corresponding to an upward movement on the [SRK](#) hierarchy. Thus, this analysis points to the need for different levels of information support for the expert and the novice inspector. It can also provide guidelines to define training requirements for novice inspectors based on identifying expert inspector behaviors.

VISUAL INSPECTION PROCESSES	BEHAVIOR CATEGORIES		
	SKILL-BASED	RULE-BASED	KNOWLEDGE-BASED
PRE-ATTENTIVE SEARCH	<ul style="list-style-type: none"> ▪ Missing rivet ▪ Hole in skin 		
FOVEAL (PURE SEARCH)	<ul style="list-style-type: none"> ▪ Missing rivet ▪ Hole in skin ▪ Deep dents ▪ Large cracks ▪ Prominent corrosion 		
FOVEAL DECISION		<ul style="list-style-type: none"> ▪ Borderline corrosion ▪ Slight wear ▪ Dished rivets ▪ Ripples in skin ▪ Small cracks 	
EXTRA-FOVEAL SEARCH	<ul style="list-style-type: none"> ▪ Chipped paint in periphery leads to next fixation 		
DECISION-MAKING (OUTSIDE OF SEARCH)		<ul style="list-style-type: none"> ▪ Streaks around rivets trigger inspection for loose rivets ▪ Powdery contamination triggers search for corrosion ▪ Borderline defects 	<ul style="list-style-type: none"> ▪ Defect type not listed ▪ Use of meta-knowledge

Table 5.16 Visual Inspection of Rivets: Cognitive Behaviors for Different Defect Types

Tables 5.14, 5.15, and 5.16 also indicate the large role that skill-based and rule-based behaviors play in visual inspection. The visual search part of visual inspection is seen to be entirely skill and rule-based for the expert inspector (or after training to criteria). The skill-based behavior can be associated to the scanning, fixating, and detecting activities (see Table 5.14). Since skill-based performance is essentially unconscious and feedforward controlled, we can conclude that the information aid for this part of the visual search should be something that does not require active conscious use by the inspector. This points to visual environment changes (better lighting, improved contrast), and improved human detection capabilities (increasing visual lobe, increasing target conspicuity). At the same time, this also indicates training as a critical need to attain satisfactory sensory performance.

Tables 5.14 and 5.16 also highlight rule-based behavior resulting in the identification and classification of defects as a significant mode of visual inspection. Thus, finding corrosion, wear, small cracks and similar difficult defects takes place because of rule-based behavior. It is pertinent to note at this point that the work card system used in the aircraft industry to control aircraft maintenance and inspection relies heavily on a linear procedural approach (Drury, 1991; Drury, Prabhu and Gramopadhye, 1990). Rule-based behavior also accounts for search strategies based on past experience and work card instructions. Thus, we reach the conclusion that it is very important to develop procedural knowledge (workcard design), checklists, and comparison standards to support this behavior.

Knowledge-based behavior is often a slow and error-prone process and creates a high cognitive workload for the human. Often in such circumstances the human will try to minimize cognitive strain by using shortcuts in the reasoning and decision making processes, which can lead to

suboptimal performance. Thus, we should try to design the system and the information environment to minimize the need to indulge in knowledge-based behavior. Knowledge-based behavior in visual inspection will be more evident in a novice inspector; this provides a strong impetus to the design of adequate training programs to bring the novice to expert levels and thus minimize knowledge-based behavior. Once a certain level of expertise is attained the knowledge-based behavior will be needed only in case of unfamiliar work situations. For example, this can happen if an inspector who normally works on only a specific part of the aircraft (e.g., the wing section) is asked to inspect a cargo door. Thus, it becomes important that the workcard (feedforward environment) be designed for usability and have the information needed to make a smooth transition to an unfamiliar task. Feedback information from a buddy system, and efficient communication lines with the supervisor, also have to be considered. Also important is the development of the knowledge about the spatial and functional aspects of the aircraft, which is partly built through the years of prior experience of the inspector as an aviation mechanic. This is normally five years but is decreasing due to a shortage of inspectors, with some inspectors having as little as three years of maintenance experience. There are cognitive error implications in too rapid a promotion system.

5.3.3.2.2 Non-Destructive Inspection

Moving to [NDI](#) inspection, skill-based behavior is predominant while using the probe and is a sensorimotor, feedback-controlled movement. This indicates the need for manual control training on tracking tasks (e.g., circle drawing, tracking) which transfer to this movement control task. Similarly, thought should be given to providing tracing paths (e.g., circles around rivets) which provide adequate feedback information. Templates can and are being used (although some inspectors do not like to use them due to handling difficulties) and the improved design and use of such aids should be encouraged. The rule-based behavior component of calibration points to the necessity of developing adequate and well designed checklists, along with procedural knowledge, for reliable performance. Swain and Weston (1988) point out that during the calibration procedures, powerplant technicians who very often have followed written steps, rely on memory and this increases the probability of omissions. This points to a calibration process design that is capable of providing cues to the next step on the display screen as well as detecting wrong inputs by the operator. Where calibration can be rigidly defined, the checklist is the obvious cognitive aid, already extensively used in aviation. Those calibration tasks which have some flexibility must be clearly delineated for separate treatment.

Display interpretation forms the critical portion of [NDI](#) and as such can be either rule-based, or knowledge-based, or both. The information environment should thus support both these behaviors while trying to ensure, through system design and training, that the need for knowledge-based behavior is minimized. Since rule-based behavior is based on signs which trigger stored patterns which in turn control our choices, Rasmussen and Vicente (1989) suggest that the design of the display should be such as to provide action cues as signs which also have symbolic content, thus supporting both rule and knowledge-based performances. Display screens for [NDI](#) that allow comparisons of the current pattern (curve) with known defect curves for comparative decision making should be considered. Also, the knowledge-based component found during display interpretation indicates the need to develop feedforward information (training and documentation) to provide technology knowledge, instrument knowledge, and aircraft defect history.

It must be emphasized at this point that in aircraft inspection, skill-based, rule-based, and knowledge-based behaviors are not necessarily stand-alone, discrete behavior modes. Indeed, they overlap on some occasions and support each other on others. For example, the skill-based behavior of probe movement is supported by either knowledge-based (for the novice) or rule-based (for the expert) behavior that ascertains the boundaries of the movement. For example, the probe should not cut the rivet head and a movement too close to an edge should be avoided since both of these will show defect indications without the presence of any defects. Similarly, rule-based behavior of defect identification and classification in visual inspection is sometimes supported by knowledge-based behavior that uses active reasoning based on a deeper and functional understanding of the aircraft.

For example, during visual inspection of the wing leading edge, the inspector who is looking for dents may reason that a dent forward of the aileron trim tab may be more important than one in another area because it could cause flow breakup in an area important to flight control. This and the preceding example highlight the often symbiotic relationship of the different behavior modes. Thus, while we concentrate on skill-based and rule-based behavior of the inspector (since these are the dominant behaviors), we also need to understand and support the knowledge-based behavior through adequate training schemes, documentation, and communications.

From the discussion above, it is evident that the mapping of the inspection processes to the [SRK](#) framework provides useful guidelines for, and a better understanding of, the type of information that has to be provided for aircraft inspection. This has been compiled in [Table 5.17](#) where the information categories (feedforward and feedback) identified in the aircraft inspection information model ([Figure 5.3](#)) have been assigned to the various inspection subtasks based on the type of behavior they would logically support.

INSPECTION PROCESSES	INFORMATION ENVIRONMENT	
	FEEDFORWARD	FEEDBACK
1. VISUAL (e.g. Rivet Inspection)		
• Pre-Attentive	• Training	
• Foveal Search	• Training	• Output Feedback
• Foveal Decision	• Training • Procedural Knowledge • Comparison Standards	• Cognitive Feedback • Buddy System
• Extra-Foveal	• Knowledge of Cues	• Feedback of Results
• Decision Making	• Co-Worker Information • Functional System Knowledge • Fault Causation Knowledge • Aircraft History (Defects)	• Communication Links • Buddy System • Cognitive Feedback
2. NDI (e.g. Eddy Current)		
• Calibration	• Checklists, Display Design	
• Probe Movement	• Training on Tracking and Accurate Movement Control	• Probing Aids (Templates or Markings Around Rivets)
• Display Interpretation	• Display Design • Functional System Knowledge • Technical Instrument Knowledge • Aircraft History	• Cognitive Feedback

Table 5.17 Information Requirements Identified from Mapping Inspection Processes to SRK

5.3.3.3 Analysis of Information Requirements: An Error Taxonomic Approach

In an analysis of 93 major accidents for a 24 year period from 1959 to 1983, Sears (1986) found that 12% were caused by maintenance and inspection deficiencies. Similarly, Nagel (1988) reports that approximately four out of every hundred accidents that occurred in the worldwide jet fleet from 1977 to 1988, had maintenance error as one of the chief causes. As shown in [Section 5.3.2](#), the effects of human error are becoming increasingly unacceptable and the issue of maintenance and inspection error is being closely examined and discussed in the aviation community (Drury, 1991).

Formulation of information environment requirements should include the notion of human error and its impact on aircraft inspection. Control of errors to an acceptable minimum is the implicit goal of all human-machine systems. In aircraft inspection, where the existence of certain defects in an aircraft ready to fly is almost unacceptable, it is pertinent to make this goal explicit, by defining information requirements based on human error avoidance. It can be argued that information provided at the right time, at the right place, in the right manner, is at least a necessary condition for

minimal error performance.

5.3.3.3.1 Methodology for Information Requirement Formulation

Human error can serve as an effective platform to study and formulate the information requirements of aircraft inspection just as it was used in [Section 5.3.2](#) to understand the overall inspection process. We present below a methodology that attempts to guide the design of the information environment to controlling human error:

1. Identify and define the levels of the system under consideration (e.g., management, supervisory, lead inspector, inspector).
2. At the level under analysis, define the functional requirement of the level, current allocation of human-computer functions, and interactions with the other levels.
3. Develop a human error taxonomy for the level under consideration.
4. Use the taxonomy and the functions identified in step 2 to outline the failure modes (phenotypes) and associated mechanisms of human malfunction and error shaping factors (geno-types) specific to each function.
5. Identify the component of the information system that would be necessary to control human error based on understanding of the phenotypes and genotypes of step 4.
6. Define the requirements of each information component: (1) what information to present (information quality); (2) when to present such information (information flow); and (3) how to present this information (information display), so that the human error potential is minimized.

The above methodology combines a task analytic approach with a human error taxonomy so that information requirements are formulated to control human error. Obviously, the error taxonomy development is an important part of this approach. A framework or guideline is presented, which can be used to develop a taxonomy for use in this methodology.

Rasmussen and Vicente (1990) suggest that human error analysis can be performed from two different perspectives. The first perspective tries to identify possible human errors and their effects on system performance, while the second perspective aims at improving system design to eliminate the effects identified in the analysis from the first perspective. Based on the first perspective, Drury (1991) developed an error taxonomy from the failure modes of each task in aircraft inspection. This taxonomy has been developed based on the recognition that a pro-active approach to error control is needed to identify potential errors. Thus, the taxonomy is aimed at the phenotypes of error (Hollnagel, 1989), i.e., the way errors are observed or appear in practice. In [Section 5.3.2](#) it was also noted that Rasmussen and Vicente (1990) propose a taxonomy from the viewpoint of identifying possible improvements in system design with categories of errors as related to: (a) effects of learning and adaptation, (b) interference among competing control structures, (c) lack of resources, and (d) stochastic variability. They suggest that different methods have to be adopted to control the errors associated with each of the above four categories, and that it is necessary to make the system error-tolerant to achieve reliable system performance.

We propose that the failure modes identified in the taxonomy of aircraft inspection by Drury (1991) can be classified using the systemic error mechanisms categories and the cognitive control categories proposed by Rasmussen and Vicente (1989). (An example is given in [Table 5.18](#) for error modes in the decision task.) In [Table 5.19](#), such an assignment is shown using the failure modes for the decision task. For each behavior mode (i.e, skill, rule, or knowledge) the genotypes of errors can be then postulated. Genotypes are the contributing psychological causes of errors and are representative of the characteristics of the human cognitive system (Hollnagel, 1989). [Table 5.20](#) shows the genotypes assigned to the different behavior modes.

TASK	ERROR(S)
TASK 4 -- DECISION	
4.1 Interpret indication	4.1.1 Classify as wrong fault type.
4.2 Access measuring equipment.	4.2.1 Choose wrong measuring equipment. 4.2.2 Measuring equipment is not available. 4.2.3 Measuring equipment is not working. 4.2.4 Measuring equipment is not calibrated. 4.2.5 Measuring equipment has wrong calibration. 4.2.6 Does not use measuring equipment.
4.3 Access comparison standards.	4.3.1 Choose wrong comparison standard. 4.3.2 Comparison standard is not available. 4.3.3 Comparison standard is not correct. 4.3.4 Comparison standard is incomplete. 4.3.5 Does not use comparison standard.
4.4 Decide on fault presence.	4.4.1 Type 1 error, false alarm. 4.4.2 Type 2 error, missed fault.
4.5 Decide on action.	4.5.1 Choose wrong action. 4.5.2 Second opinion if not needed. 4.5.3 No second opinion if needed. 4.5.4 Call for buy-back when not required. 4.5.5 Fail to call for required buy-back.
4.6 Remember decision/action.	4.6.1 Forget decision/action. 4.6.2 Fail to record decision/action.
OUTCOME 4: All indications located are correctly classified, correctly labelled as fault or no fault, and actions correctly planned for each indication.	

Table 5.18 Task and Error Taxonomy for Inspection, Task 4, Decision

SYSTEMIC ERROR	BEHAVIOR CATEGORIES		
	SKILL-BASED BEHAVIOR	RULE-BASED BEHAVIOR	KNOWLEDGE-BASED BEHAVIOR
1. Effects of Learning and Adaptation	4.2.1, 4.2.4, 4.2.6, 4.3.1, 4.3.5, 4.5.1, 4.6.2	4.2.4 Fail to calibrate meas. eqpt. 4.2.6 Fail to use meas. eqpt. 4.3.5 Fail to use comp. std. 4.6.2 Fail to record decision. 4.2.1 Choose wrong meas. eqpt. 4.3.1 Choose wrong comp. std. 4.5.1 Choose wrong action. 4.2.5 Meas. eqpt. wrongly calib. 4.1.1 Classify as wrong fault. 4.4.1 False alarm. 4.4.2 Missed fault.	4.1.1 Classify as wrong fault. 4.4.1 False alarm. 4.4.2 Missed fault. 4.5.2 Wrong decision on getting. 4.5.5 Second opinion. 4.5.4, 4.5.5 Wrong decision on calling for buyback. 4.2.5 Wrong calibration of meas. eqpt.
2. Interference Among Competing Control Structures		4.2.5 4.1.1, 4.4.4, 4.4.2 4.5.1 4.5.2, 4.5.3, 4.5.4, 4.5.5	4.1.1, 4.4.1, 4.4.2, 4.5.2, 4.5.3, 4.5.4, 4.5.5
3. Lack of Resources			4.6.1 Forget decision 4.5.2, 4.5.3 4.5.4, 4.5.5
4. Stochastic Variability		4.1.1, 4.4.1, 4.4.2	4.6.1

Table 5.19 Assignment of Systemic Error Mechanisms to Failure Based on Behavior Types Decision Making Component of Aircraft Inspection

BEHAVIOR TYPE	FAILURE MODE (PHENOTYPES)	MECHANISMS OF HUMAN MALFUNCTION AND ERROR SHAPING FACTORS (GENOTYP)
KNOWLEDGE-BASED	4.1.1 Classify as wrong fault. 4.4.1 False alarm. 4.4.2 Missed fault.	<ul style="list-style-type: none"> o Selectivity o Incomplete mental model o Lack of knowledge o Confirmation bias
	4.5.2, 4.5.3 Wrong decision on getting second opinion. 4.5.4, 4.5.5 Wrong decision on calling for buy-back.	<ul style="list-style-type: none"> o Overconfidence o Incomplete knowledge o Biased reviewing
	4.2.5 Measuring equipment wrongly calibrated.	<ul style="list-style-type: none"> o Overconfidence o Incomplete knowledge
RULE-BASED	4.2.4 Fail to calibrate measuring equipment.	<ul style="list-style-type: none"> o Recall error o Memory slip
	4.2.6 Fail to use measuring equipment. 4.3.5 Fail to use comparison standards.	<ul style="list-style-type: none"> o Recall error o Overconfidence o Memory slip o Mindset or rigidity
	4.6.2 Fail to record decision.	<ul style="list-style-type: none"> o Recall error, memory slip o Availability o Mindset o Wrong rules
	4.2.1 Choose wrong measuring equipment. 4.3.1 Choose wrong comparison standards. 4.5.1 Choose wrong action.	<ul style="list-style-type: none"> o Stereotype takeover o Memory slip o First exceptions o Availability
	4.2.5 Measuring equipment wrongly calibrated.	<ul style="list-style-type: none"> o Familiar shortcut o Misinterpretation
	4.1.1 Classify as wrong fault. 4.4.1 False alarm. 4.4.2 Missed Fault.	<ul style="list-style-type: none"> o Encoding deficiency o Familiar pattern not recognized
SKILL-BASED	4.2.1, 4.2.4, 4.3.1, 4.3.5, 4.2.6, 4.5.1, 4.6.2	<ul style="list-style-type: none"> o Omissions o Recency and frequency of use o Speed accuracy tradeoff

The above framework, then, allows the opportunity to examine each failure mode within the context of (a) the cognitive behavior from which it results, (b) the systemic error category in which it occurs, and (c) the internal error mechanisms that are the probable causes of these malfunctions. An analysis of this information can then form the basis of system design to minimize or eliminate the failure modes. From the information requirements viewpoint, system design considerations should then drive the specifications as to the type, location, and temporal position of the information. Preliminary recommendations on the type of information component have been listed in [Table 5.20](#). In actual use, [Table 5.20](#) should be utilized as a framework for an error taxonomy which can be applied in the task analysis methodology proposed.

5.3.3.4 Testing the Information Framework

Using the inspection program developed for [NDI \(Section 5.3.1.1\)](#) it is possible to make direct experimental tests of many of the predictions coming from the framework being developed in [Sections 5.3.3.2](#) and [5.3.3.3](#). As a demonstration of the use of the [NDI](#) inspection program, a relatively simple experiment based on the information requirements was conducted. It involved

training two groups of subjects on the inspection task, then either providing or not providing off-line feedback of performance, and finally measuring inspection performance of both groups.

As shown in [Section 5.3.3.1](#), on-the-job feedback can be a powerful performance enhancer, but it is an expensive one to implement. It involves re-inspection of an inspector's work by a (presumably more reliable) auditor, a process which adds cost in proportion to the percentage of work audited. A more realistic approach would be to provide feedback, for example by having the inspector inspect a test piece with a known set of faults, between regular inspection tasks. Feedback can easily be provided from such a test piece, but we need to measure the effectiveness of such feedback. A test of this effectiveness also provided a useful practical test of the [NDI](#) program, and indeed many pilot subjects were run and program modifications were made before the complete experiment reported here was started. The following is a brief description of the experiment and its results. These results are being presented in more detail in a separate project report.

[5.3.3.4.1 Methodology](#)

Two groups of eight subjects each were chosen randomly from a population replying to advertisements. All were currently unemployed members of the work force, with males and females and a variety of ages represented. Each subject was given two pre-tests, both of which had been shown to correlate with performance on industrial inspection tasks. The first was the Embedded Figures Test (EFT) which classifies the cognitive style of a person as Field Dependent (i.e., highly influenced by the visual context of a task) and Field Independent (i.e., more able to cognitively restructure a task independent of its visual context). The second test was the Matching Familiar Figures Test (MFFT) which measures the tendency of subjects to opt for speed or accuracy in their speed/accuracy tradeoff (see [Section 5.3.4](#)). Foveal visual acuity was also measured.

Both groups were given the same training in the principles of eddy-current inspection of rivets for cracks, in controlling of the pointer using the mouse, and in interpreting the meter needle movements. This training occupied about four hours. Following training, the eight subjects in the control group were tested on a task involving 420 rivets followed by a task involving 80 rivets, on each of four days. The experimental group was given the same task except that they were provided with feedback on the missed cracks, false alarms, and performance time on the 80 rivet task.

In the main task, the same measures of misses, false alarms, and task time were taken for each subject.

[5.3.3.4.2 Results](#)

Analyses of covariance were performed on the measures of total time, misses, false alarms, and derived measures from Signal Detection Theory ([Section 5.3.5](#)) of sensitivity (d') and criterion (X_c). Each analysis tested for differences between the two groups (G), for differences between the four days (D), as well as for their interaction (D X G). Two sets of covariates were derived from factor analysis to contain the following components:

Covariate 1: [EFT](#) Errors, [EFT](#) Times, [MFFT](#) Errors

Covariate 2: [MFFT](#) Time (negative), Visual Acuity

Covariate 1 represents poor accuracy performance and field dependence, while Covariate 2 represents fast performance with good vision. [Table 5.21](#) summarizes the analyses of covariance of the measures taken. There were no significant group effects, and only a single day effect, that on total time for the task. Covariate 2 was significant for total time and for criterion X_c . [Figure 5.8](#) shows plots of the results for times, misses, false alarms, and sensitivity (d') comparing the experimental and control groups across the four days of the experiment.

Measured Analyzed	Groups (G)	Days (D)	G X D	Covariate 1	Covariate 2
Total Time	---	P < 0.0001	---	---	P = 0.0235
Misses	---	---	---	---	---
False Alarms	---	---	---	---	*
Sensitivity (d')	---	---	---	---	---
Criterion (X ₀)	---	---	---	---	P = 0.0355

* Indicates that the Visual Acuity component of covariate 2 was significant at P < 0.0244.

Table 5.21 Summary of Analyses of Covariance for Off-Line Feedback Experiment

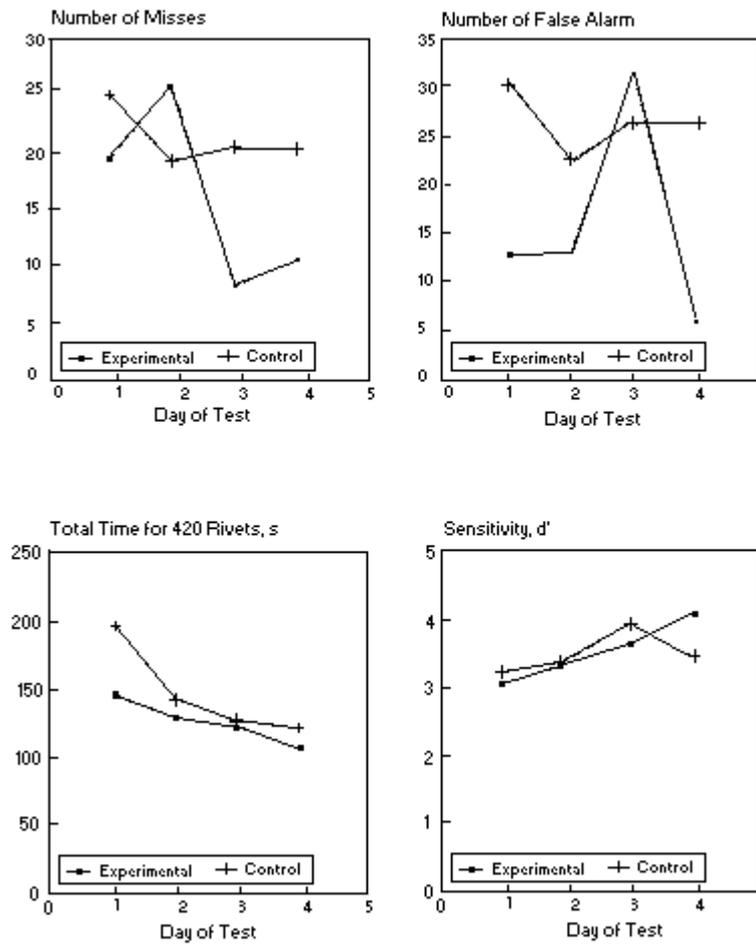


Figure 5.8 Day X Group Interactions for Off-Line Feedback Experiment

5.3.3.4.3 Discussion

The major finding of this first experiment using the [NDI](#) program was one of very high between-subject and day-to-day variability. The between-subject variability was expected, and it appears that some of this variability at least is predictable using the covariates derived. Because of this variability, the effects shown in [Figure 5.8](#) do not reach statistical significance with only eight

subjects per group. Having said this, there is an indication in all four parts of [Figure 5.8](#) that the experimental group outperforms the control group by the end of the experiment.

Over the four days of the experiment, accuracy performance, as measured by misses and false alarms, improved slightly for the control group and somewhat more for the experimental group. Despite this overall improvement, the day-to-day improvement was erratic. Total time decreased for both groups, with the experimental group being more rapid than the control group throughout. Sensitivity, as defined in Signal Detection Theory, marginally favored the control group until the final day, when the experimental group continued to improve while the control group regressed slightly.

During the course of the experiment, it was clear that the experimental group was using the off-line feedback to modify their inspection strategy. However, this process involved trial and error, which gave considerable variability of performance. The performance feedback helped somewhat, but would have been much easier to interpret if it had contained hints and steps that the inspectors could take to make the improvements they knew were needed. Cognitive feedback, as postulated in [Section 5.3.3.1](#) appears to be required if inspectors are to make use of their own performance data.

[5.3.3.4.4 Conclusions](#)

While off-line performance feedback was marginally effective, the high variability between subjects prevented significant results from being obtained. At least part of the day-to-day variability was due to subjects using the feedback in an unguided manner in an attempt to improve, suggesting that cognitive feedback may be needed to supplement off-line performance feedback. The small size of the feedback task (80 rivets) might also have failed to provide sufficient data to significantly aid in transfer of feedback results. The significant covariates for total time and criterion also indicate influence of other independent factors, namely visual acuity and cognitive style.

5.3.4 A FRAMEWORK FOR SPEED/ACCURACY TRADEOFF IN AIRCRAFT INSPECTION

In almost any discussion with aircraft maintenance personnel, maintenance managers, regulatory bodies, or the travelling public, the general issue of inspection accuracy arises. More specifically, in the post-deregulation environment of U.S. commercial aviation, the effect of time pressures on the inspection system (particularly the human inspector) is causing concern. This section reviews the functions and tasks of aircraft inspection, based upon a two-year observational study of the system, and uses prior studies of human inspection to examine the possibilities of time pressure affecting accuracy. A Speed/Accuracy Tradeoff (SATO) perspective is taken, i.e. how do speed and accuracy co-vary in inspection.

Both speed and accuracy are relatively easy to define in inspection.

Speed: The rate of inspecting items, usually measured as the reciprocal of the time (t) taken to inspect a single item or defined area.

Accuracy: False Alarm (Type 1 error) The probability of an inspector responding that a defect exists, when in truth it does not.

Miss (Type 2 error) The probability of an inspector failing to respond that a defect exists, when in truth it does exist.

This section is concerned explicitly with the co-variation of (t), False Alarms, and Misses.

From an airline management perspective, two goals need to be achieved by the system: safety and profitability. The profitability goal can only be achieved by first ensuring that the safety goal is achieved economically. These objectives are passed through sometimes complex organizational systems (Taylor, 1990) to supervisors and finally to inspectors. At the inspector's level two goals need to be achieved by the inspection system: accuracy and speed. Accuracy means detecting those

indications (faults) which must be remedied for the safe operation of the aircraft while not activating the maintenance system for non-faults. Speed means the task must be performed in a timely manner without the utilization of excessive resources. These two criteria of the inspection system can be expected to be inversely related at the inspection level (Drury, 1985).

When inspection is split into its task steps ([Table 5.1](#) of [Section 5.1](#)), it can be seen that all of the tasks require both speed and accuracy for their completion. However, the most error-prone activities in industrial inspection are the search and decision making tasks (Drury, 1984) while access is an activity whose time must be minimized for efficient operation.

The speed and accuracy with which each of the components is performed depends upon the relative utilities of the various outcomes to the inspectors. Utility is a concept that can be used in models of the inspector as a maximizer or optimizer (Drury, 1992) to give a normative model as a starting point for more realistic inspector models. Thus, the optimum speed and accuracy is not defined in terms of minimizing or maximizing one particular aspect of inspection but is defined in terms of a performance which yields the highest overall utility.

If a task can be performed at various levels of speed and accuracy, then it is possible (Wickens, 1984) to generate an operating characteristic curve (see [Figure 5.9a](#)) relating the two measures. Any point on the speed/accuracy operating characteristic (SAOC) curve shows the accuracy with which the task can be performed at a particular speed. Hence to meet the designed system objectives of speed and accuracy, it is essential that the inspectors operate at the correct point on the correct operating-characteristic curve.

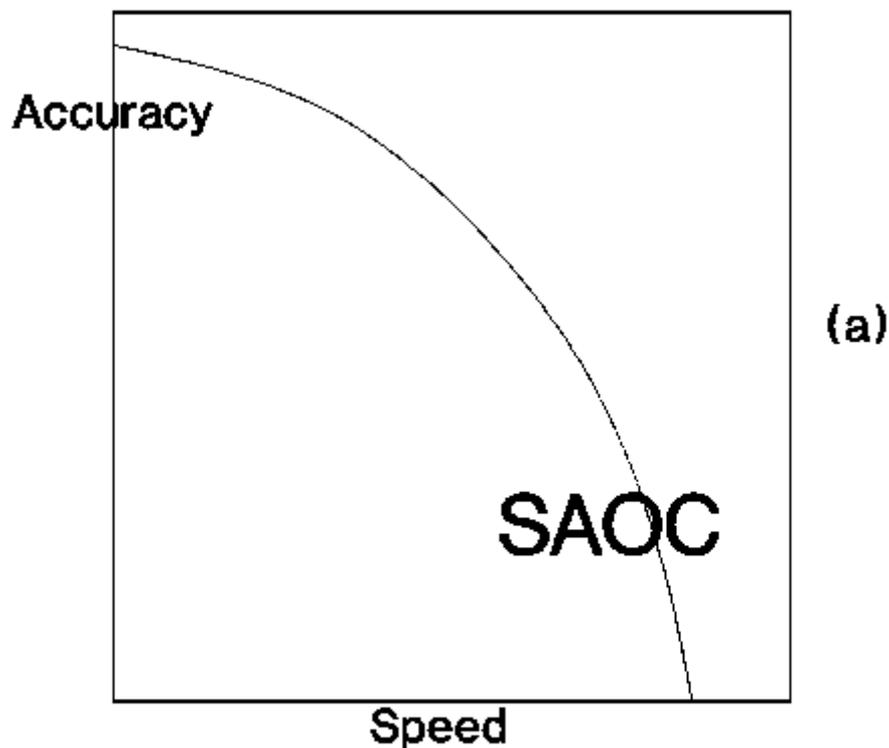


Figure 5.9a Generalized Speed/Accuracy Operating Characteristics (SAOC)

In order for an inspector to choose a particular strategy from the set of available strategies, it is necessary to determine the utility of all the candidate strategies, as a function of speed and accuracy. The utility can be computed for every point in the joint performance space (Speed, Accuracy), whether that point is achievable or not. Now by knowing the utility function it is possible to determine the optimal operating point, i.e. that which maximizes the expected utility. Typically,

contours of equal utility are superimposed upon the SAOC to show where this optimal operating point occurs (see [Figure 5.9b](#)). This section considers access, search, and decision making in turn, and uses models of each to show the form of the SAOC. Models are not developed in detail. For more information the original report (Drury and Gramopadhye, 1991) can be consulted. Each model is an optimization model, showing how an inspector may be expected to choose between alternative strategies. In large decision tasks, there is considerable evidence that they are satisfiers rather than maximizers (Wickens, 1991). However, in small task components such as those found in inspection, optimization models represent a good starting point for consideration of the factors involved (Drury, 1988; Chi, 1990).

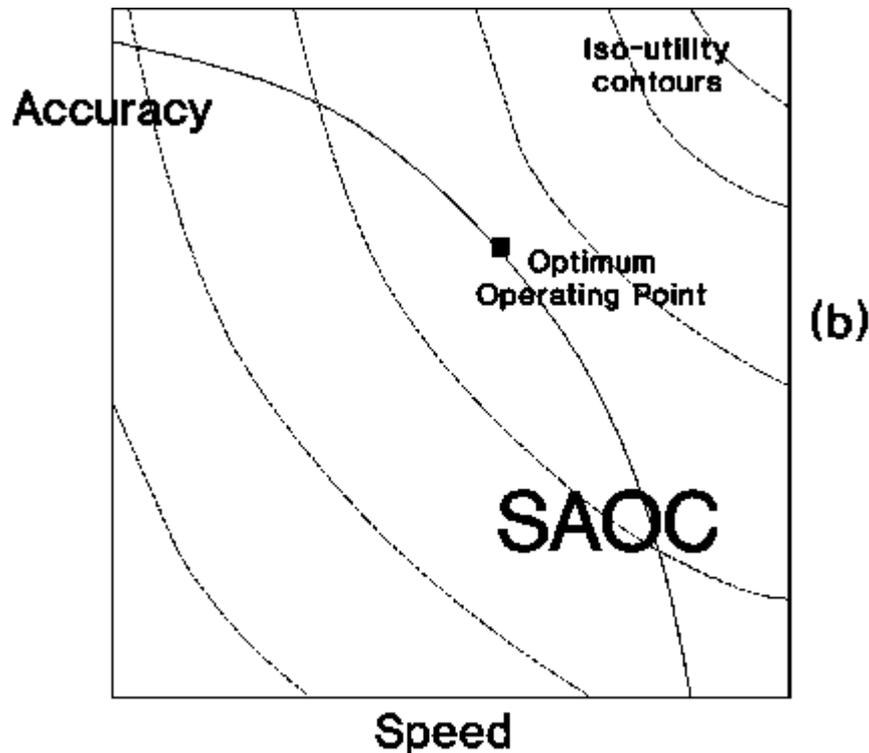


Figure 5.9b Generalized Speed/Accuracy Operating Characteristics (SAOC)

5.3.4.1 Factors Affecting Access Tasks

The access task consists of physically reaching the area to be inspected. 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 intervening structure (e.g., inspection of interiors of wing fuel tanks through access holes). All of these activities involve controlling the movement of the inspector's body, or body parts, within a restricted space. In general, control theoretic models of the human operator in control tasks (Sheridan and Ferrell, 1974; Wickens, 1984) show that as more speed is demanded, tracking accuracy decreases--a typical SATO. If the access task is modelled as moving accurately between two boundaries without making an error of exceeding a boundary, then the self-paced tracking models of Drury (1971) and Montazer, Drury and Karwan (1987) can apply. Thus, moving the body along the walkway of a scaffold without hitting (and possibly damaging) the aircraft structure on one side or the scaffold rail on the other is such a task. Moving the hand (or head) through an access hole or moving a cherrypicker along the fuselage upper skin, (although only one physical boundary exists here) are further examples.

The self-paced tracking model considers the inspector, or a vehicle controlled by the inspector, as choosing a speed which will maximize the utility to the inspector. Utility is composed of rewards for

speed and penalties for error, in this case the error of exceeding the fixed boundaries. Model results, and experimental data from a variety of studies, have shown that the speed chosen increases with space available (e.g., width) until some limiting speed is reached. The three factors affecting performance are thus space available, the ease of control (controllability) of the vehicle, and the inspector's perception of the relative utilities of speed and accuracy. Each will be considered in turn.

5.3.4.1.1 Space Available

Space available can be controlled relatively easily around the aircraft, but access within the airframe itself is largely determined at the design stage. With older aircraft there has been a history of unpleasant surprises for maintenance personnel when they reached service, but manufacturers are now using computer-manipulable human anthropomorphic models (e.g., CREWCHIEF, SAMMIE) to examine access for maintenance before structures are finalized. Note, however, that the SATO model shows that more space improves performance, so that the minimum necessary for physical access (e.g., for a 95th percentile male) will not provide optimum performance. A human anthropomorphic model only gives the space required for a person to statistically assume a posture. For movement (the essence of maintenance and inspection) more access room is required beyond this minimum. The same considerations apply to access around the aircraft. Steps and walkways should be made wide enough to provide unhindered movement, not just wide enough to accommodate a large static human. As an example, Drury (1985) reports that for movement through a doorway both performance time and errors decrease from the anthropometric minimum width of about 20 inches to the unhindered width of 36 inches. Very similar findings are used in the aviation industry to determine sizes of emergency doors for passengers.

During Phase III of this work, explicit models and experiments will be developed to test the effects of space available, and human posture, on performance and stress in inspection and maintenance activities.

5.3.4.1.2 Controllability

Controllability of the system having access is a major determinant of access performance. For most tasks, the "system" is the inspector's own body, the most naturally controllable system. However, controllability can be adversely affected by equipment carried (flashlight, tools, work cards, NDI equipment) and by the quality of clothing worn. Thus, coveralls and shoes should be minimally restrictive. Shoes should also provide good grip on a variety of surfaces under both wet and dry environmental conditions. Controllability will be decreased by any impairment of the human, for example sickness, alcohol, or drugs, reinforcing the control required over such conditions at the work place.

For control of systems such as vehicles (e.g., cherrypickers, wheeled steps, moveable access scaffolds) a considerable body of information exists (e.g., Wickens, 1984) on the human as controller. Most of these recommendations apply equally to the self-paced access tasks considered here. Thus for example, controls should move in the same directions and sense as the element they control. It should be noted that many cherrypickers have hydraulic or electrical controls which violate this principle. Direction of motion errors are to be expected with such systems, causing at best a slowing of the task and at worse damage to the aircraft structure, depending upon the operator's SATO choice. These same controls are often not progressive in operation, but "bang-bang" controls, either fully on or fully off. With such a degraded control system, any designed speed setting is a poor compromise. At times it is too slow, causing delay and frustration in making long movements, while at other times it is too rapid, causing errors and time-consuming multiple corrections in making the final accurate positioning movements. In addition, any time lags or inertia in the system controlled will have a negative impact on controllability.

Within the maintenance hangar, there are other constraints on design (or choice) of access equipment. Any equipment must be available if it is not to cause delays, suggesting both that a

sufficient supply exists, and that it is well-scheduled. The difficulty with maintaining a sufficient supply is that such equipment is both expensive and space-consuming. The typical management response is to have a mixture of special-purpose equipment, such as empennage access scaffolding, and standardized, flexible equipment, such as stepladders, cherrypickers, and standard moveable platforms. When only a single aircraft type is to be serviced, as in most large airlines and specialist repair centers, purpose-built equipment should, and does, predominate. In more general purpose organizations, the emphasis is on standardized, flexible equipment. However, there are still times when schedules demand more access equipment than is instantaneously available. It is at these times that available equipment is substituted for correct equipment to avoid delays. The result is lower system controllability, with the potential for errors affecting both job performance and personnel safety.

5.3.4.1.3 Perception of Utilities

Given the space available and the controllability of the system, the balance between speed and accuracy is still finally chosen by the operator's own [SATO](#). As discussed earlier, this is where any gate pressures or schedule demands can have an effect. As access is a task of inspection which appears non-critical, it can be one where time is saved for tasks perceived as more important. In addition, access is where pressures from other members of the maintenance team can be acute. Co-workers will at times need the access equipment the inspector is using or *vice versa*, leading to time pressures over a short time scale even where none exist on the longer-term scale of a whole maintenance visit.

Inspectors' errors in access are defined as reaching or exceeding the boundary of available space. They thus include both damage to aircraft structure, and injury to the inspector. Humans are likely to misperceive the risks associated with such rare events, both in terms of the consequences and probabilities involved. Particularly with highly experienced personnel, such as inspectors, the probabilities of error are typically rated lower than their objective values. This can be expected to lead to a choice of [SATO](#) strategy favoring speed rather than accuracy.

5.3.4.2 Factors Affecting Search Tasks

The process of visual search of an extended area, such as the area called out on a workcard, has been successfully modeled since the start of human factors engineering. A human searcher (e.g., the inspector) makes a sequence of fixations, centered on different points in the area. During a fixation, which typically lasts 0.25-0.5 seconds, the inspector can detect defects in an area, called the visual lobe, around the fixation center. Between fixations the eye moves very rapidly and can take in very little information. The sequence of fixations can either be modelled as random (e.g., Krendel and Wodinski, 1960) or systematic with repeated scans (e.g., Williams, 1966). For both of these models, equations can be developed relating the probability of detection to the time spent searching (Morawski, Drury and Karwan, 1980).

In general, the longer an inspector searches an area, the greater the probability of a target being located, with diminishing returns as search time is increased. Such curves are the [SAOC's](#) of visual search, and are shown in [Figure 5.10](#). Given such [SAOC's](#), then the optimum time for searching can be calculated (Morawski, Drury and Karwan, 1992) based upon the reward for speed and the penalty for error.

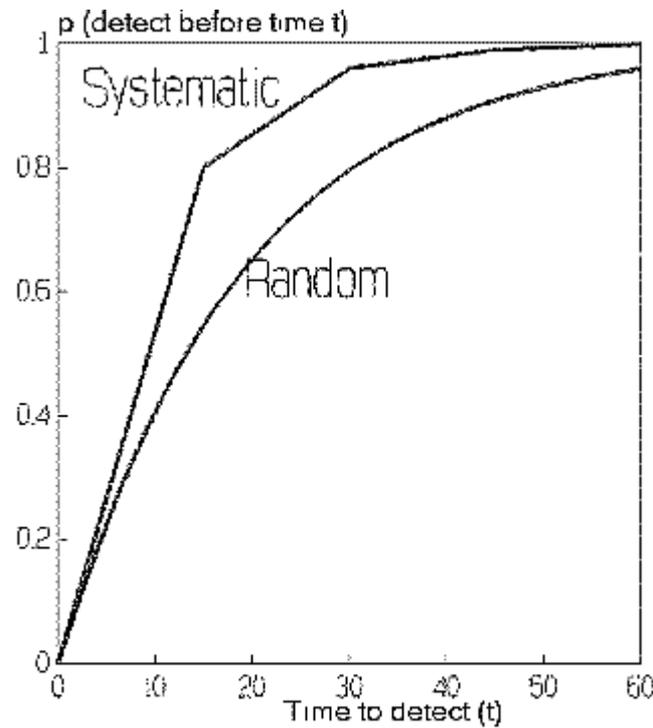


Figure 5.10 Typical Cumulative Search Time Distribution. Giving the SAOC for Visual Search

From the visual search models, three groups of factors determine search performance:

1. Factors affecting the visual lobe.
2. Factors affecting the search strategy.
3. Factors affecting the [SATO](#) and stopping policy.

Based on the defect type, severity level, and location, the defects can be classified into critical and non-critical defects. Critical defects are those defects which affect the airworthiness of the aircraft, hence whose detection is critically important. Non-critical defects do not immediately affect the airworthiness of the aircraft but have to be detected in the long run. There is clearly a heavy penalty for missing critical defects, but the entire area needs to be searched for both critical and noncritical defects within a specified time period. Thus, two goals need to be achieved by the inspector, speed and accuracy, for which the inspector needs to be efficient as well as effective. In order to understand the Speed/Accuracy Tradeoffs in search where the inspector is looking for multiple defect types, the factors which affect this tradeoff and indeed the whole search process must be examined.

5.3.4.2.1 Visual Lobe Factors

According to Engel (1971) fault conspicuity is defined as that combination of properties of a visual object in its background by which it attracts attention via the visual system, and is seen as a consequence. Monk and Brown (1975) have shown that mean search times increase as a function of the number of non-targets in the target surroundings. They have also shown that isolated targets are more easily detected than those surrounded by non-targets. Williams (1966) has shown that the color and size of the targets can be used by subjects to direct their eye movements. Studies of information processing within a single fixation have shown that the probability of target detection increases with increased target size and brightness contrast, and decreases with angular distance from the fixation point (Overington, 1973). This decrease with off-axis angle provides the basis for determining visual lobe size, i.e., the area within which a target may be detected (Bloomfield, 1975).

In aviation, this search performance has been extensively studied and modeled to determine human performance in detection of military targets (for example, ground targets or hostile aircraft). In terms of aviation maintenance inspection, the implication is that lighting and other target/background amplification devices should be used to make the conspicuity of a defect as high as possible, and hence increase visual lobe size.

There are, of course, individual differences in visual lobe size. Eye movement studies have shown that subjects who have larger visual lobes are more efficient, or they detect targets (faults) earlier on in the search process (Schoonard, et al., 1973; Boynton, 1960). Johnston (1965) provided evidence to suggest that subjects who obtain high peripheral acuity scores exhibit relatively shorter search time. There is evidence that the visual lobe size is amenable to training (Gramopadhye, Palanivel, Knapp, and Drury, 1991). There is no evidence that better inspectors have shorter fixation times, only that they make fewer fixations, presumably because of the larger visual lobe size.

The implication for aviation inspectors is that individual differences may be quite large, but are amenable to training. Other evidence (Gallwey, 1982; Drury and Wang, 1986) suggests that selection tests for visual lobe size may well be task-specific, in that the ability to search for defect (D) in background (B) may be unrelated to the ability to search for a different defect (D') in a different background (B'). As Drury and Gramopadhye (1990) have noted, training appears to be a more powerful intervention strategy than selection for inspection tasks.

5.3.4.2.2 Search Strategy Factors

As noted earlier in this section, search strategy can be modeled as random or systematic, with humans believed to lie in between these two extremes. A systematic search strategy is always more efficient than a random strategy. Scanning strategy is dependent on an inspector's:

1. Familiarity with the task (experience).
2. Ability to obtain and utilize feedforward information from cues regarding defect locations and defect types uncertainty). Gould and Carn (1973) and Monk (1977) have shown that in tasks which do not lend themselves readily to the adaptation of systematic search strategy, search times increase with increased fault uncertainty.

Search strategy in visual search is a global term which reflects many parameters of saccadic movement. The speed with which search is performed is dependent on the eye movement parameters, such as those listed by Megaw and Richardson, 1979: fixation times, spatial distribution of fixation, interfixation distance, duration of eye movements, and sequential indices. Fixation times have already been considered in the previous section on visual lobe factors.

Inspectors do not have uniform coverage of the area inspected (Schoonard, et al., 1973), with the central portions given more attention than the edges. In addition, inspectors may not always choose a correct distance between successive fixation centers (Gould and Schaffer, 1967; Megaw and Richardson, 1979). The scan path of an inspector changes with experience (Kundel and Lafollette, 1972; Bhatnager, 1987) to reflect a more consistent path, more even coverage, and more coverage where there is a higher probability of a fault being located.

The studies of search strategy are not conclusive on how to take practical steps to improve that strategy, although they do point to structuring of the search field as a way to increase the likelihood of systematic search.

With a structured field, the current fixation point will serve as a memory aid to which areas have already been searched. Suitable structuring devices may be panel lines, physical elements of complex parts (doors, landing gear), or superimposed temporary structures, such as inspectors' markings on aircraft.

Any such structuring lines should be made clear on the graphics included with workcards, and in any training materials.

There are likely to be large individual differences in search strategy, differences which are relatively stable over time. The issue of training of search strategy is the subject of one of the experiments presented in the training section ([Section 5.3.5](#)).

[5.3.4.2.3 SATO and Stopping Policy Factors](#)

Choice of operating point on the [SAOC](#) is determined by the perceived utilities of speed and accuracy. The only error possible on a search task is a Miss, so that high accuracy implies locating all potential defects in the structure. Inspectors are highly motivated for accuracy, as noted earlier ([Shepherd, et al., 1991](#)), so that one would expect an operating point on the [SAOC](#) representing long search times, with repeated search being common. In practice, inspectors appear to stop at the end of a single scan of the area, only repeating a fixation if some indication has been found. It appears that inspectors recognize the "diminishing returns" aspect of search performance, and are confident enough in their abilities that a single scan at the appropriate level of detail is seen as optimal. Such a policy certainly reduces the memory load and potential vigilance effects associated with multiple scans. However, the inspector will need to be "recalibrated" at periodic intervals by retraining or by providing test sessions to ensure that the speed of inspection chosen is appropriate to the accuracy demanded.

[5.3.4.3 Factors Affecting Decision Making](#)

Decision making is the task during which any potential defect (indication) located by the search task is evaluated to determine whether it should be reported. In this task both Type 1 errors (False Alarms) and Type 2 errors (Misses) can occur. These have their own tradeoff relationship, so that some combined accuracy measure must be derived before any tradeoff between speed and accuracy can be considered.

One particular model of the human as a rational economic maximizer which has received widespread support in inspection is Signal Detection Theory (SDT). Originally proposed by Swets and various co-workers (e.g., Swets, 1967) as a model for how humans detect signals in noise, it was subsequently applied successfully to inspection (Wallack and Adams, 1969, 1970; Sheehan and Drury, 1971; Drury and Addison, 1973).

In the [SDT](#), the inspector is assumed to be making a choice for each item inspected of whether the item contains a defect ("signal") or does not ("noise"). As the evidence for signal or noise is somewhat equivocal, there is assumed to be an "evidence variable" which increases when a signal is present and decreases when only noise is present. An example would be the judgement of whether a dent in a stabilizer leading edge should be reported. Dents can range from almost imperceptible to obviously reportable. The evidence variable (dent visual severity) must be judged against both written size standards and the likely effect of the dent on flight characteristics.

[SDT](#) shows that the two error probabilities, p (miss) and p' (false alarm), can be derived from a model in which the inspector chooses a criterion (X_c) to report on the presence of a defect. As this criterion varies from high (defects rarely reported as present) to low (defects often reported as present), an Operating Characteristic Curve is traced out. This curve has become known as the Receiver Operating Characteristic (ROC) in [SDT](#) literature. A different [ROC](#) curve is traced out for different levels of signal/noise ratio, known as discriminability and symbolized by d' .

Wickens (1984) has divided tasks into those which are resource limited and those which are data limited. In the former tasks, as the operator brings more resources to bear on a problem (e.g., devotes greater time to it) performance improves. In a data limited task, the quality of the data received by the operator is the limiting factor, so that more resources yield no better performance. It appears that [SDT](#) tasks are only resource limited up to short times, after which they are data limited. Because aircraft inspection is typically a matter of minutes and hours rather than seconds, a reasonable assumption is that its decision making aspects are data limited. Thus there is unlikely to

be a marked [SATO](#) for decision making during the inspection task. However, the grosser aspects of decision may still show a [SATO](#). For example, if the inspector is unable to reach a decision, the supervisor (or other senior personnel) may be called in to assist. Here the inspector is attempting to improve accuracy at the cost of increased time.

From the [SDT](#) model, there are three groups of factors which can affect the overall speed and accuracy:

1. Discriminability or sensitivity.
2. Choice of criterion.
3. Choice of [SATO](#) operating point.

[5.3.4.3.1 Factors Affecting Sensitivity](#)

Most factors affecting discriminability or sensitivity are physical, and can be characterized as the perceived difference between the observed indication and a standard. Thus, indications obviously well above or below the standard will have high d' values. Examples would be large areas of corrosion, cracks noticeably larger than those allowed, or completely missing rivets. None would require difficult (i.e., error prone) decisions. But "perceived difference" implies both high signal and low noise in [SDT](#) terminology. Low noise means low levels of visual distraction (i.e., competent cleaning), low levels of fatigue (i.e., frequent task breaks), and very clear standards (i.e., well-defined and well-presented job aids). All of these can be improved in aircraft inspection.

Comparison standards at the work place have been shown to be effective in improving discriminability (Drury, 1990b). It should be possible for the inspector to make a direct side-by-side comparison of an indication with a standard. For example, the critical amount of corrosion beyond which a report must be made should be indicated by a life-sized diagram on the workcard. Also, if different corrosion types are present, life-sized photographs help in positive identification (Harris and Chaney, 1969).

[5.3.4.3.2 Factors Affecting Criterion](#)

From [SDT](#), the two factors affecting the choice of criterion are the relative costs of errors (misses and false alarms) and the true rate of defects (p'). From these factors, the optimum criterion can be calculated, but this is rarely the exact criterion used by the inspector. In laboratory tasks, and in non-aviation inspection tasks, inspectors choose a criterion in a conservative manner. Thus, if the criterion should be low (i.e., they should be very willing to report indications as defects), inspectors choose a criterion which is not low enough. Similarly, they choose a criterion which is not high enough when the criterion should be high. Because of this conservatism inspectors may not react quickly enough in changing their criterion as costs and probabilities change. Thus, it is important to provide accurate and up-to-date feedforward information on the probabilities of defects in different areas to allow the inspector to make rapid criterion changes.

There are also known criterion shifts with both changing defect rate and time on task. There is little to be done about increasing the defect rate: it is fixed by the state of the aircraft. The reduction in hit rate at very low defect rates may well set a limit to the use of humans as detectors of rare events. Paradoxically, as maintenance improves to give fewer defects, the capability of the inspector to detect the few remaining defects worsens. There is clearly a need for more research into human/machine function allocation to alleviate this low defect rate problem. Time on task, the vigilance phenomenon, only causes a reduced detection rate due to criterion shift under special circumstances, i.e. uninterrupted performance. This may not be a problem in aircraft inspection, although the heavy use of night shift inspection where interruptions are less frequent and the human less vigilant, requires further study.

[5.3.4.3.3 Factors Affecting SATO](#)

The influence of decision time on sensitivity (d') was seen earlier, where it was suggested that it may not be of great importance. The ability of the inspector to integrate signal information over time may only extend for very short periods, at least compared to the time spent on search. However, this signal integration is not the only temporal aspect of decision making. When an indication is found from search, time is taken not so much in obtaining signal input as in locating and using standards, and performing the response. Thus, an inspector may have to locate the relevant standard on the workcard (which is a relatively rapid task) or in a manual (a longer task), or even through interpretation by others in management, quality control, or engineering (a much longer task). The response requires time to write, and a memory load. This response will also produce more work for the maintenance team, and hence potentially delay return to service. All of these represent indirect time pressures on the inspector.

In practice, inspectors do not appear to respond to such time pressures as much as may be expected. Their training and management reinforcement is biased towards accuracy in any [SATO](#). However, the managers of inspectors do feel these pressures, and also feel the need to insulate "their" inspectors from the pressures.

[5.3.4.4 A General Framework for Improving Speed/Accuracy Tradeoff](#)

In this section, a wide variety of temporal effects on inspection have been noted. In addition to the direct effect of time pressures (SATO), effects of time-on-task and time-of-day can be expected where vigilance or fatigue are relevant issues.

The main focus, however, has been on the joint performance measures of time-per-item and inspection errors, or their complements--speed and accuracy. Models have been presented which show how speed and accuracy are jointly determined. Access, search, and decision making all show a predictable speed/accuracy tradeoff.

If the objective is ultimately to bring speed and accuracy jointly under control, then the same concepts apply to all three key tasks. The equations defining the speed/accuracy operating characteristics have been given in detail, but the essence of all is the same: the [SAOC](#) defines the envelope of possible performance, determined by the physical functioning of the human operator within a physically-defined system. The choice of operating point on the [SAOC](#) is determined by the perceived costs of time and errors, and by the perceived probability of a defect being present. Thus, there are two control modes, hopefully applied in sequence:

1. Obtain the best [SAOC](#) envelope.
2. Obtain the best operating point on the envelope.

Clearly, the first control mode gives the prospect of simultaneous improvement in speed and accuracy, whereas the second control mode only substitutes one undesirable consequence (time) for another (errors). The analogy with inspection instrumentation (a close analogy for decision making) is that the first control mode represents increasing the signal-to-noise ratio of the instrument, while the second control mode is equivalent to choosing an optimum threshold setting.

The first control mode can be represented for the three tasks considered as:

Access: Changing the controllability of the vehicle or the unaided human movement.

Search: Changing the visual lobe size, area to be searched, and fixation time.

Decision Making: Changing the sensitivity/discriminability of the defect.

All of these three parameters (k, t, d') will take effort to improve, as they imply a change in either the physical system or the human training to deal with that system. The benefit from these changes, however, is seen in both speed and accuracy, and will be obtained. However, the speed/accuracy tradeoff is set (within broad limits).

In contrast, the second control mode implies altering the human's perception of costs/payoffs and probabilities to ensure that the balance the inspector chooses between speed and accuracy is the one which is optimal. For all of the models, this comes down to the costs and probabilities of errors and the costs of time. Error costs come from peers and other co-workers, from the management, and ultimately from society and its institutions (e.g., FAA) Costs of time come from perceived urgency of job completion. Examples are gate pressures, and the requirement for inspection to be completed early so that repairs can be scheduled. If there are conflicts and inconsistencies between these costs from different sources, or even their perceived costs, then confusion and inconsistency will result. For all convex [SAOC](#) curves, averaging of two different operating points will produce an apparent operating point (*C*) on a lower [SAOC](#). Inconsistency in the second control model can thus appear as a worsening in the first control mode.

Control of perceived costs is largely a function of the organization: its structure and its information flows. With a complex system such as aircraft maintenance and inspection (e.g., Taylor, 1990), intervention must follow careful technical analysis of the organization. For example, the more separated the inspection subsystem is from the maintenance subsystem, the fewer the direct pressures on the inspector. However, the price of this independence may well be lack of coordination and technical understanding between two of the major groups involved in maintaining airworthiness. Observations made during this project have pointed towards a lack of perceived time pressure on inspectors, largely due to their managers' function as insulators. No quantitative data (e.g., from surveys, questionnaires, or ratings) are available to substantiate this observation, but an obvious next step is to collect such data in a formal manner. The outcome of such a data collection effort would be a baseline of how (and where) inspectors choose their operating point on the [SAOC](#). The options available for changing the [SAOC](#) and the operating point are still those given in this section.

5.3.5 A FRAMEWORK FOR TRAINING FOR VISUAL INSPECTION

In parallel with development of training systems for diagnostic tasks (e.g., Johnson, 1990) the predominance of visual inspection requires studies of visual inspection training. Earlier reviews of training in aircraft inspection ([Drury and Gramopadhye, 1990](#); [Shepherd, et al., 1991](#)) have shown how the component tasks of inspection are amenable to training interventions. Literature from industrial inspection training was reviewed and applied to aircraft inspection.

Training is aimed at reducing both search errors (all misses) and decision errors (misses and false alarms). From a review of the various training interventions available (Gramopadhye, 1992), it becomes apparent that some interventions are better suited to some component tasks. The following section presents part of this review as a research rationale which will lead to specific experimental tests of training interventions. The review in [Section 5.3.5.1](#) covers three areas which are critical to inspection performance: search, decision-making, and perception.

5.3.5.1 Results of Inspection Training Literature Review

5.3.5.1.1 Search

As noted in [Section 5.3.4.2](#), search task performance is a function of visual lobe size and search strategy. Visual lobe training has been studied by Leachtenaver (1978) for photo-interpreters, who found that practice on a search task increased visual lobe size. However, practice on a visual lobe measurement task may also increase lobe size and transfer this increase to search performance.

Search strategy training is an under-represented area in the literature. From the literature it is seen that systematic search is always more efficient than random search, so that a useful assumption is that the searcher is always trying to be systematic (Arani, Drury and Karwan, 1984). One training objective should be to ensure systematic search, i.e. search in which all areas are fixated, and none are refixated during a single scan. The major difference between systematic and random search is whether or not an area is refixated. The only logical reason for an inspector to refixate an area before

a total scan is completed is that the searcher does not remember whether or not that area has been fixated already. Hence, it is seen that it is necessary to provide a memory-aid to the inspector to indicate the points of previous fixations to avoid refixations. This could be done by training the inspectors to use feedback from eye movements, either continuously (on-line), or in a discrete manner at the end of a search task.

Feedback from eye movements can be provided regarding both the number of fixations and the interfixation distance. Literature suggests that these parameters are correlated with an inspector's efficiency in locating possible defects. Providing this sort of feedback would be expected to result in the inspector developing a more efficient search strategy.

5.3.5.1.2 Decision Making

Wickens (1984) states that training for decision making can be provided in the following ways:

- Make the decision maker aware of the nature of limitations and biases. Training operators to consider alternative hypotheses might reduce the likelihood of cognitive tunnel vision.
- Provide comprehensive and immediate feedback so that the operators are forced to attend to the degree of success or failure of their rules.
- Capitalize on the natural efforts of humans to seek causal relationships in integrating cues when correlations between variables are known beforehand. Hence, providing information to the operator so as to emphasize the co-relational structure would help in entertaining particular hypotheses.

5.3.5.1.3 Perception

When the separate features that define all objects within a category may be variable, objects are assigned to different perceptual categories. Thus, the operator needs to develop a perceptual schema, a form of knowledge or mental representation that people use to assign to ill-defined categories. The schema is a general body of knowledge about the characteristics of a perceptual category that does not contain a strict listing of its defining features (e.g., features which must all be present for a particular instance to be termed a category). Because of such fuzzy defining characteristics, the schema is normally acquired as a result of perceptual experience with examples rather than learning a simple defining set of rules.

According to Posner and Keele (1968, 1970) the development of a schema consists of two components:

- a general representation of the mean, i.e., the basic form from which all the forms are derived;
- an abstract representation of the variability.

Research in schema formation suggests that the nature of mental representation which people use to classify stimuli into categories is not a strict list of the characteristics of the prototype but that the mental representation also contains information concerning the variability around the template. This is suggested by Posner and Keele (1968) who found that exposure to a variety of instances of a schema induced better performance than repeated exposure to a single instance.

Theories proposed by Medin and Schaffer (1978) state that assignment is not made by relating each new instance to a central prototype but rather relating it to the exemplar to which it is most similar and then assigning each new instance to the residence category of that exemplar.

Thus, from the above discussion, it is seen that to help in the development of the schema the training provided should be of variable instances of the category rather than a single instance of a prototypical member or rules defining the features which would classify the members into categories. The amount of variability provided in the training should be similar to that existing in the real setting.

[5.3.5.2 Rationale for Research on Visual Inspection Training](#)

From the above discussion, training for visual search would be expected to result in reduced search errors (Type 2 errors) and reduced search time. Similarly, training for decision making and perception would be expected to result in reduced Type 1 and Type 2 errors. Although training can be used to improve visual inspection performance, specific training schemes are not associated with factors that determine improvement in visual inspection performance. Hence, ad hoc training schemes are developed that guarantee improvements for a particular task without consideration whether such a training scheme could be extended to a similar task or a different task, or whether the training is optimizing the use of instructor and trainee time. Hence, the first step in the development of a rational training scheme is to identify the factors that affect visual inspection performance. The next step is to determine which of the functions of the inspection task are trainable. This in turn will establish the sensitivity of the inspection parameters to training.

For any training scheme to be effective it should minimize both search errors and decision errors. Thus, referring to the earlier proposed model of visual inspection, it is observed that intervention strategies could be developed at various stages of the inspection process which could be hypothesized to change the inspection parameters, resulting in improved performance.

The following factors are critical to the search process:

- ability to identify salient features which can be associated with a particular defect (so that features can be searched in parallel instead of requiring foveal attention);
- visual lobe;
- eye movement scanning strategy.

In order to improve visual inspection performance, it is necessary to develop training schemes which predict improvements in the above factors. In the following section various training schemes are briefly described.

[5.3.5.2.1 Visual Lobe Training](#)

The visual lobe is a very important determinant of search performance. Johnston (1965) states that observers with a larger visual lobe require fewer fixations than observers with a smaller visual lobe. He concluded that a large visual lobe or peripheral acuity may account for superior search performance. We still need to know how a large visual lobe can affect search performance and how people can be trained so as to increase the size of the visual lobe. If the above questions are answered, this would then result in a strategy for improving the visual lobe. The more general question which arises is: how does lobe size training generalize across tasks (e.g., targets and backgrounds). We are interested in understanding whether the visual lobe training on a given target type would result in an improved search performance for a different target type and the sensitivity of the search parameter to this type of training. Thus, it is essential to identify whether such a cross-over effect exists. If it does, then it is sufficient to train the person on one target type. If not, then it is essential to identify various target subsets, say T_1 , T_2 , within which cross-over does occur. The people could be provided visual lobe training on a single target belonging to each target subset.

[5.3.5.2.2 Feedback Training](#)

A person needs rapid and accurate feedback in order to correctly classify a defect, or to know the effectiveness of a search strategy. Every training program should begin with frequent feedback and gradually delay this until a level of proficiency has been reached. Additional feedback beyond the end of the training program will help to keep the inspector calibrated (Drury and Kleiner, 1990). The following feedback could be provided:

- Feedback regarding the correctness of classifying defective items into categories.

- Feedback of search strategy from monitoring eye movements.
- Feedback of fixation times from the eye movement search.

The first is known to be essential to learning in perceptual tasks (Annett, 1966). It provides the novice information regarding the critical difference between a defective item and perfect item, thus helping to develop a mental template which has the internal characteristics of the defective item. We are, however, still unsure as to what has improved. For example, has learning resulted in a new internal conceptual model of the task (i.e., is the inspector using only certain dimensions of the fault to classify it)?

It has been shown that an important difference between the best and the poorest search performance is the length of the sweeps between eye fixations during a search task (Boynton, Elworth, and Palmer, 1958). Thus, there exists a difference between how a novice and an expert move their eyes across the visual field. Gould (1973), in a visual inspection study of circuit chips, found that most of the eye fixations occur within a definite boundary, which is the area most likely to contain the targets. It is demonstrated that eye movements in a visual search scenario occur based on knowledge of the location of faults and on the probability of them occurring. The question that needs answering is: does feedback information regarding the eye movements help improve the scanning strategy? Here we hypothesize that providing such feedback information would aid the inspectors by allowing them to identify areas not covered or areas where one spends excessive time, and helping them develop a strategy to cover the entire area more effectively.

5.3.5.2.3 Feedforward Training

When a novice inspector has no knowledge of the type of faults, probability of faults, and occurrence of faults, visual search would be expected to be inefficient. Providing feedforward information should result in an improved search strategy because the uncertainty is reduced by the inspector knowing both where to look and what to look for. Perhaps the inspector could use the information to achieve a more systematic search strategy, guided by the knowledge of the fault characteristics. The inspector could use feedforward information in the following ways: 1) to ignore the information completely, 2) to selectively incorporate some of the information, or 3) to incorporate this information only at later stages of inspection, that is, only after gaining some verification. Kleiner (1983) suggests that experienced inspectors make use of feedforward information that complements their sensitivity to the fault. If the fault is one that is not easily detected, then the inspector relies heavily on the information provided. According to McKernan (1989), inspection tasks that will most likely benefit from the addition to prior information include those in which the value of the fault is greater than the value of inspection time, those in which the fault is particularly difficult to detect, and those in which the product may contain rare, detrimental, and easily overlooked, faults.

5.3.5.2.4 Attribute Training

Consider an item A. Let the item be faulty on attributes A1, A2, A3 and A4. The inspector could be trained on each of the above attributes. Such training would allow the inspector to set a response criterion for each attribute. The training should be generalizable in the sense that the inspector should be able to classify the items as defective if the items are faulty on one or more of the attributes. The inspector could be trained on which attributes to match first based on the probability of the item being faulty on the attributes and the ease with which the matching occurs. Experience and training of the inspectors determine how defect attributes are arranged (Goldberg and Gibson, 1986).

A similar training scheme has been proposed by Salvendy and Seymour (1973) for developing industrial skills. Here, separate parts of the job are taught to criterion, and then successively larger sequences of the job are integrated. Czaja and Drury (1981) and Kleiner (1983) used such progressive part training very effectively in inspection.

5.3.5.2.5 Schema Training

It is essential that the subject develop a valid mental template (internal representation) schema of the fault. The key to the development of a schema is that it should provide for successful extrapolation to novel situations which are still recognizable instances of the schema.

We need to know how schemas are developed, whether inspectors can be trained to develop schemas, and what sort of training (rule based or knowledge based) should be provided to the inspectors for effective development of such schemas.

The effects of two methods of training need to be evaluated in schema development: "active training" and "passive training". In active training, the inspector is presented with various instances of the fault and no-fault, and has to classify them as defective/non-defective. Feedback is provided regarding the correctness of classification. In contrast, passive training is where the inspector is merely presented with various instances of the faults without requiring an active response.

5.3.5.3 Testing the Visual Inspection Training Framework

In order to test whether the above predictions of training intervention/task component match are correct, a sequence of five experiments are to be undertaken as follows. All use the visual inspection simulator described in [Section 5.3.1](#). Brief synopses of each experiment are presented, with more detail given for Experiment 5, which has been completed.

Experiment 1: Feedback Training. This compares a control group and three feedback groups, using on-line and off-line feedback of both cognitive factors and performance factors (c.f. [Section 5.3.3](#)).

Experiment 2: Feedforward Training. Again, a control condition is used as a baseline against which to compare rule-based feedforward, knowledge-based feedforward, and combined feedforward.

Experiment 3: Attribute Training. Training for decision making using attributes training, i.e., providing the trainee with several levels of severity and complexity, is compared to a control condition where narrative descriptions are provided for the fault attributes.

Experiment 4: Schema Training. Schema development will be encouraged by exposing trainees to a wide variety of schema instances (corrosion levels and patterns) in both active and passive schemes.

Experiment 5: Visual Lobe Training. This experiment tests for the possible cross-over effects on the size of visual lobe measured for different fault types.

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 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 fault (cracks and loose rivets) and two types of area fault (corrosion and dents) to determine whether visual lobe training on one fault would generalize to other faults of the same or different classes.

5.3.5.3.1 Method

Twenty-four subjects were used for this study and were randomly assigned to four different groups, G1, G2, G3 and G4. Subjects were tested for 20/20 vision and color blindness. All the subjects were administered the [EFT](#) (Embedded Figure Test) and [MFFT](#) (Matching Familiar Figure Test), which have been shown to correlate with different aspects of industrial inspection performance.

Group G1: Subjects assigned to this group initially performed the visual search task on the above four fault types (randomly ordered) followed by visual lobe training on rivet cracks. The visual lobe training consisted of performing the visual lobe task five times. The training session was followed by a search task on the four fault types.

Group G2: Subjects assigned to this group also initially performed visual search tasks on all four

targets (ordered randomly). They followed this by visual lobe training on one area fault and dent. The visual lobe training consisted of performing the lobe task five times. The training session was followed by a search task on all four fault types.

Group G3: Subjects assigned to this group performed the visual search task in a similar manner to subjects in Groups G1 and G2. However, this was followed by visual lobe training on a neutral target, a computer-generated character. This training session was followed by a similar visual search task.

Group G4: Subjects assigned to this group performed similar visual search tasks. However, they did not undergo any visual lobe training. Subjects in Group 4 performed a computer task for a duration equal to the time required for the completion of the visual lobe training session in Groups G1, G2, and G3. This was followed by a visual search task.

5.3.5.3.2 Tasks

Visual Search Task. The visual search task was the simulated airframe visual inspection task described in [Section 5.3.1](#). Subjects had to search for a single fault type in a given area. Visual search performance of the subjects was evaluated on four faults which were classified into two types:

1. Area Faults - 1) corrosion, and 2) dent
2. Rivet Faults - 1) rivet crack, and 2) loose rivets (indicated by streaks of dirt on the rivet edge).

The task was unpaced. During each of the four visual search tasks, the subjects had to search for one of the predefined faults. Subjects were instructed to work as rapidly as possible consistent with accuracy. Subjects verified their response by clicking on the fault with the mouse button. Once a fault was located in a given area subjects inspected the next area.

Visual Lobe Task. The purpose of the lobe task was to determine the size of the visual lobe; i.e., how far into the periphery a subject could see in a single fixation. The basic procedure consisted of determining at what distance from the central fixation point the target was completely seen by the subject in a single fixation of the fault screen. The exposure duration was kept sufficiently short (0.33 s) to allow the subject a single fixation only. Subjects had to identify a single fault (a rivet fault in group G1, an area fault in group G2 and a neutral fault in group G3). The fault would appear on the horizontal center line of the target screen, at six equally spaced predetermined locations on the horizontal center line, three positions on either side of the central fixation point. No prior information concerning the position of the target was provided to the subjects. The subjects identified the position of the target, either to the left or to the right of the origin and accordingly pressed the key "Q" and "P" to register their response. Subjects were requested to avoid guessing and register responses only if they were sure as to the position of the targets. The fault screen alternated with a fixation screen, consisting of crosswires at the central fixation point exposed for a period of 2 seconds. The purpose of the fixation screen was to help the subjects fixate in the center of the screen after each viewing of the target screen.

5.3.5.3.3 Hypotheses Tested

1. Visual lobe training on one rivet fault (rivet crack) will result in improved visual search performance in detecting rivet faults (rivet cracks and loose rivets).
2. Visual lobe training on one area fault (dent) will result in improved search performance in detecting area faults (dents and corrosion).
3. Search performance on a fault will be superior in the case of subjects who underwent visual lobe training on the particular fault than for subjects who under went training on a neutral target, or subjects who did not undergo any visual lobe training.

Hypotheses 1 and 2 tested for crossover effects of visual lobe training and hypothesis 3 tested for the effectiveness of visual lobe training in improving visual search performance.

5.3.5.3.4 Experimental Design

The design was a 4 groups x 2 trials factorial design with six subjects nested within each group. The following performance measures were collected:

1. Number of correct responses for each of the six fault positions in the visual lobe task.
2. Time to detect a fault in each screen for the visual search task.

5.3.5.3.5 Results

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 ($F(2,15) = 11.05, P < 0.0011$), training trial ($F(4,60) = 13.46, P < 0.0000$) and their interaction ($F(8,60) = 1.75, P < 0.1046$) 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. These analyses are summarized in [Table 5.22](#), showing no main effects of groups, but highly significant group X trial interaction. [Figure 5.11](#) shows these group X trial interactions, where 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 other computer tasks, had no beneficial effect.

Search Time Analyzed	Group (G)	Trial (T)	Group X Trial
Loose Rivet	$P > 0.25$	$P < 0.005$	$P < 0.05$
Rivet Crack	$P > 0.25$	$P < 0.005$	$P < 0.10$
Dent	$P > 0.25$	$P < 0.01$	$P < 0.05$
Corrosion	$P > 0.15$	$P < 0.05$	$P < 0.05$
Overall	$P > 0.25$	$P < 0.0001$	$P < 0.005$

Table 5.22 Summary of Analyses of Variance of Mean Search Times

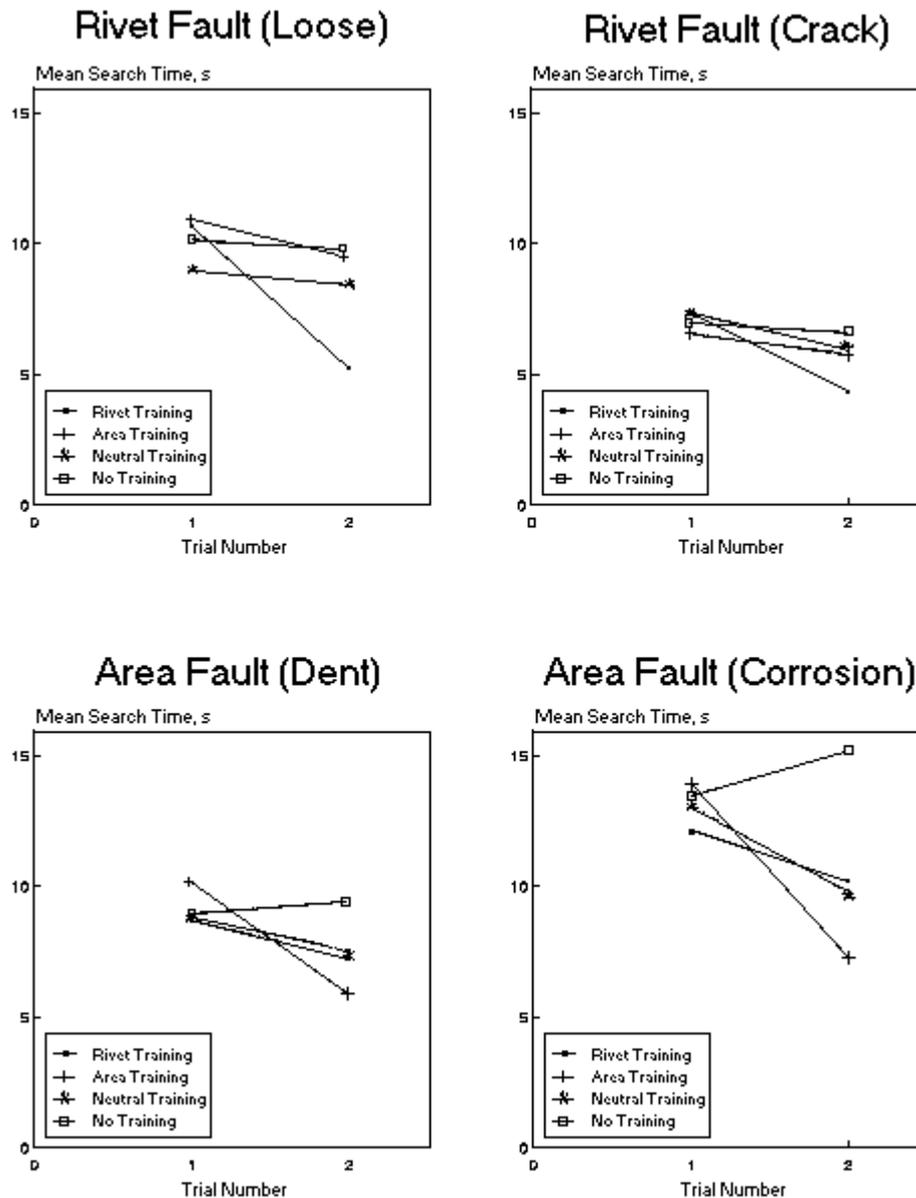


Figure 5.11 Search Performance Before and After Visual Lobe Training

Similar results can also be seen when the changes in visual lobe size during training are related to the changes in search time after training. [Table 5.23](#) relates the dependence of search time for each fault type to the increases in lobe size, using the coefficient of determination (r^2) as the measure of dependence.

Percent Increase in Visual Lobe Size For:	Percent Decrease in Search Time For:			
	Loose Rivet	Rivet Crack	Corrosion	Dent
Group 1 (Loose Rivet)	0.75	0.36	0.01	0.21
Group 2 (Dent)	0.09	0.00	0.68	0.85
Group 3 (Neutral)	0.16	0.05	0.74	0.00

Table 5.23 Dependence (r^2) of Percent Changes in Search Time on Percent Changes in Visual Lobe Size for Each Group

There was a direct transfer from the fault used in visual lobe training to that fault in visual search, with a smaller transfer to the other fault in the same group (rivet or area). The neutral fault visual lobe training transferred only to one area fault.

5.3.5.3.6 Discussion and Conclusions

Providing training, even just repeated practice, in rapidly detecting a fault in peripheral vision, does indeed increase the size of the area in which that fault can be detected in a single glimpse, i.e., the visual lobe. This increased visual lobe is not merely a result of increased familiarity with the experimental visual lobe task, as it transfers to a more realistic inspection task, visual search. Thus, even such a basic aspect of inspection performance as the visual lobe can be improved through training. For each fault type there was a 20-30% increase in lobe size over just five practice trials. This transferred to the search task with percentage changes in overall visual search time of:

Group 1 (Loose Rivet) 30%

Group 2 (Dent) 32%

Group 3 (Neutral) 18%

Group 4 (No Training) -4%

There is 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.

From [Section 5.3.4](#), it was seen that visual search follows a speed/accuracy tradeoff curve, so that what has been measured here as search times, can also be interpreted as search accuracy in a given, fixed time. Thus, this experiment has demonstrated the value of training in increasing the inspector's ability to receive and interpret peripheral visual information. The implication is that tasks similar to the visual lobe task given here need to be derived and used with inspectors. The benefits of a simple, simulator-based study in rapidly determining the feasibility of new training techniques has also been demonstrated. A study based on actual faults on a real aircraft structure would have been impossible as single glimpses cannot be repeated without the inspector learning the true identity of each fault. A study using hardware to simulate the faults would be extremely cumbersome, with hundreds of fuselage samples identical apart from fault location being required.

5.3.6 INTERNATIONAL COMPARISONS IN AIRCRAFT INSPECTION

As noted in [Section 5.1](#), a joint study of inspection practices in the U.K. and U.S.A. was undertaken as part of a Memorandum of Agreement between the [CAA](#) and [FAA](#). The Lock and Strutt report (1985) was in fact prepared earlier in that decade, so that the [CAA](#), for whom the report was produced, initiated an update by M. W. B. Lock in 1990. As the techniques of observation were similar to those used by the [FAA/AAM](#) team, a joint venture was created to allow direct comparison of U.S.A. and U.K. practices. Both C. G. Drury and M. W. B. Lock were participants, and have issued a joint report (Drury and Lock, 1992), so that only a briefer summary is presented here.

The aircraft to be maintained are designed and sold for world-wide markets, so that much of the inspection and maintenance is pre-determined by the manufacturers. However, the various regulatory authorities around the world (e.g., [FAA](#), [CAA](#), [JAA](#)) have different requirements. In addition, the way in which an airline chooses to meet these requirements leaves some latitude for local and cultural variations.

Although many points of difference were noted, perhaps the most obvious is in the way in which the inspection/maintenance job is scheduled and controlled. In the U.K., the management structures of maintenance and inspection are usually closely intermeshed. In the past it was frequently the case that the engineering manager and the quality control chief were the same person. Although this not the case in large transport aircraft, it can still be the case in smaller commuter airlines. Work arising from an inspection can be allocated by the inspector, who is often also a supervisor, or by a senior person who has responsibility for both inspection and maintenance. The inspector is frequently consulted during the defect rectification, in some cases is the actual supervisor of that work, and will usually be the person to buy back the repair.

In the U.S.A. the management structures of maintenance and inspection are separated up to a level well beyond the hangar floor. A wide variation of management authority was found whereby either maintenance, inspection, or even planning, could dominate (Taylor, 1990). In a few companies visited there was provision for coordination between maintenance and inspection by an engineer whose job was to ensure some cross talk. The engineer served as shift change coordinator. Typically though, work arising from an inspection is allocated by a maintenance supervisor so that the inspector who raised the defect has no responsibility for defect rectification and may not be the inspector who does the buy-back inspection.

The separation of the two management structures in the U.S.A. is dictated largely by the existing Federal Airworthiness Regulations, driven by a deeply-felt need for checks and balances as an error reduction mechanism. At the hangar floor level the general view is that repair and maintenance would suffer if the repairer knew that certain inspectors were 'buying back' the work, as some are known to be less stringent than others. The general view in the U.K. was that the system of having the same inspector responsible throughout for any particular defect and its rectification was preferable as the repair could be monitored at appropriate stages, ensuring that the job had been performed correctly.

Both systems lead to different requirements for training in managerial skills. Despite the greater direct management responsibilities of inspectors in the U.K., little formal training in managerial skills was evident.

A number of visits were undertaken by each participant in each country, either separately or together. There was no attempt at comprehensive sampling; rather the knowledge of each participant was used to select sites which would be illustrative of various features. For example, in the UK, visits were made to specialist third-party [NDT](#) companies which serviced civil aviation because they represent a major source of NDT expertise utilized by some airlines.

At each site, the visit was divided into two sections, although these often overlapped in coverage:

- **Systems Overview.** First the management of the maintenance of the site was probed in management interviews. The structure of the maintenance and inspection organization(s) was elicited during discussions with managers, shift supervisors, foremen, and often with staff who were outside the line management structure. These could include training personnel, archive keepers, work card preparers, planners, and so on depending upon the initial discussions with management. The aim was to be able to write a short description of how the system should operate, and the management philosophy behind this system structure and functioning.
- **Hangar-Floor Operations.** Detailed observations of the practice of inspection, and its organizational constraints, were made by following an inspector for all or part of a shift. As the inspector progressed through a job, questions were asked concerning the inspection itself

and ancillary operations, such as spares availability from stores, or time availability for training. Thus a reasonably complete task description and analysis could be written on the inspection task itself, while obtaining information on the wider context of the inspector's job. This technique also allowed the collection of anecdotal recollections of previous jobs, and other events from the past. While these had an obviously lower evidence value than direct observation of task performance, they did provide a valuable adjunct to the data collection process.

Sites visited included major air carriers, regional or second-level airlines, repair stations, and [NDT](#) companies. In addition visits were made to [FAA](#) and [CAA](#) personnel and to a Royal Air Force base where maintenance and inspection procedures are written.

5.4 CONCLUSIONS

As the [FAA/AAM](#) program on human factors moves from its second to third phases, work has progressed from observation to demonstrations of concepts for doing maintenance and inspection. The original approach, developed in Phase I and reported in [Shepherd, et al., \(1991\)](#) was to have human factors engineers study aircraft inspection and maintenance so as to determine a strategy. Enough depth and breadth of study was maintained to be able to find critical intersections between human factors knowledge and techniques on one hand, and field problems of inspection and maintenance on the other. This involved both top-down analysis, taking a systems view, and bottom-up analysis, performing detailed task analyses of inspector's jobs.

Phase II has rather closely followed the recommendations made in Phase I. Observation of field activities has been scaled down and re-focussed onto very specific areas. These have evolved into the on-going sequence of demonstration projects. While results from the first two such projects are not scheduled to be available until the summer of 1992, the concept appears to be working well. Airline personnel at all levels recognize that improvements are possible, and thus, are being most cooperative with the human factors team.

As Phase III approaches, more of the projects listed in the Phase I report will be performed, as well as new ones added. For example, the whole field of inspection and maintenance scheduling could benefit from human factors research into combined human/automated scheduling systems (e.g., Sanderson, 1989). When projects are completed, a dissemination of results and lessons learned will be needed, presumably by presentations and published papers. Both the [FAA](#) and the airline maintenance organizations need to consider the best ways for rapid dissemination and application of demonstration project results.

The detailed application of human factors knowledge (often models) to specific problems ([Sections 5.3.1, 5.3.2, 5.3.3, 5.3.4, and 5.3.5](#)) has yielded insights for the experimental program and the demonstration projects. Feedback is now required from the industry on whether it finds this work adds to its operational understanding. The experimental program is just starting, following hardware procurement and software development. As this progresses, the same simulations should be available for specific experiments supported by industry, as well as for the on-going programs presented here.

The long-term aim of the whole project is to provide phased solutions of practical use to industry to improve the already high performance of aircraft inspection and maintenance.

5.5 REFERENCES

- Annett, J. (1966). Training for perceptual skills, *Ergonomics*, 9, No. 6, 459-468.
- Arani, T., Drury, C.G. and Karwan, M.H. (1984). A variable-memory model of visual search. *Human Factors*, 26 (6), 631-639.

- Bhatnager, V.W. (1987). *The effects of experience on search strategy, performance and stress during microscope inspection of real IC chips*. Unpublished Ph.D. Dissertation, State University of New York at Buffalo.
- Bloomfield, J.R. (1975). Studies of visual search. In Drury, C. G. and Fox, J. G. (Eds.), *Human Reliability in Quality Control*. London: Taylor and Francis, 31-43.
- Boynton, R. M. (1960). *Concluding remarks in visual search techniques*. Washington, D.C.: National Academy of Science, National Research Council, Publication 712, 231-239.
- Boynton, R. M., Elworth, C. L. and Palmer, R. M. (1958). *Laboratory studies pertaining to visual air reconnaissance, part III*. Wright Paterson AFB, Ohio, Wright Air Development Center, WADC-TR- 304.
- Brown, I. D. and Groeger, J. A. (1990). Errors in the operation of transport systems. Special Issue of *Ergonomics*, 33.11, 1183-1430.
- Chi, C.F. (1990). *The human as optimizer in inspection*. Unpublished Ph.D. Dissertation, State University of New York at Buffalo.
- Czaja, S. J. and Drury, C. G. (1981). Training programs for inspection. *Human Factors*, 23(4), 473-484.
- Drury, C. G. (1992). Inspection performance. In Salvendy, G. (Ed.), *Handbook of Industrial Engineering*, Second Edition, Chapter 88, 2282-2314.
- Drury, C. G. (1991). Errors in aviation maintenance: Taxonomy and control. In *Proceedings of the Human Factors Society 35th Annual Meeting*, San Francisco, California, 42-46.
- Drury, C. G. (1990a). The ergonomics audit. In Lovesey, E. J. (Ed.), *Contemporary Ergonomics*, London: Taylor and Francis, 400-405.
- Drury, C. G. (1990b). The information environment in aircraft inspection. In *Final Report - Second Federal Aviation Administration Meeting on Human Factors Issues in Aircraft Maintenance and Inspection, Information Exchange and Communications*, 98-109.
- Drury, C. G. (1988). Inspection performance and quality assurance: Chapter 65. *Job Analysis Handbook*, New York: J. Wiley.
- Drury, C. G. (1985). Stress and quality control inspection. In Cooper, C. L. and Smith, M. J. (Eds.), *Job Stress and Blue Collar Work*, Chapter 7. Chichester, U.K.: John Wiley and Sons, Inc.
- Drury, C. G. (1984). Improving inspection performance. In Salvendy, G. (Ed.), *Handbook of Industrial Engineering, Chapter 8.4*, New York: John Wiley and Sons.
- Drury, C. G. (1978). Integration of human factors models into statistical quality control. *Human Factors*, 20.5, 561-572.
- Drury, C. G. (1971). Movements with lateral constraints. *Ergonomics*, 14.2, 293-305.
- Drury, C.G., and Addison, J.L. (1973). An industrial study of the effects of feedback and fault density on inspection performance. *Ergonomics*, 16, 159-169.
- Drury, C. G. and Gramopadhye, A. (1991) (draft). Speed and accuracy in aircraft inspection.
- Drury, C. G. and Gramopadhye, A. (1990). Training for visual inspection. In *Proceedings of the Third Federal Aviation Administration Meeting on Human Factors in Aircraft Maintenance and Inspection: Training Issues*. Atlantic City, New Jersey.
- Drury, C. G. and Kleiner, B. M. (1990). Training in industrial environments. In *Proceedings of Human Factors Association of Canada Meeting*. Ottawa, Canada. pp. 99-108.
- Drury, C. G. and Lock, M.W.B. (in press) (1992). *Reliability of aircraft inspection: U.K. and*

U.S.A. perspectives. Joint CAA/FAA Report.

Drury, C. G. and Prabhu, P. (1991) (working paper). Feedforward and feedback in aircraft inspection.

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

Drury, C.G. and Wang, M.J. (1986). Are research results in inspection tasks specific? In *Proceedings of the Human Factors Society 30th Annual Meeting*, 393-397.

Engel, F.L. (1971). Visual conspicuity, directed attention, and retinal focus. *Visual Research*, 11, 563-576.

Gallwey, T.J. (1982). Selection tests for visual inspection on a multiple fault-type task. *Ergonomics*, 25.11, 1077-1092.

Goldberg, J. H. and Gibson, D. C. (1986). The effects of checklists and error chunking upon the accuracy of visual inspection. *Trends in Ergonomics/Human Factors III*, Amsterdam: North-Holland, 359-368.

Goranson, U.G. and Miller, M. (1989). Aging jet transport structural evaluation programs. In *Proceedings of the 15th ICAF Symposium: "Aeronautical Fatigue in the Electronic Era"*. Jerusalem, Israel, 319-353.

Gould, J. D. and Schaffer, A. (1967). Eye movement parameters in pattern recognition. *Journal of Experimental Psychology*, 74, 225-229.

Gould, J. D. (1988). How to design usable systems. In Helander, M. (Ed.), *Handbook of Human-Computer Interaction*, New York: North-Holland, 757-789.

Gould, J. D. and Carn, R. (1973). Visual search, complex backgrounds, mental counters, and eye movements. *Perception and Psychophysics*, 14, 125-132.

Gould, J. D. (1973). Eye movements during visual search and memory search. *Journal of Experimental Psychology*, 98, 184-195.

Gramopadhye, A. (1992). *Training in visual inspection*. Ph.D. dissertation, State University of New York at Buffalo.

Gramopadhye, A., Palanivel, Knapp, and Drury, C. G. (1991). Training in visual search: Effect on visual lobe. (to be submitted to Human Factors)

Gramopadhye, A., Reynolds, J. and Drury, C. G. (1992) (draft). Design of the aircraft inspection/maintenance visual environment.

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

Hollnagel, E. (1989). The phenotype of erroneous actions: Implications for HCI design. In G.R.S. Weir and J.L. Alty (Eds.), *Human-Computer Interaction and Complex Systems*. London: Academic Press.

IES lighting handbook. (1984). Kaufman, J. E. (Ed), New York: Illuminating Engineering Society.

Johnston, D. M. (1965). Search performance as a function of peripheral acuity. *Human Factors*, 527-535.

Johnson, W. B. (1990). Application of new technology for aviation maintenance training: An industry status report. In *Proceedings of the Human Factors Society 34th Annual Meeting*, 1171-1175.

Johnson, W. B. and Rouse, W. B., 1982. Analysis and classification of human errors in

- troubleshooting live aircraft power plants. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-12, No. 3, 389-393.
- Kinney, G. C., Spahn, M. J. and Amato, R. A. (1977). The human element in air traffic control: Observations and analyses of the performance of controllers and supervisors in providing ATC separation services. McLean, VA: METREK Division of the MITRE Corporation, MTR-7655.
- Kleiner, B., 1983. *The development of an effective industrial training system*. Unpublished Master's Thesis, State University of New York at Buffalo.
- Krendel, E. S. and Wodinski, J., 1960. Visual search in unstructured fields. In A. Morris and E. P. Horne (Eds.), *Visual Search Techniques*, National Academy of Sciences, Washington, D.C., 151-169.
- Kundel, H. S. (1974). Peripheral vision, structured noise and film reader error. *Radiology*, 1144, 269-273.
- Kundel, H. S., and LaFollette, P. S., Jr. (1972). Visual search patterns and experience with radiological images. *Radiology*, 103, 523-528.
- Langan-Fox, C. P. and Empson, J. A. C. (1985). "Actions not as planned" in military air-traffic control. *Ergonomics*, 28, 1509-1521.
- Leachtenaver, J. C. (1978). Peripheral acuity and photointerpretation performance. *Human Factors*, 20(5), 537-551.
- Lock, M. W. B and Strutt, J. E. (1985). *Reliability of in-service inspection of transport aircraft structures*. CAA Paper 85013, London.
- Latorella, K. and Drury, C. G. (1991) (working paper). Errors in aviation inspection and maintenance: a system modeling approach to error management.
- Mackworth, N. H. and Bruner, J. S. (1970). How adults and children search and recognize pictures. *Human Development*, 13, 149-177.
- McKernan, K. (1989). *The benefits of prior information on visual search for multiple faults*. Unpublished Master's Thesis. State University of New York at Buffalo.
- Medin, D. L. and Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85, 207-238.
- Megaw, E. D. and Richardson, J., 1979. Eye movements and industrial inspection. *Applied Ergonomics*, 10, 3, 145-154.
- Meister, D. (1971). *Human factors: Theory and practice*. New York: Wiley.
- Monk, T. M., (1977). Sequential expectancy in visual search. *Human Factors*, 18, 607-612.
- Monk, T. M. and Brown, B. (1975). The effect of target surround density on visual search performance. *Human Factors*, 17, 356-360.
- Montazer, M. A., Drury, C. G, and Karwan, M. (1987). Self-paced path control as an optimization task. *IEEE Transactions*, SMC-17.3, 455-464.
- Morawski, T. B., Drury, C. G. and Karwan, M., (1992) (in press). The optimum speed of visual inspection using a random search strategy. *IIE Transactions*.
- Morawski, T. B., Drury, C. G. and Karwan, J. (1980). Predicting search performance for multiple targets. *Human Factors*, Vol. 22, 707-718.
- Nagel, D. C. (1988). Human error in aviation operations. In Wiener, E. L. and Nagel, D. C. (Eds.) *Human Factors in Aviation*, New York: Academic Press, Inc.
- Norman, D. A. (1988). *The psychology of everyday things*. New York: Bawic Books, Inc.

- Norman, D.A. (1981). Categorization of action slips. *Psychological Review*, 88(1), 1-15.
- Norman, D. A. (1980). *Errors in human performance*. Report #8004, Center for Human Information Processing, San Diego: University of California.
- Overington, I., (1973). *Vision and acquisitions*. London: Pentech Press.
- Prabhu, P., Sharit, J. and Drury, C. G. (1992). Classification of temporal errors in CIM systems: Development of a framework for deriving human-centered information requirements. *International Journal of Computer Integrated Manufacturing*, (in press).
- Posner, M. I. and Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology*, 83, 304-308.
- Posner, M. I. and Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353-363.
- Rasmussen, J. (1987). Reasons, causes and human error. In J. Rasmussen, K. Duncan and J. Leplat (Eds.), *New Technology and Human Error*, New York: J. Wiley, 293-301.
- Rasmussen, J. (1986). *Information processing and human-machine interaction, An approach to cognitive engineering*. New York: North-Holland.
- Rasmussen, J. (1983). Skills, rules, knowledge: signals, signs and symbols and other distinctions in human performance models. *IEEE Transactions: Systems, Man and Cybernetics*, Vol. SMC-13, 257-267.
- Rasmussen, J. (1982). Human errors: A taxonomy for describing human malfunctions in industrial installations. *Journal of Occupational Accidents*, 4, 311-333.
- Rasmussen, J. (1976). Outline of a hybrid model of the process plant in operator. Sheridan, T. B. and Johansson, G. (Eds.), *Monitoring Behavior and Supervisory Control*, New York: Plenum Press, 371-383.
- Rasmussen, J., Duncan, K. and Leplat, J. (1987). *New technology and human errors*. New York: J. Wiley and Sons.
- Rasmussen, J. and Vicente, K. J. (1989). Coping with human errors through system design: Implications for ecological interface design. *International Journal of Man Machine Studies*, 31, 517-534.
- Reason, J. (1990). A framework for classifying errors. In J. Rasmussen, K. Duncan and J. Leplat (Eds.), *New Technology and Human Error*. London: John Wiley.
- Rodgers, S. (Ed.) (1983). *Ergonomics design for people at work: Vol. 1*. New York: Van Nostrand Reinhold, 180.
- Rouse, S. H., Rouse, W. B. and Hammer, J. M. (1982). Design and evaluation of an onboard computer-based information system for aircraft. *IEEE Transactions Systems, Man, Cybernetics*, Vol. SMC-12, 451-463.
- Rouse, W. B., 1985. Optimal allocation of system development resources and/or tolerate human error. *IEEE SMC-15(5)*, 620-630.
- Rouse, W.B. and Rouse, S.H. (1983). Analysis and classification of human error. *IEEE Transactions Systems, Man & Cybernetics*, Vol. SMC-13(4), 539-549.
- Salvendy, G. and Seymour, J.W.D. (1973). *Prediction and development of industrial work performance*. New York: John Wiley and Sons.
- Sanderson, P. M. (1989). The human planning and scheduling role in advanced manufacturing systems: An emerging human factors domain. *Human Factors*, 31, 635-666.

- Sanderson, P.M. and Harwood, K. (1988). The skills, rules and knowledge classification: A discussion of its emergence and nature. In Goodstein, L. P., Andersen, H. B. and Olsen, S. (Eds.), *Tasks, Errors and Mental Models*, New York: Taylor and Francis, 21-34.
- Schoonhard, J. W., Gould, J. D., and Miller, L. A (1973). Studies of visual inspection. *Ergonomics*, 16, 365-379.
- Sears, R. L. (1986). A new look at accident contributions and the implications of operational and training procedures. Unpublished report. Boeing Commercial Airplane Company.
- Sheehan, J.J. and Drury, C.G. (1971). The analysis of industrial inspection. *Applied Ergonomics*, 2.2, 74-78.
- Shepherd, W. (1990). A program to study human factors in aircraft maintenance and inspection. In *Proceedings of the Human Factors Society 34th Annual Meeting, Volume 2*, 1168-1170.
- Sheridan, T.B. and Ferrell, W.R. (1974). *Man-machine systems*. MIT Press, Cambridge, Massachusetts.
- Stager, P. and Hameluck, D. (1990). Ergonomics in air traffic control. *Ergonomics*, 33.4, 493- 499.
- Swain, A. D. and Guttman, H. E., 1983. *Handbook of human reliability analysis with emphasis on nuclear power plant applications*. NUREG/CR-1278, US-NRC.
- Swain, A. D. and Weston, L. M. (1988). An approach to the diagnosis and misdiagnosis of abnormal conditions in post-accident sequences in complex man-machine systems. In L. Goodstein, H. Andersen and S. Olsen (Eds.), *Tasks, Errors and Mental Models*, London: Taylor and Francis.
- Swets, J. A., 1967. Signal detection theory applied to vigilance. In Mackie, R. R. (Ed.), *Vigilance: Theory, Operational Performance and Psychological Correlates*, New York: Plenum Press.
- Taylor, J.C. (1990). Organizational context for aircraft maintenance and inspection. In *Proceedings of the Human Factors Society 34th Annual Meeting, Volume 2*, 1176-1180.
- van Eckhout, J. M. and Rouse, W. B., 1981. Human errors in detection, diagnosis, and compensation for failures in the engine control room of a supertanker. *IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-11, No. 12*, 813-816.
- Wallack, P. M. and Adams, S. K. (1969). The utility of signal detection theory in the analysis of industrial inspector accuracy. *AIIE Transactions*, 1, 33-44.
- Wallack, P. M. and Adams, S. K. (1970). A comparison of inspector performance measures. *AIIE Transactions*, 2, 97-105.
- Wickens, C.D. (1984). *Engineering psychology and human performance*. Scott, Foresman & Co., Glenview, IL.
- Wickens, C. D. (1991). *Engineering psychology and human performance*, Second Edition. New York: Harper Collins.
- Williams, L. G. (1966). Target conspicuity and visual search. *Human Factors*, 8, pp. 80-92.
- Wright, P. (1991). Cognitive overheads and prostheses of some issues in evaluating hypertexts. In *Hypertext '91 - Proceedings*, San Antonio, Texas.
- Woods, D. D. (1989). Coping with complexity: the psychology of human behavior in complex systems. In Goodstein, L. P., Andersen, H. B. and Olsen, S. E. (Eds.), *Tasks, Errors and Mental Models*, London: Taylor and Francis, 128-148.

Chapter Six

A Human Factors Guide For Aviation Maintenance

6.0 INTRODUCTION

The U.S. air carrier industry and the Federal Aviation Administration are dedicated to the highest level of safety in commercial aviation. To achieve this goal, they must rely on effective and efficient maintenance operations. Proper maintenance support is indispensable to safety, to aircraft availability, and to airline profitability. The safety requirements dictate that maintenance be effectively error-free. Aircraft inspectors and aircraft mechanics must work in an environment and use procedures and equipment all carefully structured to work well and to minimize any potential for error. The design of procedures and equipment must ensure that errors are not built into the system.

The maintenance effort to ensure continuing airworthiness of the air carrier fleet is demanding and costly. The maintenance industry continues to grow in parallel with that of airline operations. [Table 6.1](#) shows that, in 1991, about 59,000 mechanics were employed in this industry, with maintenance expenses of approximately \$9 billion. These numbers reflect significant growth over the last decade but do not indicate the changing character of the industry. Maintenance operations are being recast to account for the introduction of new and more complex aircraft and the use of more sophisticated maintenance and inspection procedures.

Mechanics employed	=	58,819
Maintenance expenses	=	\$8.8 billion (11.5% of operating expenses)
Major carriers contract approximately 11% of maintenance work		

Table 6.1 Maintenance Parameters for U.S. Scheduled Airlines (1991 data) *ATA (1992); Office of Technology Assessment (OTA) (1988).*

Aviation maintenance is in fact a large industrial system which includes many elements such as the aircraft, the maintenance facility, supervisory forces, inspection equipment, repair equipment, and the maintenance technician. All of these elements together comprise the "maintenance system" ([Figure 6.1](#)). Within this system, the technician functions and should be viewed as one would view any other element. A maintenance technician has a set of operating characteristics. Conceptually this human can be considered in essentially the same manner as other system elements such as, for example, items of electronic equipment. The major difference is that the human is significantly more complex and not nearly as predictable. However, anyone responsible for designing or operating a system, such as a maintenance system, must understand the operating characteristics of each element within the system, and this includes the human.

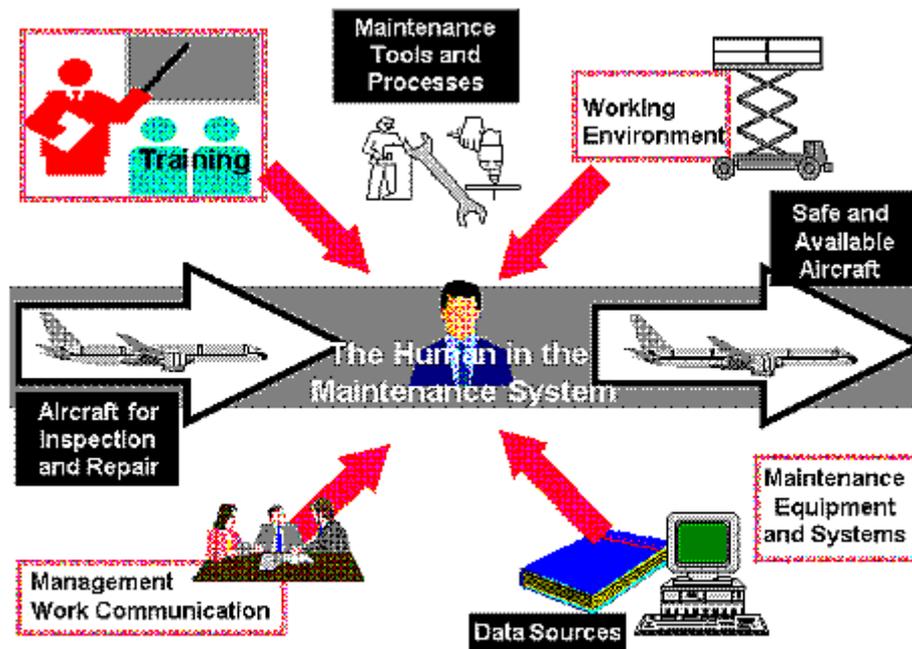


Figure 6.1 The Maintenance System

Human factors is a discipline which seeks to understand the laws of human behavior, the capabilities and limits of humans, and the effects of environmental and other factors on human behavior. A key goal of human factors is to provide guidelines for the optimum use of humans in operating systems. An allied goal is to define the manner in which variables internal and external to a system affect human performance within the system.

The operation of any system can only be optimized if every system element is working properly and if each element is carefully coordinated with every other element. The manager of a system such as a maintenance operation should have all necessary information concerning maintenance technicians and, in particular, those features of the maintenance environment which serve either to enhance or to degrade technician performance. The manager or supervisor of a maintenance activity can be aided through use of a Human Factors Guide that will provide this information in a form suitable for day-to-day reference use.

A Human Factors Guide will present established principles of job design and work. These principles, if well applied, can make a major contribution toward the control of human error in aircraft maintenance and inspection. Issues of communications, equipment utilization, work scheduling and load, work environment, and management relations all are of importance in determining worker effectiveness. A Human Factors Guide should cover these and other issues of human performance that can be applied in aviation maintenance. In presenting these principles, the Guide should see that the information is especially addressed to aviation maintenance and inspection needs.

The preparation of a Human Factors Guide is timely for a number of reasons. The most important of these reasons include:

1. There is a need. Safety is always a matter of concern. The Guide can contribute to maintenance efficiency and to the control of human error in maintenance. This in turn will support continuing safety. There also is the matter of cost control. Maintenance effectiveness contributes to cost reduction.
2. Human factors is a mature and growing discipline. The knowledge within this discipline should be used to support maintenance operations in the same manner as information from

the engineering sciences support specific maintenance procedures.

3. Considerable information concerning human factors in aviation maintenance has been developed both through the research conducted by the OTA Human Factors Team and through the conduct of human factors meetings during which beneficial recommendations have been made by representatives of air carrier maintenance. All of this information should be incorporated into a Human Factors Guide.

The Human Factors Guide, as envisioned now, will be directed primarily toward those concerned with the development and operation of air carrier maintenance organizations. However, the Guide will be structured to meet the needs also of a larger audience interested in and responsible for aviation maintenance. In general, the Guide should provide human factors principles and data of use to:

- Maintenance planners and supervisors
- Maintenance inspectors and technicians
- OTA management and OTA inspectors
- Air carrier operators
- Designers of maintenance equipment
- Aircraft design teams
- Aircraft manufacturers

6.1 DEVELOPMENT OF A HUMAN FACTORS GUIDE

The development of a Human Factors Guide for Aviation Maintenance is underway. The first step in this development was to consider the premise on which the Guide should be constructed. This premise is that the Guide, or any such document, is of little if any value if it is not used. The aviation maintenance community must use the Human Factors Guide if the Guide is to serve any real purpose. For the Guide to be used, it must meet ongoing needs of maintenance personnel and must be prepared in such a manner as to foster use by this group.

In order to collect information to satisfy the above requirements, a sampling of aviation maintenance personnel was conducted. The information solicited was designed to ensure that the real needs of maintenance personnel would be met and that the Guide could be consistent with the ways in which this sample stated they were likely to use such a Guide.

Approximately 60 individuals affiliated in some manner with the air carrier maintenance industry were contacted to provide guidance on significant maintenance topics. Names were selected from the list of attendees at earlier OTA Human Factors Meetings. The list included persons both from the United States and from foreign countries. The role of these persons in aviation maintenance, based on their replies, is shown in [Table 6.2](#). The fact that most replies were received from "Inspection/ Maintenance Managers" is to be expected since this job category constituted the bulk of the initial mailing.

Work Classification	Number
Inspection/Maintenance Manager	22
Educator/Trainer	7
Aircraft Designer	2
Other (Senior Management, Quality Assurance, Consultant, Research, AMT Associate, Crew Systems Analyst)	7

Table 6.2 Occupation of Respondents

The next question concerned the value users would place on a Human Factors Guide for their work. Not surprisingly, almost all of those who replied indicated a Human Factors Guide would be "very valuable" or "valuable." Since these replies were given by persons who had evidenced interest in this topic by attending human factors meetings, these replies were anticipated.

The individuals were questioned on the anticipated frequency of use for a Human Factors Guide, if the Guide contained appropriate information. This question was asked in order to determine whether the Guide should be prepared as a working document (as a job aid) or as a reference manual. [Table 6.3](#) presents the replies to this question.

<u>Frequency of Use</u>	<u>Number of Replies</u>
Review initially	1
Daily	3
Weekly	20
Monthly	13
Rarely	1

Table 6.3 Frequency of Anticipated Use for a Human Factors Guide

The replies indicate the preferred use for a Human Factors Guide would be as a working document consulted on a number of occasions during the year.

6.2 HUMAN FACTORS COVERAGE

The coverage provided in a Human Factors Guide is of great importance if the Guide is to be truly useful. Certainly, the topics included in this Guide should be those which members of the maintenance community consider important. In order to collect information concerning desired coverage, an outline of a prototype Guide was prepared. Each person was presented a list of chapter headings from the prototype outline and asked to judge the importance of the topic on a five-step scale ranging from "very important" (weighting of five) to "not important" (weighting of one). With this system, had each of the respondents judged a given topic to be "very important," that topic would have received a total score of 190. Results for this question are presented in [Table 6.4](#). While there is a dispersion of total scores, it is quite apparent that most topics were judged either as "very important" or "important." The topics in [Table 6.4](#) are listed in terms of decreasing order of judged importance.

<u>Topic</u>	<u>Weighted Score</u>
Human Error in Maintenance	178
Information Exchange and Communications	173
Maintenance Training and Practices	173
Human Capabilities and Limits	169
Human Performance	166
Work Requirements	163
The Maintenance Workplace	160

Table 6.4 Importance of Specific Topics for Inclusion in a Human Factors Guide

Prior to the contact with the 60 individuals, a question had arisen about the desirability of including a section within the Human Factors Guide concerning emotional factors. For this reason, a separate question was included that asked "Should the Human Factors Guide contain a section, not usually included in texts of this type, that addresses social and emotional factors that can affect the

performance of a worker?" The following responses were received:

Yes = 32 **No** = 6

Obviously, the majority of the respondents believe that a section containing the above information should be included in the Guide.

To ensure that no appropriate topics were missed, each respondent was asked to note any additional topics believed important for a Guide of this type. Quite a few replies were received; most appeared to be variants of the topics in the initial list presented in the mailing. However, a few were indeed new and are listed below:

1. Requalification, limitations, and competency verification for aviation maintenance technicians.
2. Minimum individual qualifications (eyesight, color blindness, and manual dexterity) for specific maintenance functions.
3. Sexual harassment. (This could become increasingly important as workforce demographics change.)
4. Working with the handicapped. (The recently passed Americans With Disabilities Act gives impetus to this topic.)

6.3 FORMAT

For a Guide to be useful, it not only must contain appropriate information but also must be presented in a manner designed to make it easy to use. Several questions addressed the general issue of format. The first question concerned optimum length. The contacted individuals were asked "To be most usable, what size should a Human Factors Guide be?" [Table 6.5](#) presents the responses.

Length	Responses
Less than 100 pages	18
100 - 300 pages	14
Over 300 pages	0
Size is of no concern	6

Table 6.5 Desired Length of a Human Factors Guide

The above replies clearly point to a shorter rather than a longer Guide. These data are supported by a comment submitted by one respondent:

A Human Factors Guide should be 50-75 pages for handout to line management personnel. It should be 100-300 pages for managers and supervisors with decision making capabilities for resources and monies.

The next item asked "What format would you find most useful?" This question is considered quite important since the manner in which information is presented can affect the extent to which individuals will seek and use information concerning the topic being presented. [Table 6.6](#) shows the results for this question.

Format	Number of Selections
Key information and recommendations in bullet form, with illustrations	20
Short statements, with illustrations	8
Running prose, with illustrations	8
Other (Please Specify) (Combine short statements, with illustrations, and key information and recommendations in bullet form, with illustrations; use running prose - segmented by topic statements; use electronic/digital format with key word search.)	2

Table 6.6 Preferred Format for a Human Factors Guide

Results indicate the desired format would be one in which information is presented tersely and concisely, either in bullet form or using short statements. Illustrations should support the materials as needed.

A question next was asked which relates both to the length of the Guide and the manner in which materials are presented. Two alternatives were given with a request for a preference between these two. [Table 6.7](#) lists the two alternatives and shows the replies.

Alternatives	Number
A shorter guide, presenting brief discussions and recommendations, with supporting data elsewhere (possibly in another book or in a computer data base)?	21
A longer guide, with supporting data included as appendices?	17

Table 6.7 Preference for Physical Structure of a Human Factors Guide

These results show a slight preference for a shorter Guide, even if one has to look elsewhere for data supporting and elaborating the concise information presented in the Guide. One supporting comment illustrates this: "A shorter Guide. Computerized supporting data would be very nice."

6.4 SUMMARY

The replies of the maintenance personnel, combined with other discussions with those likely to use a Human Factors Guide, lead to the following conclusions concerning the content and structure of the Guide.

Audience. The principal users of a Human Factors Guide will be air carrier maintenance planners and supervisors. However, care must be taken that the structure of the Guide not be oriented entirely toward this group. A Human Factors Guide also can be used to advantage by other groups, including OTA management and OTA inspectors as well as aircraft design teams and designers of maintenance

equipment. The document also could be used profitably in training operations.

Content. Three topics have been identified as most important for inclusion in a Human Factors Guide. These topics are:

- Human error in maintenance
- Information exchange and communications
- Maintenance training and practices

All topics must be given appropriate coverage in the Guide. Greatest attention, however, will be given to the three topics listed above.

Size. The Human Factors Guide should not be a large document and probably should not exceed 200 pages in length. A larger document might well impact use, particularly if the document is to be carried around within the maintenance facility. Current thinking is that supporting materials, which could be quite lengthy, would best be contained in a computerized data base in a [CD-ROM](#) system. With proper search strategies, data supporting the Guide could be obtained quite rapidly.

Style. Information within the Human Factors Guide, such as basic human factors principles applied to specific maintenance labor, should be presented concisely, possibly using a bullet format, with supporting illustrations. Introductory chapters and materials can be more in a running prose form. The language should be simple and straightforward English. This will make it more likely that the message is conveyed as intended. Use of simple English also will help should the Guide be translated into a foreign language for use in overseas maintenance activities.

6.5 REFERENCES

Air Transport Association. (1992). *Air Transport 1992. Annual report of the U.S. scheduled airline industry*. Washington, DC.

Office of Technology Assessment. (1988). *Safe skies for tomorrow*. (Report No. OTA-SET-381). Washington, DC: U.S. Government Printing Office.

Chapter Six Appendix

Sample Section Of A Human Factors Guide

The following section illustrates the manner in which the above concepts and rules would be applied in the preparation of a section for the Human Factors Guide. The section is presented in highly abbreviated form simply to show the appearance and general content of a part of the Guide. This section does not indicate the depth of coverage planned for individual topics.

1.0 SECTION I: AREA AND TASK LIGHTING

1.1 Importance of Lighting in Industrial Operations

Lighting conditions in an industrial workplace are important both for worker productivity and for worker comfort. Numerous studies have examined the effect on worker productivity of varying levels of task illumination (see Cushman, 1987). In general, these studies show that performance under low illumination improves to a point as the illumination level is increased. [Figure 1](#) shows the reduced time required to complete a typical industrial task (reading a micrometer) as the level of illumination on the task is increased. Note that when the illumination reaches about 100 footcandles,

no additional improvement is seen. In general, industrial tasks show smaller and smaller improvement in performance as illumination is increased. However, the point where performance finally levels off is task-dependent. Tasks that are visually difficult, as might be true for inspection activities, will require more light to achieve best performance than will easier tasks.

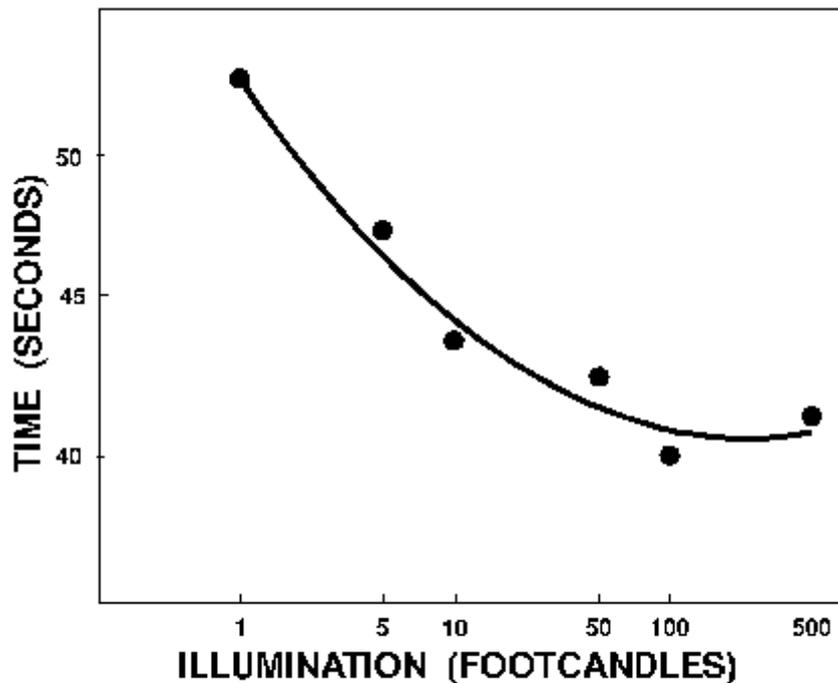


Figure 1 Effect of illumination level on time to complete a typical industrial task (micrometer readings). Adapted from *Sanders and McCormick, 1987*.

Research conducted to assess the effectiveness of illumination on performance must deal with two issues which can affect the results. First, motivational factors must be controlled. Subject, or workers, who know they are in a study will tend to perform better independent of the illumination level. Second, the age of subjects is important. Workers who are over 45 years of age will show more improvement with increasing illumination than will younger workers. If a work group contains older workers, illumination should be increased to account for this.

Recommendations for proper illumination levels for various activities have been prepared by the Illuminating Engineering Society and are presented in [Table 1](#).

Activity	Level (footcandles)
Working spaces with occasional visual tasks	10 - 20
Performance of visual tasks of high contrast or large size	20 - 50
Performance of visual tasks of medium contrast	50 - 100
Performance of visual tasks of low contrast or very small size	100 - 200
Performance of visual tasks of low contrast and very small size over a prolonged period	200 - 500

Table 1 Recommended Illuminance Values for Different Types of Activity. Adapted from *Kantowitz & Sorkin, 1983*.

Sanders and McCormick (1987) point out problems in arriving at recommendations for adequate illumination to ensure proper task performance. Interestingly, they note that recommended levels continually increase through the years. Current recommended levels are about five times greater than levels recommended 30 years ago for the same tasks.

Even though proper levels of illumination are provided, task performance can be degraded if glare sources are present. Glare is of two types. Direct glare is produced when a bright light source is in the visual field. Indirect glare, often called reflected glare, is reflected from the work surface and reduces the apparent contrast of task materials. Either direct or indirect glare can degrade task performance. [Table 2](#) offers suggestions concerning ways to control the effects of glare sources.

<u>To Control Direct Glare</u>	<u>To Control Indirect Glare</u>
Position lighting units as far from the operator's line of sight as practical	Avoid placing lights in the indirect-glare
Use several low-intensity lights instead of one bright one	Use lights with diffusing or polarizing lenses
Use lights that produce a batwing light distribution and position workers so that the highest light level comes from the sides, not front and back	Use surfaces that diffuse light, such as flat paint, non-gloss paper, and textured finishes
Use lights with louvers or prismatic lenses	Change the orientation of a workplace, task, viewing angle, or viewing direction until maximum visibility is achieved
Use indirect lighting	
Use light shields, hoods and visors at the workplace if other methods are impractical	

Table 2 Techniques for Controlling Glare, Adapted from *Rodgers, 1987*.

1.2 Lighting Conditions in Aviation Maintenance

A study of illumination conditions within major air carriers was accomplished as part of an OTA audit (Thackray, 1992). In these facilities, overhead lighting typically is supplied by mercury vapor, metal halide, or high-pressure sodium lights. The principal difference here is in terms of the color rendition of the lights. While color rendition is probably not too important for aircraft exterior maintenance tasks, the level of illumination could be. [Table 3](#) shows average illumination levels

measured at different maintenance work areas, both for day shifts and night shifts. [Table 3](#) also presents recommended illumination levels for aircraft repair and inspection tasks. Although slightly below recommended levels, the illumination for work on upper and lateral surfaces of an aircraft appear adequate. For repair and inspection conducted below wings, the fuselage, and within cargo and engine areas, measured illumination levels are not adequate and supplemental light sources are required. In general, supplemental lighting is provided through quartz halogen stand lights, dual 40-watt fluorescent stand fixtures, single hand-held fluorescent lamps, and flashlights.

<u>Measured (Footcandles)</u>		
	<u>Day</u>	<u>Night</u>
Hangar area	66	51
Below wings, fuselage and in cargo areas	26	15
Within fuselage	23	18
Visual inspection (2 D-cell flashlight)	100-500	
<u>Recommended (Footcandles)</u>		
	<u>Min. Level</u>	
Aircraft repair, general	75	
Aircraft visual inspection		
Ordinary area	50	
Difficult	100	
Highly difficult	200	

Table 3 Measured Illumination Levels at Major Air Carriers Compared with Recommended Levels. Adapted from *Thackray, 1990*.

Use of supplemental lighting does not necessarily solve existing lighting problems. The OTA audit of major carriers found that supplemental lighting systems frequently were placed too far from the work being performed and were too few in number. The result was that, even with supplemental lighting, the illumination directly at the work site was less than adequate.

Aircraft inspectors generally use small flashlights as supplementary sources. At times, small lights mounted on headbands may be used. The flashlights provide illumination ranging from 100 to 500 footcandles and are acceptable for visual inspection. However, use of the flashlight means that one hand cannot be used for manipulation of the systems being inspected.

In an attempt to produce more even lighting within maintenance bays, some carriers have painted the walls and even the floors with a bright white reflective paint. While this does tend to reduce shadow effects, other problems can arise. The principal one is glare. Reflected light from bright sources produces glare which can both cause discomfort and reduce visibility of key features of the maintenance task. The glare tends to obscure or veil part of the visual task.

1.3 Guidelines

The goal of controlling human error in aviation maintenance requires that maintenance be conducted under proper lighting conditions. This is true both for area lighting, that which illuminates the full working area, and task lighting, that directed toward specific work activities. Improper or insufficient lighting can lead to mistakes in work tasks or can simply increase the time required to do the work. In a program directed toward proper lighting conditions, the following guidelines should be observed:

- Area lighting within a maintenance facility should be a minimum of 75 footcandles. A level of 100-150 footcandles is preferred.
- Care must be exercised to see that the light level available for night maintenance activities in particular does not drop below recommended levels. Any lighting studies must be conducted both during the day and at night.
- Task lighting for aircraft inspection requires a minimum of 100 footcandles of illumination. For difficult inspections or fine machine work, 200-500 footcandles of illumination is necessary.
- Supplemental lighting must be adequate for the task at hand, best judged by the worker. Task lighting should be placed close to the work being done and, if feasible, should leave both of the worker's hands free for the work. If systems must be manipulated, lights mounted on headbands are preferred to flashlights.
- If the workforce contains a substantial percentage of older workers, i.e. those greater than 45 years of age, recommended lighting levels should be increased, probably on the order of 50 percent.
- Glare sources should be controlled. Supplemental lighting should be placed as far from a worker's line of sight as practical. Reflected glare can be changed by reorienting the work surface or changing the position of lights. Worker complaints are the best means for identifying offending glare sources.

1.4 Procedures for Evaluating Light Conditions

The best procedure for determining if lighting conditions are adequate is through the services of either the industrial hygiene department or the safety department of the air carrier. Individuals in these departments typically are trained in procedures for conducting an environmental audit, possess the necessary measurement equipment, and understand the problems involved in obtaining meaningful measurements. Specialists from these departments also will be able to provide a proper evaluation of the audit results.

If the services of specialists are not available, maintenance managers can assess lighting conditions themselves. Photometric equipment is available which will provide accurate (generally plus or minus five percent) measurement of facility lighting. Illuminometers/photometers are available commercially for a price in the order of \$1,000. Catalogs of scientific equipment describe these items.

1.5 REFERENCES

Cushman, W.H. & Crist, B. (1987). Illumination. In G. Salvendy (Ed.), *Handbook of human factors* (pp. 670-695). New York: John Wiley & Sons.

Kantowitz, B.H. & Sorokin, R.D. (1983). *Human factors: Understanding people-system relationships*. New York: Wiley & Sons.

Rodgers, S.H. & Eggleton, E.M. (Eds.). (1983). *Ergonomic design for people at work*. (Vol. 1). Rochester, NY: Eastman Kodak Co.

Sanders, M.S. & McCormick, E.J. (1987). *Human factors in engineering and design* (6th ed.). New York: McGraw-Hill Book Co.

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

Chapter Seven

The Effects of Crew Resource Management (CRM) Training in Maintenance: An Early Demonstration of Training Effects on Attitudes and Performance

7.0 INTRODUCTION

There are encouraging results from initial data on the effects of a [CRM](#) training program for technical operations managers on attitudes and maintenance performance.

The study is intended as an illustration of results beginning to emerge after only a few months experience in a long term program of effective communication to improve safe, dependable and efficient performance. Specifically, the analyses reported here compare managers' pre- and post-training attitudes about a variety of management and organizational factors as well as pre- and post-performance measures in several maintenance (and related) departments. Some highlights are as follows:

1. Participants' immediate response to the training was very positive. They were even more positive than those from other studies investigating [CRM](#) in maintenance or in flight operations.
2. Changes in relevant attitudes measured immediately before and after training reveal strong and positive changes following training, for three of the four indices measured. Follow-up results several months after training, reveal that these changes are stable; forthcoming data should continue to strengthen these conclusions.
3. Most maintenance performance measured before and after the [CRM](#) training sessions, examined in the present paper, show significant changes in the expected direction, indicating a positive effect of training.
4. A pattern of relationships in the expected direction were observed between the post-training attitudes and the post-training performance. In particular, those measures predicted (by the program's trainers and managers) to be particularly sensitive to the effects of the [CRM](#) training were affected.
5. Finally, in follow-up [surveys](#), managers responses indicate that, when transferring skills and knowledge, they tend to initiate interactive behaviors instead of passive ones. Anecdotal evidence is also beginning to confirm these changes and their positive impact on performance.

7.1 THE EFFECTIVE-COMMUNICATION PROGRAM: CREW RESOURCE MANAGEMENT TRAINING

Resource management training for airline flight crews was introduced in the late 1970's (Helmreich, 1979). It has spread to many air carriers in the U.S. commercial aviation industry, to several foreign carriers and to various sectors of U.S. and Canadian military aviation. That training has been extended from the cockpit to cabin crews, to maintenance teams and to air traffic centers and is now referred to as Crew Resource Management (CRM) training. Although specific programs differ from one organization to another, Crew Resource Management typically involves training in several team-related concepts: communication skills, self-knowledge, situational awareness, and assertiveness skills.

The effect of Crew Resource Management training in airline flight operations has been widely

studied during the 1980s. Numerous reports document CRM's positive impact on the attitudes and performance of flight crews (cf., Helmreich, Foushee, Benson, & Russini, 1986; Helmreich, Predmore, Irwin, Butler, Taggart, Willhelm, Clothier, 1991). Taken together, the evidence shows that team coordination among aviation "managers," and between them and subordinates, improves system effectiveness and safety.

As a result of recent work researching team concepts in aviation maintenance, further investigations have been recommended by both industry and government groups as a national priority (Federal Aviation Administration, "The National Plan for Aviation Human Factors," Washington, DC: 1991). A first instance of an airline applying [CRM](#) to maintenance operations was reported by Taggart (1990). Others are in beginning or planning stages. This report describes the case of a second company's program to apply [CRM](#)-type training in maintenance.

7.2 THE PRESENT STUDY

The analyses reported below will assess the relationships among managers' pre- and post-training attitudes about a variety of management and organizational factors and the levels of maintenance performance measures in a large U.S. airline (hereinafter called "the company").

7.2.1 The Purpose of the Program and the Course

The program's champion is the company's Senior Vice President for Technical Operations. He has stated that his aim for the training and evaluation program is to improve human resource (HR) management using science-based tools and techniques for diffusion and evaluation. Further, the effectiveness of this training, as measured by the ongoing evaluation of it, can help to direct the industry's HR practices in the future; and to guide the development of future ATA and [FAA](#) training policies and regulations. The training is a technical operations program entirely. It is managed and administered by technical operations people, and the trainers (assisted by professional communications training consultants) are technical operations people too. An Editor from *Aviation Week & Space Technology* participated in the training at the invitation of the company. The two articles he wrote (Fotos, 1991) further describe top management's reasons for undertaking the program, and provide impressions of the training itself.

Course objectives. The purpose of the training, as stated by trainers on the first day of each training session, is "To equip all Technical Operations personnel [management first] with the skill to use all resources to improve safety and efficiency." The objectives (the more specific goals of the training) are also clearly stated during the trainers' introductory remarks. They are as follows:

1. Diagnose organizational "norms" and their effect on safety.
2. Promote assertive behavior.
3. Understand individual leadership styles.
4. Understand and manage stress.
5. Enhance rational problem solving and decision making skills.
6. Enhance interpersonal skills

The course as designed for the objectives. The aims and objectives of the training are facilitated through a course syllabus containing 12 modules ([Appendix A1](#), [Appendix A2](#), [Appendix A3](#), and [Appendix A4](#) contain the current syllabus).

**Technical Operation Division
Crew Coordination Concepts Syllabus**

Module	Time	Facilitator	Method	Objective
Day 1 Introduction	8:15 (30 min)	TO	Introduce self. Cover facilities, restrooms, telephones, breaks, lunches, closing time. What have you heard about course? Review PURPOSE, OBJECTIVES. WORKING AGREEMENTS. Collect Pre-seminar questionnaire. Explain purpose & importance of ID#. Group introduces themselves.	Position program Get group talking Set the tone of the Clarify expectatio Remove teacher-s
Portland Video	8:45am (45 min)	HF	Show video; identify resource mgmt. problems; relate Portland problems to work place. Prepare flipchart.	Attention getter. problems become
Expectations	9:30am (15 min)	HF	Develop expectations of individuals in course. Write on chart, compare to course objectives.	Get expectations f illustrate differenc
Break	9:45am (15 min)			
Testing Assumptions/ DC-10 Video GUM	10:00am (120 min)	TO	Introduce concept of Perception vs Reality Show DC-10 Video GUM. Tablework: What were "Chain of Events" that led to the accident? Discuss "ASSUMPTIONS" that led or contributed to accident. How can we test our assumptions? What could each person have done differently?	Test Assumptions Advocacy - "Spea Inquiry - "Ask Qu Active Listening -
Lunch	12:00pm (60 min)			

Appendix A Part 1

Module	Time	Facilitator	Method
Behavior	1:00pm (75 min)	HF	Discuss use of Instruments; admin SDI; develop Behav Dim Model; discuss concept of "Assertive Behavior."
Break	2:15pm (15 min)		
Behaviors Cont'd	2:30pm (60 min)	HF	Scores SDI; Interpret SDI; draw arrows; discuss/apply "Assertiveness" Behavior modification approx. for effective supervision.
Break	3:30pm (15 min)		
Stress Management	3:45 pm (60 min)	TO/HF	TO-Introduce/understand stress What are sources of stress in our jobs? Develop list of Stressors. HF-Identify body's reaction to stress. Ways to deal with stress. TO-Work on how to deal with 2 examples. ETR'S, MANPOWER, PARTS AVAILABILITY

Appendix A Part 2

<u>Module</u>	<u>Time</u>	<u>Facilitator</u>	<u>Method</u>
Day 2 Opening DC-10 Video GUM 041	8:15am (30 min)	TO	Show GUM 041 video with breakfast. Address loose ends from Day 1; Outline schedule for Day 2. Review lessons learned in context of video (Gum 041)
Sub Artic Survival	8:45am (60 min)	TO	Purpose of simulation. Make individual decisions (step 2).
Break	9:45am (15 min)		
SAS II	10:00am (60 min)	HF	Define/develop rational decision process; apply process to simulation; complete SAS; critique team effect develop lessons learned/apply to job.
Break	11:00am (15 min)		
Norms/ EAL 855	11:15am (45 min)	TO	Identify role of norms in Tech Ops. Introduce concept, give examples. Allow time to review EAL 855. Assign each table a role. What were norms that led to accident? Groups develop lists of good/bad CO Norms. Discuss how to manage norms and prevent accidents.
Lunch	12:00pm (60 min)		

Appendix A Part 3

Module	Time	Facilitator	Method	Objective
Listening & Communicating	1:00pm (60 min)	HF	Tie into SAS; Sleep exercise Communication model; listening barriers; listening tips.	Listening is a learning listening limits quality of making.
Break	2:00pm (15 min)			
Supporting/ Confronting	2:15pm (60 min)	TO/HF	Use interactive dilemmas. Conduct 1st dilemma. Critique/lessons learned from 1st. Conduct second dilemma.	Application of Belbin Understand the value of is important. Use skills/lack of and
Break	3:15pm (15 min)			
Wrap-up Evaluation/ Questionnaire	3:30pm (45 min)	To/HF	Take Home concepts? USC/CAL Questionnaire explanation I.D #s on questionnaires.	Pledge to do some Feedback for program feel good about ex

Appendix A Part 4

State of training completed. To date, less than half of the technical maintenance operations stations, departments, or functional divisions in the company have been able to send all of their managers to the training -- this can only be done over time to avoid the absence of all managers and supervisors, at the same time, from their departments or stations while training is conducted. Thus, the present analysis uses data from roughly one-quarter of the whole technical operations management staff who will eventually complete the training. Overall, however, most cities among the company's 31 locations have had at least one-third of their maintenance managers attend a session. By the end of 1991, the majority of people remaining to complete the course are located in the company's three largest cities -- they comprise mainly maintenance supervisors (over 50% of whom still need to attend); and assistant supervisors (nearly 75% who have yet to attend).

Maintenance work units as the focus of the analyses. The analyses described in this report are intended to illustrate the effect of changes in work-unit attitudes (data collected from individual managers who underwent the training, combined into averages for the units to which they belong) and the performance data (classified into safety, dependability, and efficiency categories) constructed into measures for the same work units.

How is the course experienced? Summary results of answers to a post-training question concerning the perceived usefulness of the CRM training are shown in Figure 7.1. This figure compares the sample from the present study with two flight operations samples from other companies (Helmreich, 1989) and with one other technical operations sample (Taggart, 1990). Like the other maintenance department measured, the data from the Technical Operations division of the present company reveals high enthusiasm about the training. None of the respondents in either maintenance department sample said that the CRM training was either a "waste of time" or only "slightly useful." In addition, a very high percentage (68%) of the present company found the training to be "extremely useful." This training appears to be very well received by its "customers."

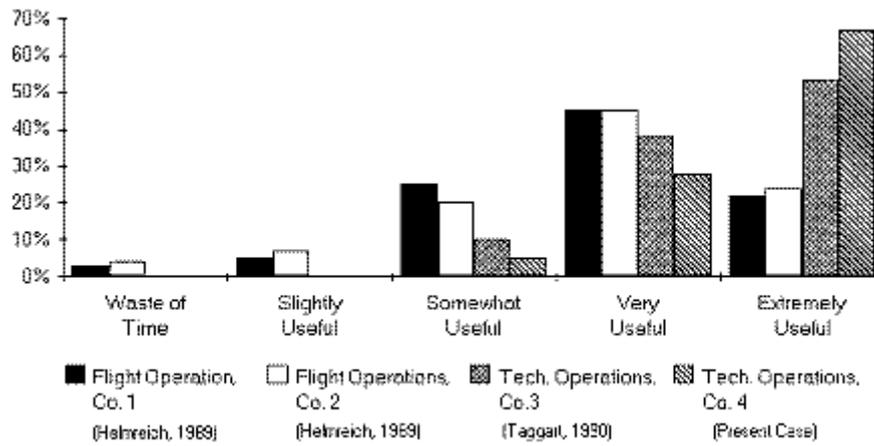


Figure 7.1 Ratings of Usefulness of CRM Training for Flight Operations and Technical Operations

7.3 CREW RESOURCE MANAGEMENT/TECHNICAL OPERATIONS QUESTIONNAIRE

7.3.1 Background Related to Measurement of CRM Training

A questionnaire called "CMAQ" (for Cockpit Management Attitudes Questionnaire) has long been a recognized measure for assessing flight crew attitudes (Helmreich et al., 1986). The [CMAQ](#) contains 25 items measuring attitudes that are either conceptually or empirically related to [CRM](#). Taggart (1990) revised the [CMAQ](#) for use in a technical operations department, and reported positive initial results following [CRM](#) training.

A recent study involved the analysis of the [CMAQ](#) instrument through the use of Factor Analysis, a technique to explore for a consistent internal structure (Gregorich et al., 1990). In their study these authors showed the relationships among the 25 [CMAQ](#) items clustered into the following four constellations of attitudes: Sharing Command Responsibility, Value of Communication and Coordination, Recognizing Stressor Effects, and Avoiding Conflict.

Because multiple items tap into specific aspects of [CRM](#) (e.g., "communication," and "interpersonal skills"), Gregorich et al., (1990) combined the items into composite indices. Such index scales permit more detailed assessment of separate but related attitudes than a single total score for the entire questionnaire, and they also provide more accurate and reliable results than are available from each of the individual questionnaire items alone.

7.3.2 Survey Used in the Present Study

Measurement of attitudes. The "Crew Resources Management/Technical Operations Questionnaire" (CRM/TOQ) as used in the present study is a modified version of Taggart's (1990) revised [CMAQ](#).

The [CRM/TOQ](#) contains 26 multiple response items. The company's modifications of the [CMAQ](#) involved removing five questions and adding six others. The five questions were removed because they either lacked predictive validity (as reported by earlier flight crew studies; Helmreich et al., 1986) or, in the company's opinion, lacked relevance to technical operations.

Six questions were also added to the [CRM/TOQ](#), based on items intended to measure respondents' perceptions of behaviors dealing with attainment of work goals (Geirland & Cotter, 1990).

Measurement over time. At the time of this writing (some six months into an 18 month training process) the data sets are necessarily incomplete. Over 500 of 1,800 total managers have already attended the two- day training program. Virtually all those attending the training have completed the [CRM/TOQ](#) both pre- and post-training. Some 400 of the 500 pre- and post-training [CRM/TOQ](#) completed have been entered into the data base and will be used in the present demonstration. It is still too early in the training and evaluation process to have abundant data for the [CRM/TOQ](#) follow-up (two and six month), or many months of performance data subsequent to the training; but the present report will describe the results of some 60 two-month and 50 six-month follow-up questionnaires returned so far. This follow-up version of the [CRM/TOQ](#) was mailed in early October, 1991, to the homes of those managers who had attended the [CRM](#) training in May or June for the "six-month" measure, while attendance during July and August qualified for the two-month follow-up.

Collecting baseline data for testing statistical goodness of data. During the last week in May 1991 and before the training program began, the [CRM/TOQ](#) was sent to the homes of all directors, managers, supervisors, and assistant supervisors (1,787 total). Within five weeks over 900 questionnaires had been returned for a return rate of over 50%. A return rate this high is considered quite acceptable with paper and pencil surveys of this type (Borg & Gall, 1986) -- especially since a reminder or prompt was not possible, given that the training program was due to be announced and start within two weeks of the mailing. This questionnaire was termed the "baseline" survey. A sample of the "baseline" questionnaire and one of the "follow-up" questionnaires are included here in [Appendix B](#) and [Appendix C](#), respectively.

The mean scores of the baseline results could reveal bias with a 50% return, but continuing comparisons between those baseline data and immediate pre-training results (where very close to 100% return rates are realized) similar relative levels for the six composite index scales are revealed. [Figure 7.2](#) presents the comparison profiles. These results show few differences in absolute scores and similar profiles.

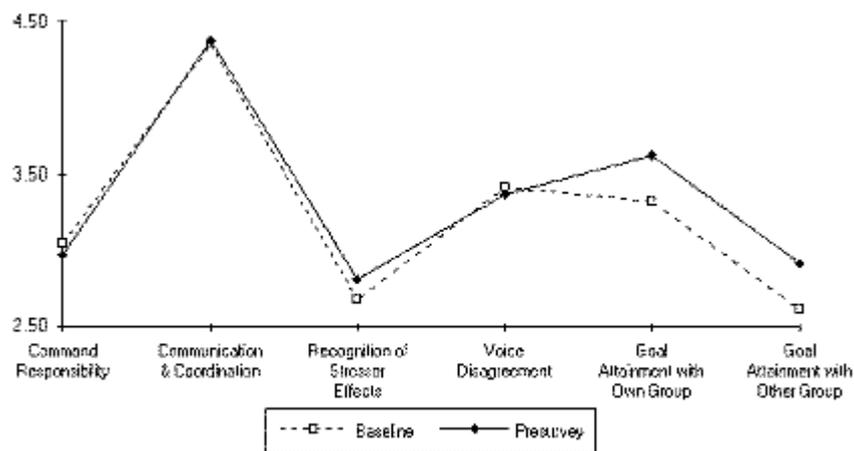


Figure 7.2 Mean Scores of Indexes for Baseline and Pre-Seminar for Technical Operations

7.4 PERFORMANCE DATA DESCRIBED

Technical Operations managers in the company already collect performance data in abundance. [Table 7.1](#) presents the 14 measures used as end-result criteria in the present study. Three conditions were met in order to include these measures in the work-unit analysis reported here. First (and obviously) the performance measures need to be available by work unit, and not just by department or function. Second the measures must be ones that people in the work unit can affect by their actions and not merely ones that are conveniently assigned to a unit -- but for which it can do little.

The third condition applied was that the measures not be directly related to or completely determined by, other measures in the set.

I. MEASURES OF SAFETY	
A. <u>Safety of Aircraft</u>	
1.	Number of Ground Damage Incidents ^a (44)
2.	Number of Air Turnbacks/Diversions caused by human error ^{a,d} (31)
B. <u>Personal Safety</u>	
3.	% Occupational Injury Days Lost ^a (60)
4.	% Sick Days Lost ^a (60)
II. MEASURES OF DEPENDABILITY	
A. <u>Departure Performance</u> (Line Station data only, n=31)	
5.	% Departure within 5 minutes ^a
6.	% Departure within 15 minutes ^a
7.	% Departure within 60 minutes ^b
8.	% Departure over 60 minutes, but not canceled ^a
9.	Number of Delays due to late from Maintenance ^{a,d} (n=35)
B. <u>Service Levels</u> (Materials Services only, n=9)	
10.	% Rotable parts available ^c
11.	% Expendable parts available ^c
C. <u>Base performance to plan</u> (n=4)	
12.	% Heavy Checks on time to initial plan ^b
III. MEASURES OF EFFICIENCY	
A. <u>Productivity</u>	
13.	Ratio of Hours paid to those applied to production ^b (7)
B. <u>Cost</u>	
14.	% Overtime paid to total wage bill ^b (60)
(maximum number of work units possible are in parentheses)	
a= Measures evaluated most sensitive to effects of CRM training	
b= Measures evaluated moderately sensitive to effects of CRM training	
c= Measures evaluated least sensitive to effects of CRM training	
d= These measures exhibit skewness and lack of variability	

Table 7.1 Technical Operations Performance Measures Available by Work Unit

The 14 measures are classified into three performance categories: Safety, Dependability and Efficiency; and this classification is shown in [Table 7.1](#). The trainers and administrators of the [CRM](#) course evaluated the 14 performance measures and predicted which of them would be more sensitive to effects of [CRM](#) training. Their conclusions were that six measures were the most readily improved by the training. These six performance measures included both aircraft safety items (ground damage and turnbacks), days lost to occupational injury, dependability based on departures within 5 and 15 minutes and delays due to maintenance error. Their second ranked performance measures likely to improve in response to training included sick days lost, departures within 60 minutes, canceled flights due to late from maintenance, base maintenance performance to plan, hours applied to production and overtime paid. The performance indicators rated least likely to improve as a function of the training were departures later than 60 minutes and warehouse parts service levels.

The present analysis employs four months of performance data. June and July 1991 are considered "pre-training," since they are coincident with, or precede, most of the training sessions measured here. August and September are used as the post-training measures. These performance data are available for 31 line maintenance stations, 4 base maintenance stations, 3 shops, 9 materials services warehouses, and 13 inspection/quality assurance units, for a maximum of 60 data points each month. The planning and engineering units in technical operations do not yet have unit performance data which they can be measured against, although the company is developing such measures for inclusion in later analyses. As noted above, the attitudinal data are not yet available for all 60 work units; and in some cases performance data is unavailable for certain work units for the four months analyzed for this report. Because of this, the full number of maintenance units available for the present analysis is somewhat less than the full set.

7.5 THE ANALYSIS PLAN DESCRIBED

7.5.1 A Model for Testing Relationships

The theory tested is that training in teamwork and communications influences maintenance personnel attitudes and perceptions, and these in turn produce positive changes in group behavior which impact performance. The analysis of the model, specified in [Figure 7.3](#), tests these influences by postulating systematic relations among a set of explanatory variables. [Figure 7.3](#) provides a diagram of the supposition or small "theory" examined in this model.

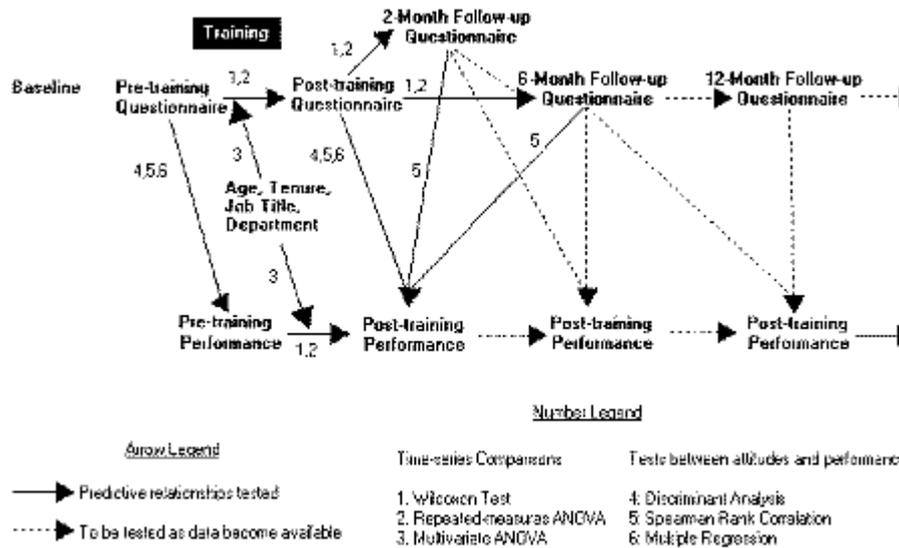


Figure 7.3 A Model of CRM Training, Attitude Change, and Technical Operations Performance

The conclusive assessment will be a causal analysis proving the strength of influence the training has on subsequent safety, dependability and efficiency. The present data do not yet establish true causal linkages. However, the present data do show associations between pre- and post- training attitudes and perceptions (measured by the [CRM/TOQ](#)), and between post- training [CRM/TOQ](#) results and performance.

7.6 RESULTS OF STATISTICAL ANALYSIS

7.6.1 Factor Analysis of Survey Items

Among other statistical procedures undertaken to assess the goodness of the [CRM/TOQ](#) measure, two Factor Analyses were performed, using data from the 900-plus completed questionnaires from the "baseline" survey. The first Factor Analysis was conducted to determine the underlying internal structure among the answers to the 20 attitude questions, and the second was run for the six questions that measured perceptions of behaviors dealing with attainment of work goals.

First, the structure of the attitude data was found to be very similar to the structure reported by Gregorich et al. (1990). On the strength of this similarity, it was decided to apply the same names to the first three of the four [CRM/TOQ](#) factors that Gregorich applied to the [CMAQ](#). Because all factors derived from the revised [CMAQ](#) were statistically strong, even the "avoiding conflict"

composite (not subsequently used by Gregorich, et. al., (1990)) was retained here as the reflected index "Willingness to Voice Disagreement." The item sums for each factor were averaged to form four index scores for each respondent. The four factors were titled (1) Command Responsibility, (2) Communication and Coordination, (3) Recognition of Stressor Effects, and (4) Willingness to Voice Disagreement. The reliabilities were good for scales of this length, ranging between .54 and .56 for Indices 1, 2, and 4. The reliability coefficient of .39 for Index 3 was lower than desired for confidence in a stable index, but the association among the contributing items "communalities" were reasonably high at .58 and .59.

Secondly, items measuring attainment of work goals loaded on two separate factors. One factor consisted of items assessing "Goal Attainment with My Group and the other "Goal Attainment with Other Groups;" their respective reliability coefficients were .77 and .74. Two additional indices were formed based upon these results.

7.6.2 Attitude Changes Over Time

Pre- and post-training scores. The "before" and "after" scores for the four composite attitude scales were compared and scores of the 385 individual managers (whose pre- and post-questionnaires could be matched) combined and averaged by the work units they belong to. Several analyses allow comparison of pre-training attitudes, to post-training attitudes, as measured by the [CRM/TOQ](#). Both the Wilcoxon Matched-Pairs Signed-Ranks Test and the Repeated-Measures Analysis of Variance tests were used to assess differences between pre- and post-measures and each test has its advantages, depending upon the nature of the data set (SPSS - User's Guide, 1990). In addition, Multivariate Analysis of Variance (MANOVA) was employed to test the effects, across time, of demographic variables, such as job title, job tenure or age, on attitudes and of different departments on performance measures. [Figure 7.4](#) shows the comparison of overall averages for the four composite attitude scales.

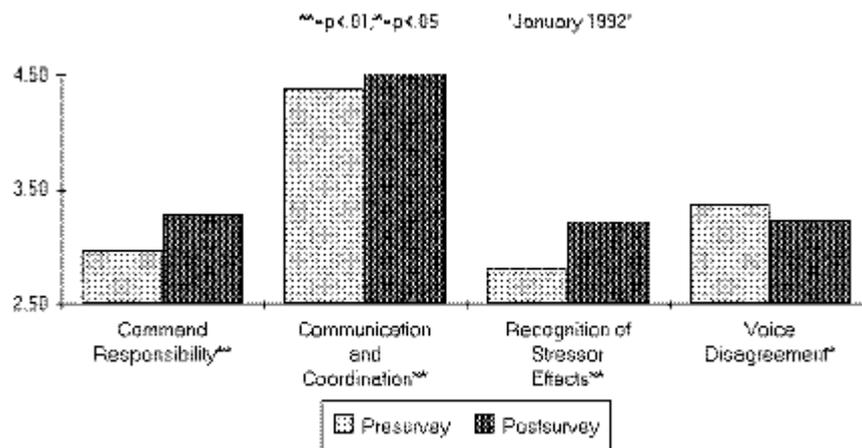


Figure 7.4 Means Scores of Indexes for Pre- and Post-Seminar Technical Operations

All of the pre-training and post-training changes are in the expected direction except for "Willingness to Voice Disagreement," which shows a shift away from voicing disagreement and toward avoiding conflict. All of the differences are significant at the .05 level or better.

Willingness to Voice Disagreement: Effect of Age and Job. [Figures 7.5](#) and [7.6](#) present the change in "Willingness to Voice Disagreement" for the age of the maintenance personnel and for the given categories of job classifications, respectively. Each of these categories interacted with the repeated variable (pre-post survey), affecting the magnitude of negative shift (Pre-Post x Age: $F = 1.96$, $p < .10$; Pre- Post x Job: $F = 5.20$, $p < .01$).

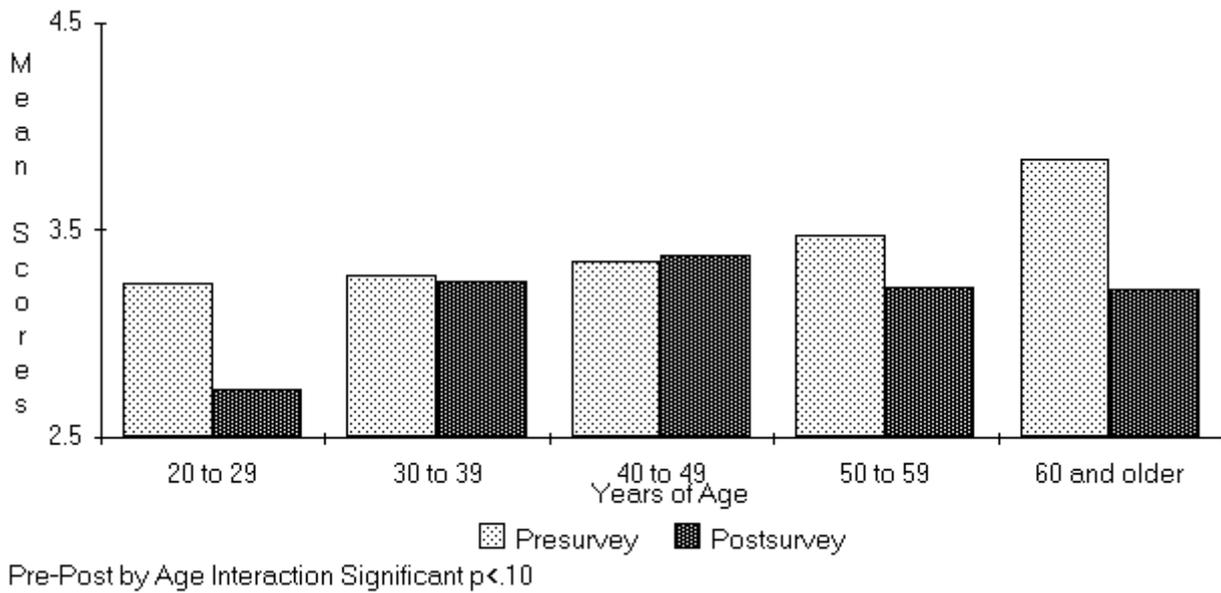


Figure 7.5 Means Scores for Pre- and Post-Seminar 'Willingness to Voice Disagreement' by Years of Age

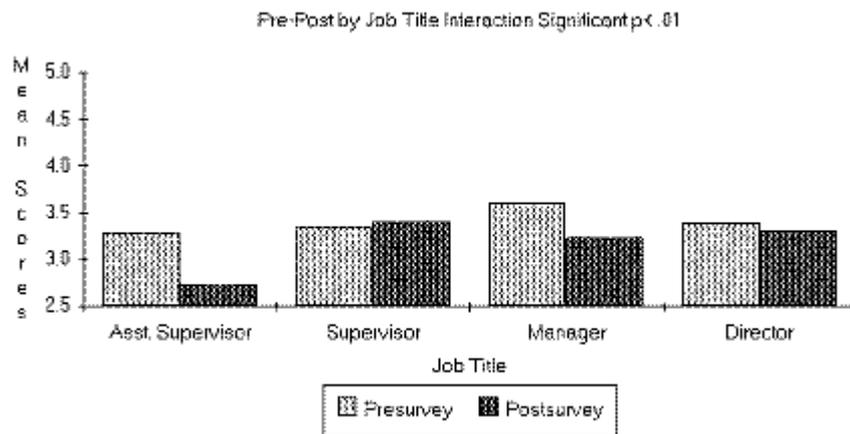


Figure 7.6 Mean Scores for Pre- and Post-Seminar 'Willingness to Voice Disagreement' by Job Title

These interactions are interesting because they further explain the negative shift in attitude on the "Willingness to Voice Disagreement" scale.

Figure 7.5 shows that the oldest and youngest members of the sample exhibit the greatest negative shift, whereas the remainder show very little change in "Willingness to Voice Disagreement" immediately following the training. In addition, Figure 7.6 shows that the assistant supervisors (who are over-represented by the youngest and oldest members of the sample) and managers have the greatest negative shift in attitude.

Recognition of Stressor Effects: Effects of Department. The magnitude of change in attitudes on the index "Recognition of Stressor Effects" were affected by the department affiliation of the respondent, revealing a significant interaction between attitude change and department (Pre-Post x Dept: $F = 2.25, p < .05$). Figure 7.7 shows that the attitudes of base maintenance and quality departments change little over time whereas those of line maintenance and planning departments improve considerably.

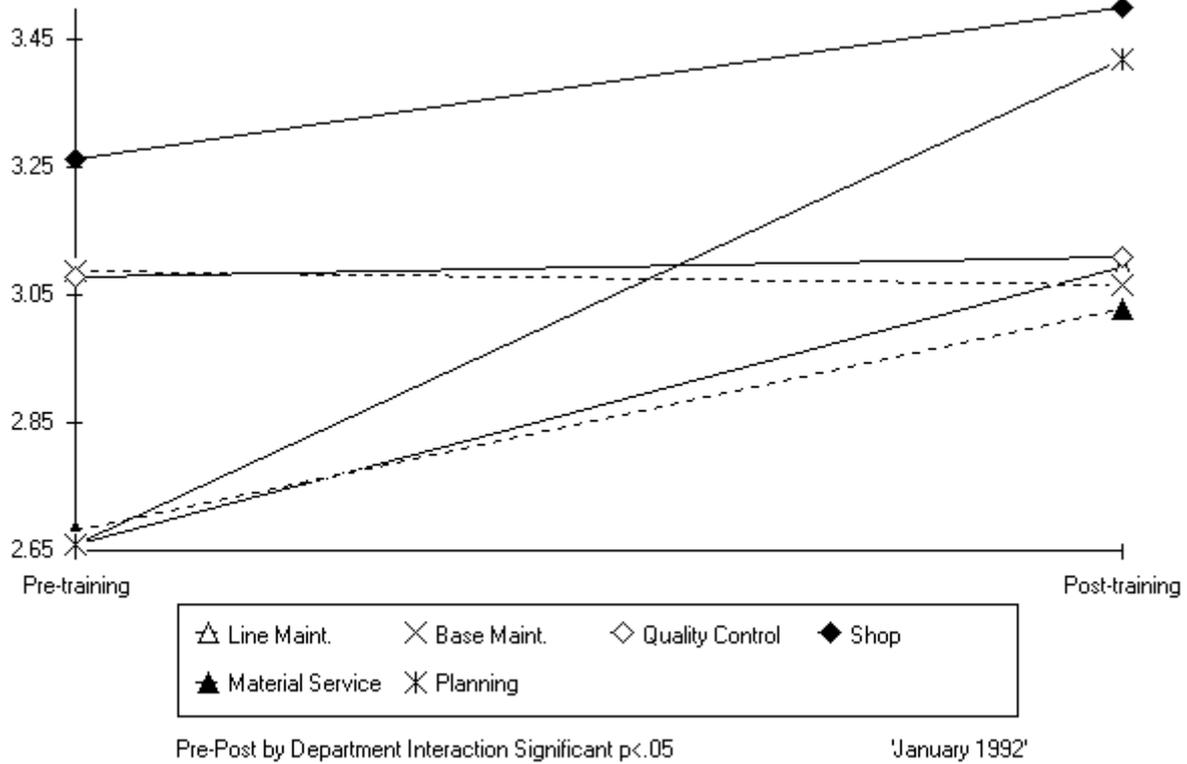


Figure 7.7 Means Scores, Pre- and Post-Training, for 'Recognition of Stressor Effects' by Department

Goal Attainment. The scores for two goal attainment scales can also be compared before and after training. Since these are perceptions of behavior, and not measures of attitudes, goal attainment scales were initially not expected to show changes because there would not be time to change the behaviors which were the focus of the measures. Because perceptions are not behavior, like attitudes, they can be more quickly influenced by exposure to new information, or by reconsidering initial assessments. One statistically significant change was found in the goal attainment with one's group immediately following training. [Figure 7.8](#) shows the pre-post averages for both of these measures.

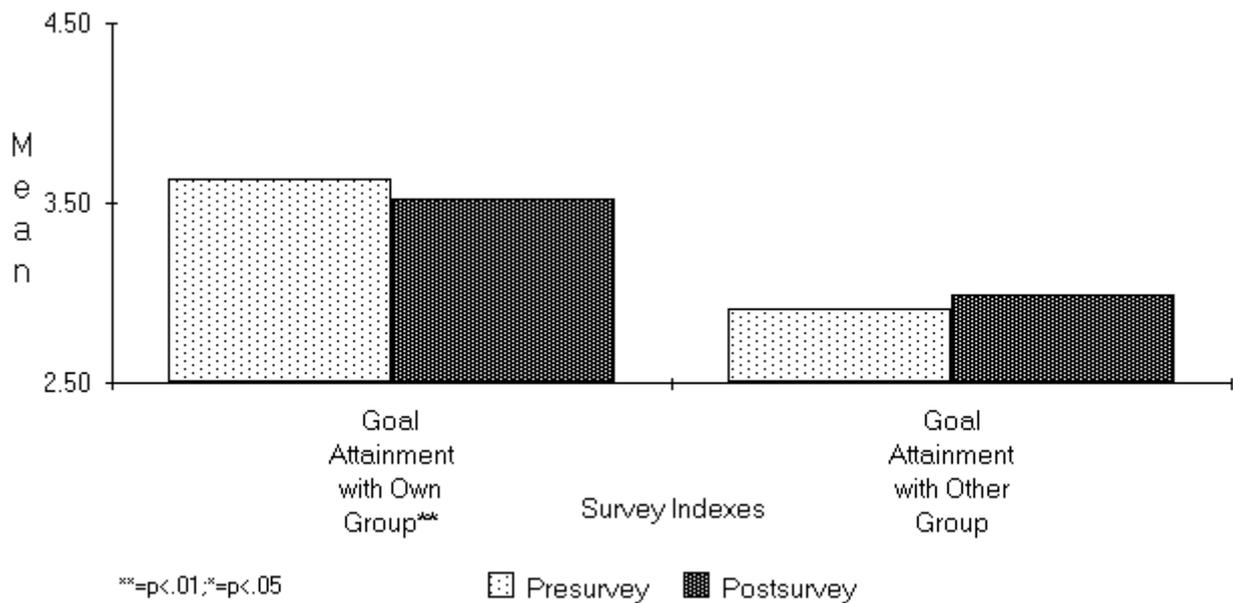


Figure 7.8 Mean Scores Pre and Post-Seminar for Goal Attainment Indexes

It is interesting to note the direction of changes in [Figure 7.8](#). For the scale focused on behaviors to achieve goal attainment in one's own group a negative change occurred (i.e., the levels of discussion and encouragement were reported to be higher before the training than afterwards). This might be explained as unrealistic assessments (i.e., over-optimistic view of their own groups) of this group behavior which were challenged by exposure to the case analyses and discussions during the [CRM](#) training.

The slight positive shift in perception of activity to share goals with other groups' approaches shows a tendency, but is not significant ($p > .15$). This trend may also be explained as a probable effect of training which provided new information (and more importantly new experiences with people from other departments during the training) about the similarity of goals pursued by other groups and functions in technical operations.

Post-training attitudes compared with two- and six-month follow-up. Scores for the four indexes were compared with results from the follow-up surveys. These results compare the post-test scores with the 50 or 41 individual managers whose post-training and 2-month or 6-month follow-up questionnaires could be matched. None of the differences are significant at the .05 level or better (in fact none approached significance any stronger than $p < .15$).

Taken together, the results show that attitudes about communication/ coordination and sharing command responsibility remain high in the months following [CRM](#) training. The goal attainment measures were not included in [Figures 7.9](#) and [7.10](#) and no statistically significant changes were found for them either.

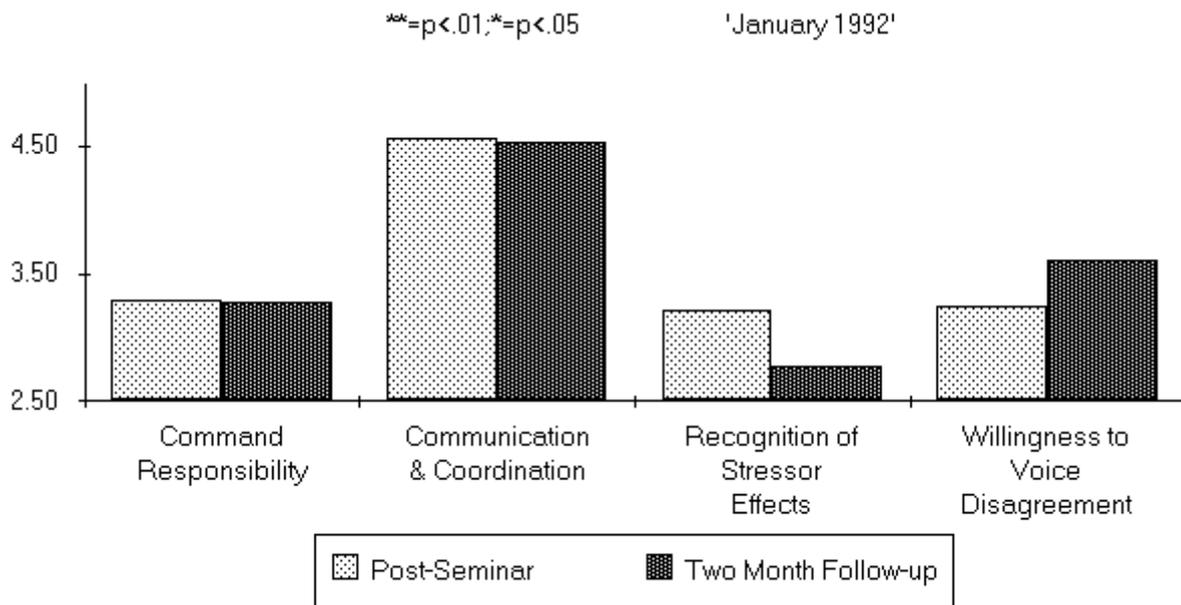


Figure 7.9 Mean Scores of Indexes for Post-Seminar and Two Month Follow-up for Technical Operations

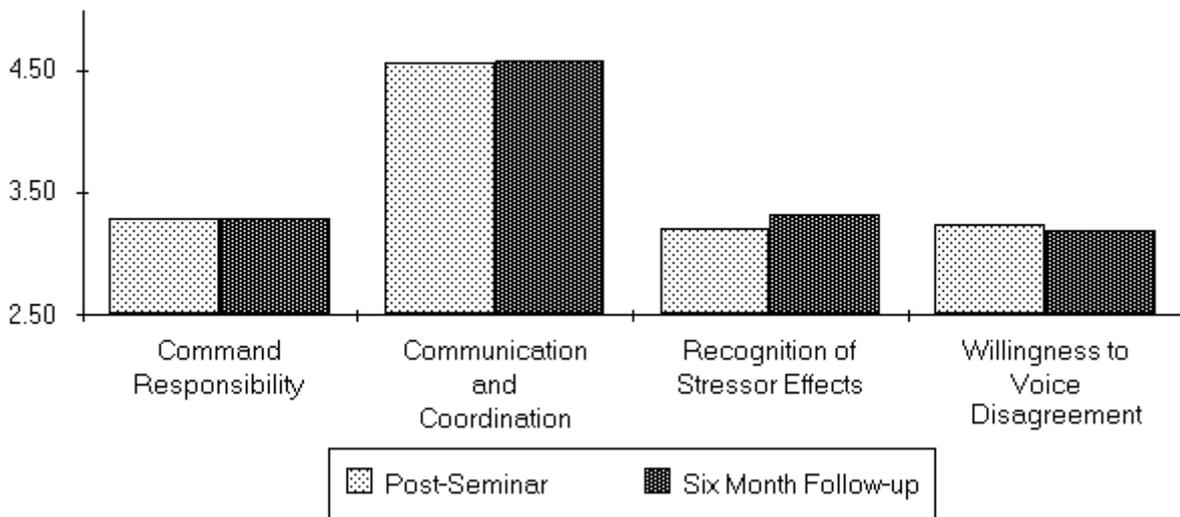


Figure 7.10 Mean Scores of Indexes for Post-Seminar and Six Month Follow-up for Technical Operations

7.6.3 Performance Changes Over Time

Pre- and post-training scores. The "before" and "after" scores for the 14 performance measures were compared and tested for statistical significance. Once again the Wilcoxon test, the Repeated-Measures analysis and the Multivariate Analysis of Variance (MANOVA) were used. These tests represent a unit by unit comparison of the scores of the work units for which the specific measures applied. Many of the differences are significant at the .05 level or better, but some are in the

opposite direction from that expected if the changes are to be attributed to changes from the CRM training. The next four figures (7.11, 7.12, 7.13 and 7.14) present the overall scores averaged over all the work units for which the performance measures apply. The statistically significant differences are displayed with the histograms in the figures as noted below.

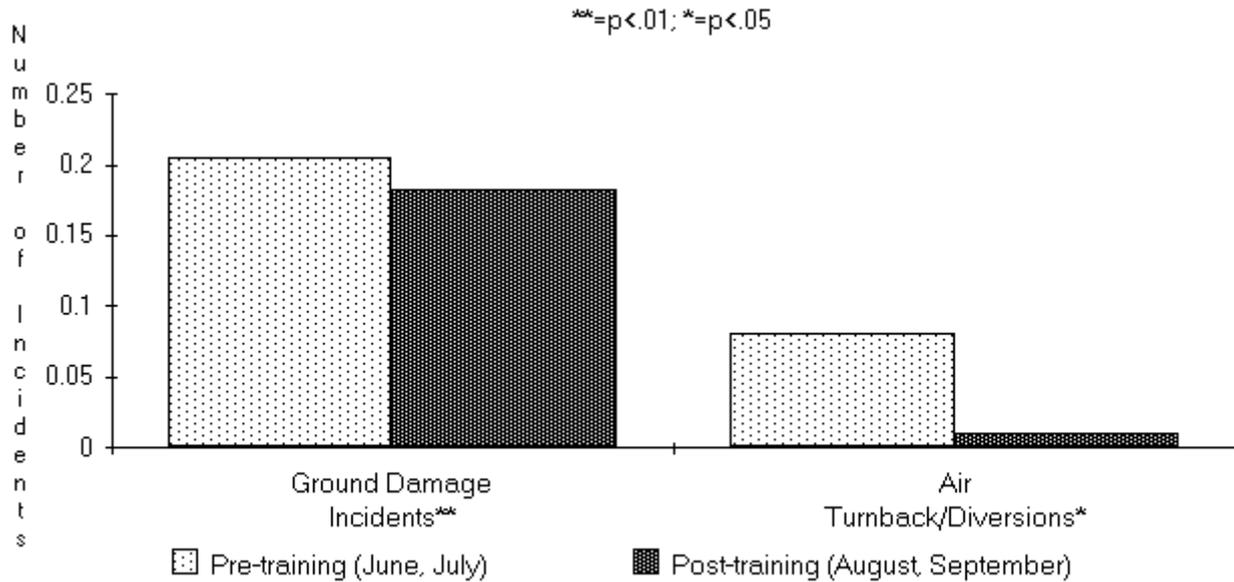


Figure 7.11 Mean Scores, Pre- and Post-Training, for Safety of Aircraft Performance Measures

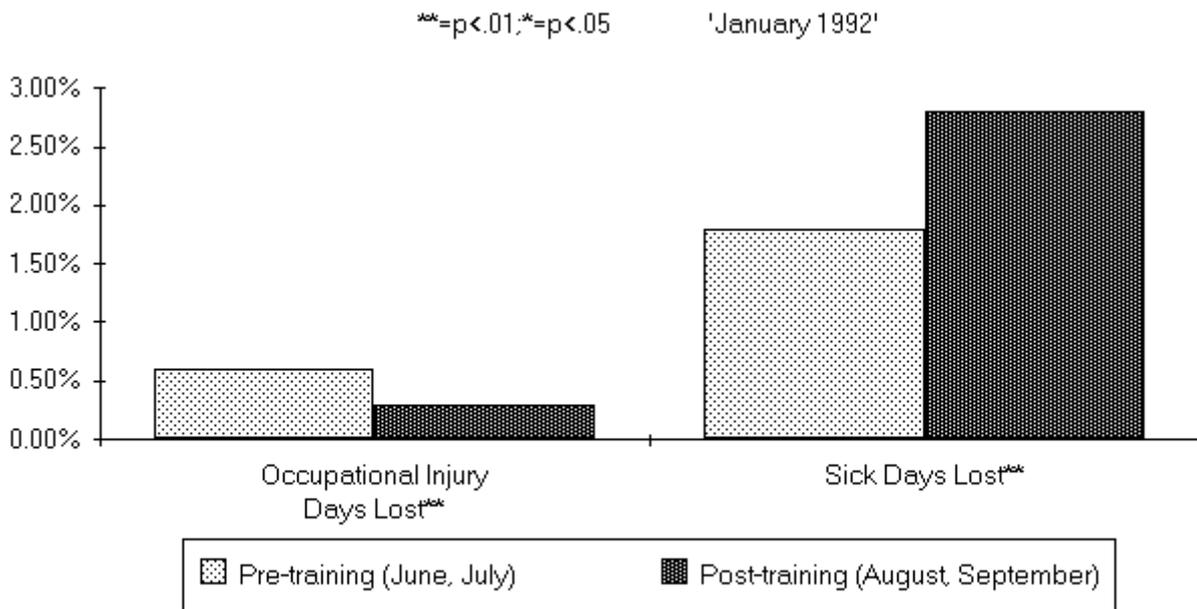


Figure 7.12 Mean Percentage, Pre- and Post-Training, for Personal Safety Performance Measures

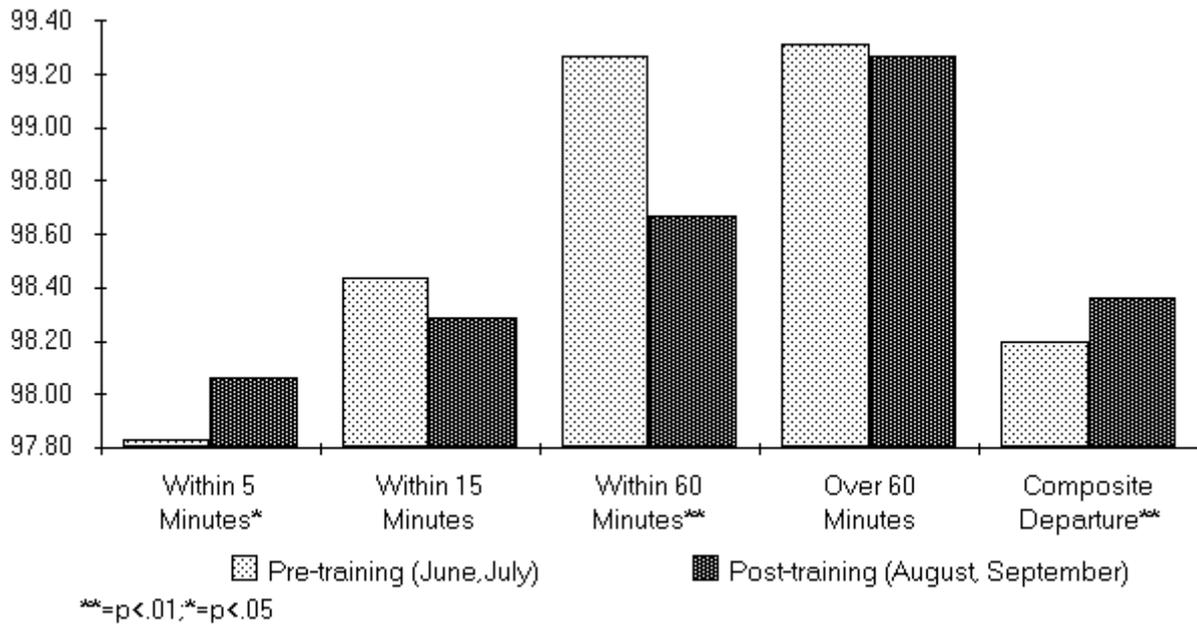


Figure 7.13 Mean Percentage, Pre- and Post-Training, for Departure Performance

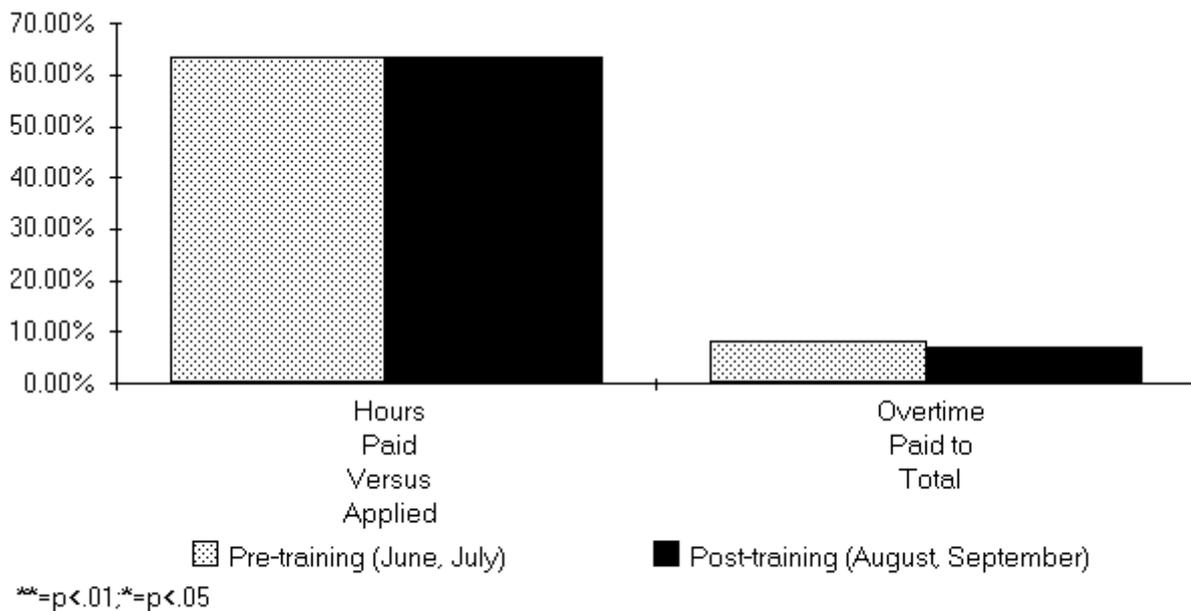


Figure 7.14 Mean Percentage, Pre- and Post-Training, for Efficiency Performance Measures

Measures of Safety Performance: Safety of Aircraft. [Figure 7.11](#) presents before-and after-training comparisons of two measures of aircraft safety performance -- number of Ground- Damage incidents for 44 work units, and number of Turnbacks or Flight Diversions caused by human error for 31 line stations. Both of these differences are in the predicted direction and significant ($p < .05$). Although there are ample alternative explanations for these improvements between June/July and August/September, it is important to note that [CRM](#) training could be responsible for at least some of the observed changes. The trainers and training administrators had predicted that these two safety measures would be among the most likely to improve because of the training.

Measures of Safety Performance: Personal Safety. [Figure 7.12](#) presents comparisons for the two

measures percent Days Lost to Occupational Injury, and percent Sick Days Lost for the 55 and 54 work units respectively for which data were available.

Both of these differences are statistically significant, but only one is in the predicted direction. Occupational Injury days decreased significantly from the pre- to post-training period. Sick days lost, on the other hand, significantly increased during the period. The training planners and managers had predicted that improved lost time due to injury would improve more readily than sick days lost following the training.

Measures of Dependability. [Figure 7.13](#) presents before- and after-training comparisons for departures within 5 minutes, within 15 minutes, within 60 minutes, over 60 minutes, and an overall departure performance measure which includes all of the above.

Three of these differences are statistically significant, two of which are in the expected direction. Departures within 5 minutes and the overall measure improve significantly from pre- to post-training periods and this result is consistent with the assessments of the trainers.

Percent of departures within 60 minutes significantly decreases (the reverse direction to that predicted) pre- to post-training. Initially, this result appears puzzling, but, upon reflection, it is not unexpected. The two performance indicators (departures within 5 minutes, and within 60 minutes), although not completely independent of one another in their measurement, can vary widely in practice. To illustrate this, consider the busy summer season when all line stations strive for on-time performance. The typical station (unable to substantially increase its manpower or overtime) could decide to sacrifice the schedule of aircraft that require as much as an hour to repair, in order to assign the line maintenance crews to those aircraft which can be dispatched "on- time." This tactic causes performance on the 5 minute mark to improve while performance on the 60 minute mark declines.

One other measure of dependability, "expendable parts service level," decreases (an unexpected result) between pre and post periods (pre = 95% < post = 93.3%; $Z = -1.83$, $n = 4$, $p = .07$). The other measures of dependability -- "number of delays due to late from maintenance" ($n=35$), "rotable parts service" ($n=8$) and "heavy checks on time to initial plan" ($n=4$) -- did not show measurable change in pre and post levels of performance.

Measures of Efficiency. [Figure 7.14](#) shows the pre- post comparisons for the percentage of hours applied to production and for overtime charged.

Although, neither measure shows change after training, the results of a series of analyses conducted to assess possible relationships between department and performance measured revealed that only overtime charged was related to department. These results show significant main effects of Pre-Post change ($F=4.14$, $P < .05$) **and** department ($F = 2.71$, $p < .05$), indicating all departments, except Quality control, reduced overtime after training. [Figure 7.15](#) displays these changes in bar chart form.

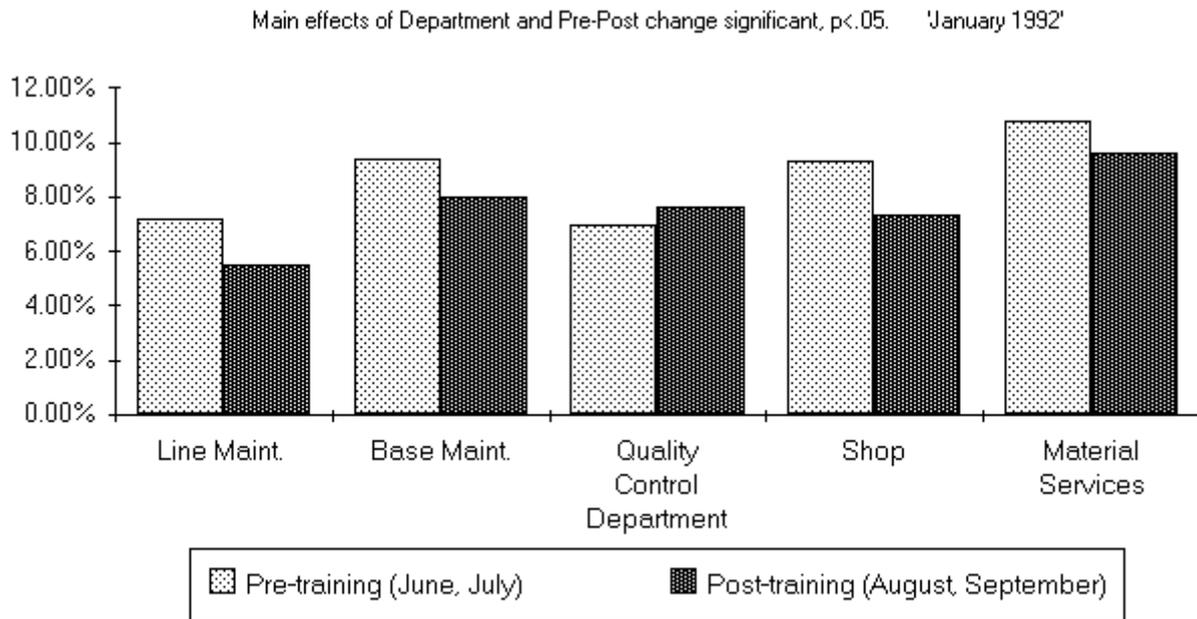


Figure 7.15 Mean Percentage, Pre- and Post-Training, for Overtime Paid to Total Department

7.6.4 Relationships Between Post-training Attitudes and Performance

The predictor variables to be tested include only the post-training results of the four attitude indices: "Sharing Command Responsibility," "Communication and Coordination," "Recognition of Stressor Effects," and "Willingness to Voice Disagreement." Spearman Rank-order Correlations that test the relationship between these indices and the 12 performance measures are presented in [Tables 7.2, 7.3, and 7.4](#). Because the present data are limited in sample size statistical significance conventions are extended to show not only probability levels of 1% and 5%, but also 10% and 15% to aid in examining patterns of expected results.

ATTITUDE INDEX			
"SHARING COMMAND RESPONSIBILITY"	"COMMUNICATION AND COORDINATION"	"RECOGNITION OF STRESSOR EFFECTS"	"WILLINGNESS TO VOICE DISAGREEMENT"
MEASURES OF SAFETY			
Number of Ground Damage Incidents ¹ (n=35)			
-.04	+.19	-.03	+.30*
% Occupational Injury Days Lost ¹ (n=37)			
+.28*	+.14	-.09	-.05
% Sick Days Lost ¹ (n=36)			
-.25 [†]	-.42	-.02	-.05
** p < .01; *p < .05			
[†] : These items score lower when performance is positive, thus the direction of the statistics are mirrored to reflect positive when relationships are in the expected direction.			

Table 7.2 Spearman-Rho Correlations Between Post-Training CRM/TOQ Results and Post-Training Maintenance Safety Performance

Table 7.2 Spearman-Rho Correlations Between Post-Training CRM/TOQ Results and Post-Training Maintenance Safety Performance

ATTITUDE INDEX			
"SHARING COMMAND RESPONSIBILITY"	"COMMUNICATION AND COORDINATION"	"RECOGNITION OF STRESSOR EFFECTS"	"WILLINGNESS TO VOICE DISAGREEMENT"
MEASURES OF DEPENDABILITY			
% Departure within 5 minutes (n=25)			
+32*	+30†	+27†	+24†
% Departure within 30 minutes (n=25)			
+29†	+3†	+26†	+5
% Departure within 60 minutes (n=25)			
+34*	+69	+29†	+2†
% Departure over 60 min, but not cancelled (n=25)			
+66	+69	+69	+23†
% Reliable parts available (n=4)			
+80†	+20	+20	+20
% Dependable parts available (n=4)			
+40	+40	+50	+80†
% Heavy checks on time to initial plan (n=4)			
+10*	+40	+10*	+80†
** p < .01 * p < .05 † p < .10 ‡ p < .15			

Table 7.3 Spearman-Rho Correlations Between Post-Training CRM/TOQ Results and Post-Training Maintenance Dependability Performance

Table 7.3 Spearman-Rho Correlations Between Post-Training CRM/TOQ Results and Post-Training Maintenance Dependability Performance

ATTITUDE INDEX				
"SHARING COMMAND RESPONSIBILITY"	"COMMUNICATION AND COORDINATION"	"RECOGNITION OF STRESSOR EFFECTS"	"WILLINGNESS TO VOICE DISAGREEMENT"	"WILLINGNESS TO VOICE DISAGREEMENT"
MEASURES OF EFFICIENCY				
Ratio of Hours paid for applied to production (n=5)				
-.77*	.23	-.20	-.31	-.31
% Overrun parts total weight (n=17)				
-.19†	-.09	-.12	-.20†	-.20†
** p < .01 * p < .05 † p < .10 ‡ p < .15				
† These items score lower when performance is positive, thus the direction of the statistics are mirrored to reflect positive when relationships are in the expected direction.				

Table 7.4 Spearman-Rho Correlations Between Post-Training CRM/TOQ Results and Post-Training Maintenance Efficiency Performance

Table 7.4 Spearman-Rho Correlations Between Post-Training CRM/TOQ results and Post-Training Maintenance Efficiency Performance

The matrix of results reveals that post-training attitudes predict the performance of the work units. These relationships are especially consistent for attitudes regarding the willingness of units to share command responsibility and to voice disagreement. These results are discussed below in more detail.

It's useful to note that the **total** number of **positive** and **significant** relationships relative to their proportion to the total number of tests presented in [Tables 7.2](#), [7.3](#), and [7.4](#) is substantially above that expected by chance alone. Twenty Spearman "Rhos" were less than or equal to a probability of .15, from a total of 48 tests conducted (20/48 = 42%); and seven of them were less than .05 (7/48 = 15%).

Measures of Safety Performance. Post-training responses for some of the attitude indices were associated with the measures of Safety performance (See [Table 7.2](#)). Specifically "Willingness to Voice Disagreement" was related to the number of ground-damage incidents. This relationship implies that the more willing managers were to voice disagreement the better that their units performed (e.g., had fewer incidents of ground damage).

Similarly, attitudes about "Sharing Command Responsibility" were positively associated with percentage of occupational injury days lost. That is, the more favorable the attitudes were about "Sharing Command Responsibility" the better were the outcomes; positive attitudes were associated with fewer days lost to occupational injury. Finally, attitudes on this same index, "Sharing Command Responsibility," were negatively related to percentage of sick days lost. This relationship is not in the expected direction, suggesting that more sick days are taken among units who are more willing to share command responsibility.

Measures of Dependability. Departures within 5 and 15 minutes of schedule (the two most salient Dependability measures) and 60 minutes display significant positive relationships with most or all of the four attitudes (See [Table 7.3](#)).

Taken together, these results suggest that positive attitudes toward the concepts learned in training have a facilitative effect on the departure performance of units. For example, those units favorable towards the sharing command responsibility tend to have better departure performance measured within 5, 15, and 60 minutes. A similar pattern of positive relationships between recognition of stressor effects and these departure performance assessments emerged as well. Line managers who advocate the importance of recognizing stress are more likely to guide their mechanics to perform well on the percentage of goal achieved for departure in 5, 15, and 60 minutes.

Warehouse service levels show few positive relationships with post-training attitudes (see [Table 7.3](#)). Percentage of rotatable parts available is moderately correlated ($p < .10$) with attitudes on the index "Sharing Command Responsibility." Warehouse service levels in expendable parts, however, show a negative relationship ($p < .10$) to "Willingness to Voice Disagreement."

Finally, the dependability measure, heavy-checks-on-time-to-plan, shows two perfect positive correlations with "Sharing Command Responsibility" and "Recognition of Stressor Effects" as well as a marginally significant but high correlation with "Willingness to Voice Disagreement." Although the number of units available for this measurement are unavoidably low (available $n=4$) the consistent pattern of positive relationships suggest that positive post-training attitudes are related to achieving the standards of the initial plan.

Measures of Efficiency. [Table 7.4](#) shows that attitudes about sharing command responsibility predict units' performance for one of the measures of efficiency, "ratio of hours paid to applied to production." This relationship suggests that the more the unit agrees with the principle of sharing command responsibility the better will be the ratio of efficiency. Such a positive relationship, however, is not evident when considering the percentage of overtime-paid-to-total-wage. In fact, there is a modest negative relationship between this efficiency measure and both "Sharing Command Responsibility" and "Willingness to Voice Disagreement."

7.7 QUALITATIVE DATA OBTAINED AFTER THE CRM TRAINING

Not all indicators used are quantitative. The case study approach is also beginning to provide evidence for the program's success. In addition, there are several open-ended questions included in the immediate post-training [CRM/TOQ](#), as well as in the 2- and 6-month follow-up questionnaires. These two potential sources of evidence will be described and illustrated below.

7.7.1 The Case Study Approach

One such anecdote involves participant reaction to the stress-management module of the [CRM](#)

course, and how it led to discovering how effective the course could be in heavy maintenance planning. Meeting objectives for completion times (called estimated time for return, or "ETRs") from base maintenance overhaul are always raised as examples of "high stress" parts of the job by managers in the CRM course. During a CRM session in November 1991, six months after the onset of the CRM training in Technical Operations, Planning and Maintenance Managers from one maintenance base **did not** report ETRs as a source of stress. With a little encouragement these managers revealed that, beginning two or three months before, Maintenance and Planning functions meet together frequently to confirm ETRs, or to change them if required. The net effect is an improvement not only in the performance to plan, but in the timeliness and quality of the aircraft delivered from base maintenance. The Director of Maintenance for that base, it turned out, had attended a CRM session three months earlier where he had shown an increased acceptance for the concepts -- having ended the two day training saying, "Maybe this stuff has a place in management after all." It should be noted that quantitative results, presented earlier, evidence the relationship between several attitude indices and heavy maintenance performance to plan.

7.7.2 Written Comments from the Post-training and Follow-up CRM/TOQ

"How will the CRM training be used on the job?" Immediately following the training this question was most frequently answered "Better communication" (active transferring of information), "Better listening" (a passive improvement made within the person), "Being more aware of others" (a passive, reactive behavior), and "Dealing better with others" (interactive, problem-solving).

"How did you use the CRM training on the job?" For the sample of respondents who returned the first 2-month and 6-month follow-up questionnaires (in October, 1991) the largest numbers, saying what they had used from the CRM training, reported actions in those same four categories. Tables 7.5 and 7.6 show these comparisons for the immediate Post-training survey with the two-month and six-month follow-up surveys, respectively.

RESPONSE	SURVEY PERIOD		
	POST-TRAINING WILL USE	TWO MONTH WAS USED	TWO MONTH WILL USE
Better Communication	26%	35%	17%
Better Listening	16	20	3
Be More Aware of Others	11	15	11
Deal Better with Others	10	25	25
Use in Daily Tasks	6	5	3

Table 7.5 Percentage of Written-in Responses Indicating How Training Will be Used and Was Used on the Job for Post-training and Two Month Follow-up

RESPONSE	SURVEY PERIOD		
	POST-TRAINING WILL USE	SIX MONTH WAS USED	SIX MONTH WILL USE
Better Communication	20%	20%	26%
Better Listening	16	17	6
Be More Aware of Others	11	14	5
Deal Better with Others	10	21	23
Use in Daily Tasks	6	2	2

Table 7.6 Percentage of Written-in Responses Indicating How Training Will be Used and Was Used on the Job for Post-training and Six Month Follow-up

Most respondents in both samples emphasized that they had tried to listen better, and to "deal better with others" (often specifying team work, decision making and problem solving as ways of doing this). The largest numbers in both the two- and six-month follow-up samples stated emphatically that "Communicating better," and "Using more teamwork" (the latter accounting for over half of the responses included in "dealing better with others") was what they intended to further use from what they learned in CRM training.

These open-ended responses tend to confirm that improved interpersonal behaviors have resulted from the positive attitudes which followed the training. Furthermore, the preferred behaviors tend to have shifted from those people could do by themselves (e.g., "be a better listener" and being more aware of others), to those behaviors which involve others, such as "communicating better," and dealing better with others.

"What could be done to improve the training?" Most answers to this question, immediately after training were "do nothing" or "it's fine as it is." Two and six months later the largest single category of answers was "do more of this" or "bring us back for follow-up training." [Table 7.7](#) presents these results for both the two month and six month follow-up surveys compared with the immediate post training questionnaire. From this sample, at least, participants clearly believe they would benefit from further [CRM](#) type training.

RESPONSE	SURVEY PERIOD		
	POST-TRAINING	TWO MONTH FOLLOW-UP	SIX MONTH FOLLOW-UP
Needs Nothing	35%	16%	12%
More Training & Follow-up	11	27	28
More Role Playing	7	11	4
Add Time to Training	7	5	4
More Case Studies	6	7	7
Better Mix of Participants	6	5	7

Table 7.7 Percentage of Written-in Responses Indicating How Training Might be Improved for Post-training, Two, and Six Month Follow-up

7.8 CONCLUSIONS

7.8.1 Strengths of the CRM Training Program

The [CRM](#) training for maintenance managers evaluated here has been successful, after as short a time as six months, in improving attitudes which appear to have influenced safe, dependable and efficient performance. Some specific findings should be emphasized:

1. The timing and content of the program has been well received by participants. The immediate evaluation of the training was even more positive than had been the case in other companies using [CRM](#) in maintenance as well as flight operations. [Figure 7.1](#) presented this dramatic difference.
2. The training produces a significant improvement in most attitudes measured. Changes in relevant attitudes measured immediately before and after training reveal strong and positive changes following training, for three of the four indices measured (cf., [Figure 7.4](#)). Follow-up results several months after training tend to confirm that most (if not all) these changes are stable ([Figures 7.9](#) and [7.10](#)).
3. Performance appears to improve due to CRM training. Changes in most maintenance performance measured before and after the [CRM](#) training sessions examined in the present paper show significant changes in the expected direction (e.g., [Figures 7.11](#), [7.12](#) and [7.13](#)).
4. Specific attitude changes seem to cause specific performance. A pattern of significant relationships in the expected direction was noted between the post-training attitudes and the post-training performance. Those measures predicted (by the program's trainers and managers) to be especially sensitive to the effects of the [CRM](#) training were affected (cf.,

[Tables 7.2, 7.3](#)). In particular, positive attitudes about "Sharing Command Responsibility" and "Willingness to Voice Disagreement" both show the most association with improved performance.

5. More active change is occurring. Finally, the program is reported in follow-up surveys to be stimulating changes from the "easier" more passive behaviors (such as "better listening"), to the more interactive ones in working with others, such as holding meetings and undertaking joint problem-solving. Anecdotal evidence is beginning to confirm these changes and their impact on performance.

Continuous Improvement. A few of the strengths, such as the final one noted above, also provide guidance where the training design for this, and future programs, can be improved.

The results presented in this report reflect the program as it was in November, 1991. The training facilitators and administrators of the program receive abundant verbal participant feedback at the end of each training session. They have also received the earlier reports issued through this research project. A watchword of those managing this program is its flexibility and adaptability to the needs (as well as constraints) of the company, as well as to increased learning about this new kind of maintenance training. Several of the concerns and opportunities noted below have been successfully addressed by the trainers since November, 1991. These efforts will be briefly described.

The ambitious syllabus for the two-day training evaluated here ([Appendix A1](#), [Appendix A2](#), [Appendix A3](#), and [Appendix A4](#)) contains many important content categories; and it has little room for added exercises or activities. This tight program has created a dilemma for the trainers from its onset. For them, the opportunities for program improvement usually require classroom time to exploit. The trainers have devised an innovative solution to this dilemma in at least one area of potential improvement described below.

7.8.2 Opportunities for Improvement in Maintenance CRM Training

1. Help participants plan what and how to use their new-found skills back at work. The finding that a substantial portion of participants months later report that they are ready and willing to try more active behaviors -- those actually involving others as well as themselves -- suggests that the training program might be improved to help the "graduates" to design or develop their own approach to implementation. Such implementation could help participants actually plan, during the training, how and what they would try to change when they returned to work. Part of this awareness could include research results (such as those contained in this report) for participants so that they understand that listening and stress-management skills are useful and important, but that the assertiveness and team leadership skills could impact safety even more. Another avenue to impact active improvement would be to expand the course module on assertiveness training.

2. Focus directly on assertiveness skill training. The research evidence presented in this report shows that, of the four major attitude clusters derived from the [CRM/TOQ](#) questionnaire, "Willingness to Voice Disagreement" forms a mixed picture. Although assertiveness is mentioned throughout the training, the average scores reveal less positive attitudes about voicing disagreement (or wanting to "avoid conflict") following the course than before it. This shift is most marked for the youngest and the oldest; and for the lowest and highest in the management hierarchy. Those in the middle of the age range and hierarchy tend to show some shift toward addressing disagreements." What is intended of the [CRM](#) training in this regard? As an active social skill, assertiveness forms a theme throughout the two-day program, but it hasn't been emphasized at the expense of other aspects of the course. Perhaps more time directly with assertiveness skills is required. There are ample materials already developed and included in the participants' [CRM](#) handbook, which contain theory and skill practice sections, on assertiveness, but limited classroom time has not permitted the intensive coverage this material deserves. Greater use of role-playing exercises could also help participants learn how to be assertive and participative, and to give them confidence that they can skillfully employ these behaviors, but this too requires time.

Trainers in the present program feel very strongly about the need for practical application of all of the [CRM](#) concepts -- and assertiveness in particular. They are also sensitive about the contrived nature of even the best role-playing cases as seen by the very pragmatic audience they are training. In an effort to emphasize "real life" rather than contrived situations, the trainers have replaced one of the two role-plays dealing with "supporting and confronting others" (a major aspect of assertiveness skill) with intensive small-group and whole-group discussions dealing with the same concepts, but focused on personal illustrations volunteered from the participants. It is too soon to quantitatively measure the effects of this innovation, but trainers' initial reports are encouraging.

The results of association between post-training attitudes and unit performance (i.e., [Table 7.2](#)) tend to show that positive feelings about assertiveness and sharing authority are most related to safety and dependability. Further improving and/or expanding the training design could well leverage this advantage by increasing the numbers of participants with a positive opinion about assertiveness.

3. Plan and publicize recurrent [CRM](#) training. Another way to develop active leadership and follower skills, while recognizing the ambitious syllabus in a two-day training course, would be to plan and publicize a [CRM](#) program that includes a follow-up training module (or even successive recurrent training). Given that many [CRM](#) graduates say that they would like to receive additional training like this, expansion of the program curriculum could provide a real advantage. For instance, the [CRM](#) training program could be designed to first awaken awareness of [CRM](#) and develop easily attainable successes and then subsequently develop and exercise team and assertiveness skills.

7.9 REFERENCES

- Borg, W.R., & Gall, M. D. (1989). *Educational research* (5th ed.). New York: Longman.
- Federal Aviation Administration. (1991). *Challenges in aviation human factors: The national plan (draft)*, Washington, D.C.
- Federal Aviation Administration, Office of Aviation Medicine. (1991). "Reliability and Validity. Some measurement characteristics of Crew Resources Management/Technical Operations Questionnaire (CRM/TOQ)" (Task 2 Report).
- Federal Aviation Administration, Office of Aviation Medicine. (1991). "Assessment of Maintenance Performance Measures, and a Demonstration of CRM Training Evaluation Analysis Plan" (Task 3 Report).
- Fotos, C.P. (1991, August 26). Continental applies CRM concepts to technical, maintenance corps, [and] Training stresses teamwork, self-assessment techniques. *Aviation Week & Space Technology*, 32-35.
- Geirland, J.G., & Cotter, J. J. (1990). *Social analysis questionnaire*. No. Hollywood, CA: John J. Cotter & Associates, Inc.
- Gregorich, S.E., Helmreich, R.L., & Wilhelm, J.A. (1990). The structure of cockpit management attitudes. *Journal of Applied Psychology*, 75, 682-690.
- Helmreich, R.L. (1979). Social psychology on the flight deck. In G.E. Cooper, M.D. White & J.K. Lauber (eds.) *Proceedings of the NASA Workshop on Resource Management Training for Airline Flight Crews*, NASA CP-2120.
- Helmreich, R.L. (1989). Evaluating the effect of cockpit research management training (Abstract). *Human Error Avoidance Technical Conference*. Washington, D.C.
- Helmreich, R.L., Foushee, H.C., Benson, R., & Russini, R. (1986). Cockpit management attitudes: Exploring the attitude-performance linkage. *Aviation, Space and Environmental Medicine*, 57,

1198-1200.

Helmreich, R.L., Predmore, S.C., Irwin, C.M., Butler, R.E., Taggart, W.R., Willhelm, J.A., & Clothier, C. C. (1991). Beyond CRM: Papers from the NASA/University of Texas Crew Performance Project. *Proceedings of the Sixth International Symposium on Aviation Psychology*. Columbus OH: Ohio State University.

Norusis, M. (Ed.). (1990). *SPSS Base System User's Guide*. Chicago, Illinois.

Taggart, W. (1990). Introducing CRM into maintenance training. *Proceedings of the Third International Symposium on Human Factors in Aircraft Maintenance and Inspection*. Washington, D.C.: Federal Aviation Administration.

Chapter Seven Appendix A

Technical Operation Division Crew Coordination Concepts Syllabus

<u>Module</u>	<u>Time</u>	<u>Facilitator</u>	<u>Method</u>	<u>Objective</u>
Day 1 Introduction	8:15 (30 min)	TO	Introduce self. Cover facilities, restrooms, telephones, breaks, lunches, closing time. What have you heard about course? Review PURPOSE, OBJECTIVES. WORKING AGREEMENTS. Collect Pre-seminar questionnaire. Explain purpose & importance of ID#. Group introduces themselves.	Position program Get group talking Set the tone of the Clarify expectations Remove teacher's
Portland Video	8:45am (45 min)	HF	Show video; identify resource mgmt. problems; relate Portland problems to work place. Prepare flipchart.	Attention getter. problems become
Expectations	9:30am (15 min)	HF	Develop expectations of individuals in course. Write on chart, compare to course objectives.	Get expectations illustrate differences
Break	9:45am (15 min)			
Testing Assumptions/ DC-10 Video GUM	10:00am (120 min)	TO	Introduce concept of Perception vs Reality Show DC-10 Video GUM. Tablework: What were "Chain of Events" that led to the accident? Discuss "ASSUMPTIONS" that led or contributed to accident. How can we test our assumptions? What could each person have done differently?	Test Assumptions Advocacy - "Speak Inquiry - "Ask Qu Active Listening -
Lunch	12:00pm (60 min)			

Module	Time	Facilitator	Method
Behavior	1:00pm (75 min)	HF	Discuss use of Instruments; admin SDI; develop Behav Dim Model; discuss concept of "Assertive Behavior."
Break	2:15pm (15 min)		
Behaviors Comt'd	2:30pm (60 min)	HF	Scores SDI; Interpret SDI; draw arrows; discuss/apply "Assertiveness" Behavior modification approx. for effective supervision.
Break	3:30pm (15 min)		
Stress Management	3:45 pm (60 min)	TO/HF	TO-Introduce/understand stress What are sources of stress in our jobs? Develop list of Stressors. HF-Identify body's reaction to stress. Ways to deal with stress. TO-Work on how to deal with 2 examples. ETR'S, MANPOWER, PARTS AVAILABILITY

<u>Module</u>	<u>Time</u>	<u>Facilitator</u>	<u>Method</u>
Day 2 Opening DC-10 Video GUM 041	8:15am (30 min)	TO	Show GUM 041 video with breakfast. Address loose ends from Day 1; Outline schedule for Day 2. Review lessons learned in context of video (Gum 041)
Sub Artic Survival	8:45am (60 min)	TO	Purpose of simulation. Make individual decisions (step 2).
Break	9:45am (15 min)		
SAS II	10:00am (60 min)	HF	Define/develop rational decision process; apply process to simulation; complete SAS; critique team effect develop lessons learned/apply to job.
Break	11:00am (15 min)		
Norms/ EAL 855	11:15am (45 min)	TO	Identify role of norms in Tech Ops. Introduce concept, give examples. Allow time to review EAL 855. Assign each table a role. What were norms that led to accident? Groups develop lists of good/bad CO Norms. Discuss how to manage norms and prevent accidents.
Lunch	12:00pm (60 min)		

<u>Module</u>	<u>Time</u>	<u>Facilitator</u>	<u>Method</u>	<u>Objective</u>
Listening & Communicating	1:00pm (60 min)	HF	Tie into SAS; Sleep exercise Communication model; listening barriers; listening tips.	Listening is a lead listening limits q making.
Break	2:00pm (15 min)			
Supporting/ Confronting	2:15pm (60 min)	TO/HF	Use interactive dilemmas. Conduct 1st dilemma. Critique/lessons learned from 1st. Conduct second dilemma.	Application of Bel Understand the va is important. Use skills/lack of and
Break	3:15pm (15 min)			
Wrap-up Evaluation/ Questionnaire	3:30pm (45 min)	To/HF	Take Home concepts? USC/CAL Questionnaire explanation I.D #s on questionnaires.	Pledge to do some Feedback for prog feel good about ex

Chapter Seven Appendix B

Date _____

Airlines Technical Operations Division CCC WorkShop Survey

BASELINE

Please answer by writing beside each item the number that best reflects your personal opinion.
Choose the number from the scale below.

*** SCALE ***

1	2	3	4	5	
Disagree	Disagree	Neutral	Agree	Agree	
Strongly	Slightly	Slightly	Strongly		

- ___1. Tech. Ops. team members should avoid disagreeing with others because conflicts create tension and reduce team effectiveness.
- ___2. It is important to avoid negative comments about the procedures and techniques of other team members.
- ___3. Casual, social conversation on the job during periods of low workload can improve Tech. Ops. team coordination.
- ___4. Good communications and team coordination are as important as technical proficiency for aircraft safety and operational effectiveness.
- ___5. We should be aware of and sensitive to the personal problems of other Tech. Ops. team members.
- ___6. The manager, supervisor, or asst. supervisor in charge should take hands-on control and make all decisions in emergency and non-standard situations.
- ___7. The manager, supervisor, or addt. supervisor in charge should verbalize plans for procedures or actions and should be sure that the information is understood and acknowledged by the other Tech. Ops. team members.
- ___8. Tech, Ops. team members should not question the decisions or actions of the manager, supervisor, or asst. supervisor except when they threaten the safety of the operation.
- ___9. Even when fatigued, I perform effectively during critical phases of work.

*** SCALE ***

1	2	3	4	5	
Disagree	Disagree	Neutral	Agree	Agree	
Strongly	Slightly	Slightly	Strongly		

- ___10. Managers, supervisors and asst. supervisors should encourage questions during normal operations and in special situations.
- ___11. There are no circumstances where the subordinate should assume control of a project.
- ___12. A debriefing and critique of procedures and decision after each major task is an important part of developing and maintaining effective team coordination.

- ___13. Overall, successful Tech. Ops. management is primarily a function of the manager's, supervisor's, or asst. supervisor's technical proficiency.
- ___14. Training is one of the manager's most important responsibilities.
- ___15. Because individuals function less effectively under high stress, good team coordination is more important in emergency or abnormal situations.
- ___16. The start-of-shift team briefing is important for safety and for effective team management.
- ___17. Effective team coordination requires each person to take into account the personalities of other team members.
- ___18. The responsibilities of the manager, supervisor, or asst. supervisor include coordination between his or her work team and other support areas.
- ___19. A truly professional manager, supervisor or asst. supervisor can leave personal problems behind.
- ___20. My decision-making ability is as good in abnormal situations as in routine daily operations.

***** SCALE *****

1	2	3	4	5
Disagree	Disagree	Neutral	Agree	Agree
Strongly	Slightly	Slightly	Strongly	

In the following questions, "my management group" refers to those people who report to the same manager that I do.

- ___21. I am kept informed by others in my management group about the goals and objectives of this organization (e.g., cost, quality, service, etc.).
- ___22. Work goals and priorities are understood and agreed to by members of my management group.

In the following items, "my work group" refers to those people who report to me.

- ___23. Employees in my work group receive detailed feedback regarding the organization's performance.
- ___24. If employees in my work group disagree with the goal and priorities that have been established, they feel free to raise their concerns with supervision.
- ___25. Employees in other groups within Tech. Ops. plan and coordinate their activities effectively together with people in my work group.
- ___26. Employees in other groups, departments and divisions through the company act as if they share many of the same organizational goals that we do.

BACKGROUND INFORMATION

Year of Birth _____

Total Years at _____

Sex (M or F) _____

Current
Department

_____ Maintenance

_____ Engineering

_____ Quality Control
 _____ Planning
 _____ Logistics
 _____ Shop
 _____ Material Services

Job Title: _____

Years in present position: _____

Past Experience/Training (# of Years): _____

Military _____

Trade School _____

College _____

Other Airlines _____

Chapter Seven Appendix C

Date _____

AIRLINES TECHNICAL OPERATION DIVISION CCC WORKSHOP SURVEY

"Two-month Follow-up" Questionnaire

Please enter the five digit Personal Identification Number that you selected at the beginning of the seminar.

Identification Code _____

Now, please answer by writing beside each item the number that best reflects your personal attitude. Choose the number from the scale below. All data are strictly confidential.

***** SCALE *****

1	2	3	4	5	
Disagree	Disagree	Neutral	Agree	Agree	
Strongly	Slightly	Slightly	Strongly		

- ____1. Technical Operations team members should avoid disagreeing with others.
- ____2. It is important to avoid negative comments about the procedures and techniques of other team members.
- ____3. Casual, social conversation on the job during periods of low worked can improve Technical Operations team coordination.
- ____4. Good communications and team coordination are as important as technical proficiency for aircraft safety and operational effectiveness.
- ____5. We should be aware of and sensitive to the personal problems of other Technical Operations team members.
- ____6. The manager, supervisor, or assistant supervisor in charge should verbalize plans for

procedures or actions and should be sure that the information is understood and acknowledged by the other Technical Operations team members.

___7. The manager, supervisor, or assistant supervisor in charge should verbalize plans for procedures or actions and should be sure that the information is understood and acknowledged by the other Technical Operations team members.

*** SCALE ***

1	2	3	4	5
Disagree	Disagree	Neutral	Agree	Agree
Strongly	Slightly	Slightly	Strongly	

___8. Technical Operations team members should not question the decisions or actions of the manager, supervisor, or assistant supervisor except when they threaten the safety of the operation.

___9. Even when fatigued, I perform effectively during critical phases of work.

___10. Managers, supervisors, and assistant supervisors should encourage questions during normal operations and in special situations.

___11. There are no circumstances where the subordinate should assume control of a project.

___12. A debriefing and critique of procedures and decisions after each major task is an important part of developing and maintaining effective team coordination.

___13. Overall, successful Technical Operations management is primarily a function of the manager's, supervisor's, or assistant supervisor's technical proficiency.

___14. Training is one of the manager's most important responsibilities.

___15. Because individuals function less effectively under high stress, good team coordination is more important in emergency or abnormal situations.

___16. The start-of-shift team briefing is important for safety and for effective team management.

___17. Effective team coordination requires each person to take into account the personalities of other team members.

___18. The responsibilities of the manager, supervisor, or assistant supervisor include coordination between his or her work team and other support areas.

___19. A truly professional manager, supervisor, or assistant supervisor can leave personal problems behind.

___20. My decision-making is as good in abnormal situations as in routine daily operations.

*** SCALE ***

1	2	3	4	5
Disagree	Disagree	Neutral	Agree	Agree
Strongly	Slightly	Slightly	Strongly	

In the following questions, "my management group" refers to those people who report to the same manager that I do.

___21. I am kept informed by others in my management group about the goals and objectives of this organization (e.g., cost, quality, service, etc.).

___22. Work goals and priorities are understood and agreed to by members of my management group.

In the following items, "my work group" refers to those people who report to me.

___23. Employees in my work group receive detailed feedback regarding the organization's

performance.

____24. If employees in work group disagree with the goals and priorities that have been established, they feel free to raise their concerns with supervision.

____25. Employees in other groups within Technical Operations plan and coordinate their activities effectively together with people in my work group.

____26. Employees in other groups, departments and divisions throughout the company act as if they share many of the same organizational goals that we do.

27. How useful has the CCC training been for others? (Circle one)

A Waste of Time Slightly Useful Somewhat Useful Very Useful Extremely Useful

28. How much has the CCC training changed your behavior on the job? (Circle one)

No Change A Slight Change A Moderate Change A Large Change

29. What changes have you made as a result of the CCC training?

1 2 3 4 5
Disagree Disagree Neutral Agree Agree
Strongly Slightly Slightly Strongly

30. How will you further use the CCC training in the coming months?

31. Looking back on it now, what aspects of the training were particularly good?

32. What do you think could be done to improve CCC training?

Year of Birth _____

Total years at _____

Sex (M or F) _____

CURRENT DEPARTMENT

- _____ Line Maintenance
- _____ Base Maintenance
- _____ Quality Control
- _____ Planning
- _____ Shop
- _____ Material Services
- _____ Engineering
- _____ Other

Work Location - City

Job Title: _____

Years in present position: _____

Past experience/training (# of years): _____

Military _____

Trade School _____

College _____

Other Airline _____

This completes the questionnaire. Thanks for your help.

Phase III, Volume 1 Progress Report (1993)

Galaxy Scientific Corporation

Pleasantville, NJ 08323

Office of Aviation Medicine

Federal Aviation Administration

Washington, DC 20591

August 1993

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.

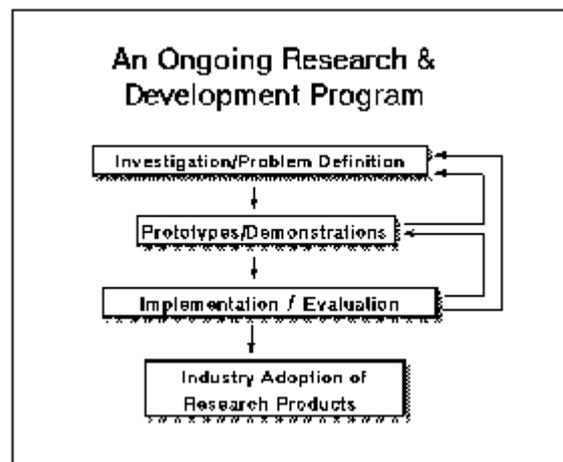


Figure 1 The Research Program

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.

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 cancelations, 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.

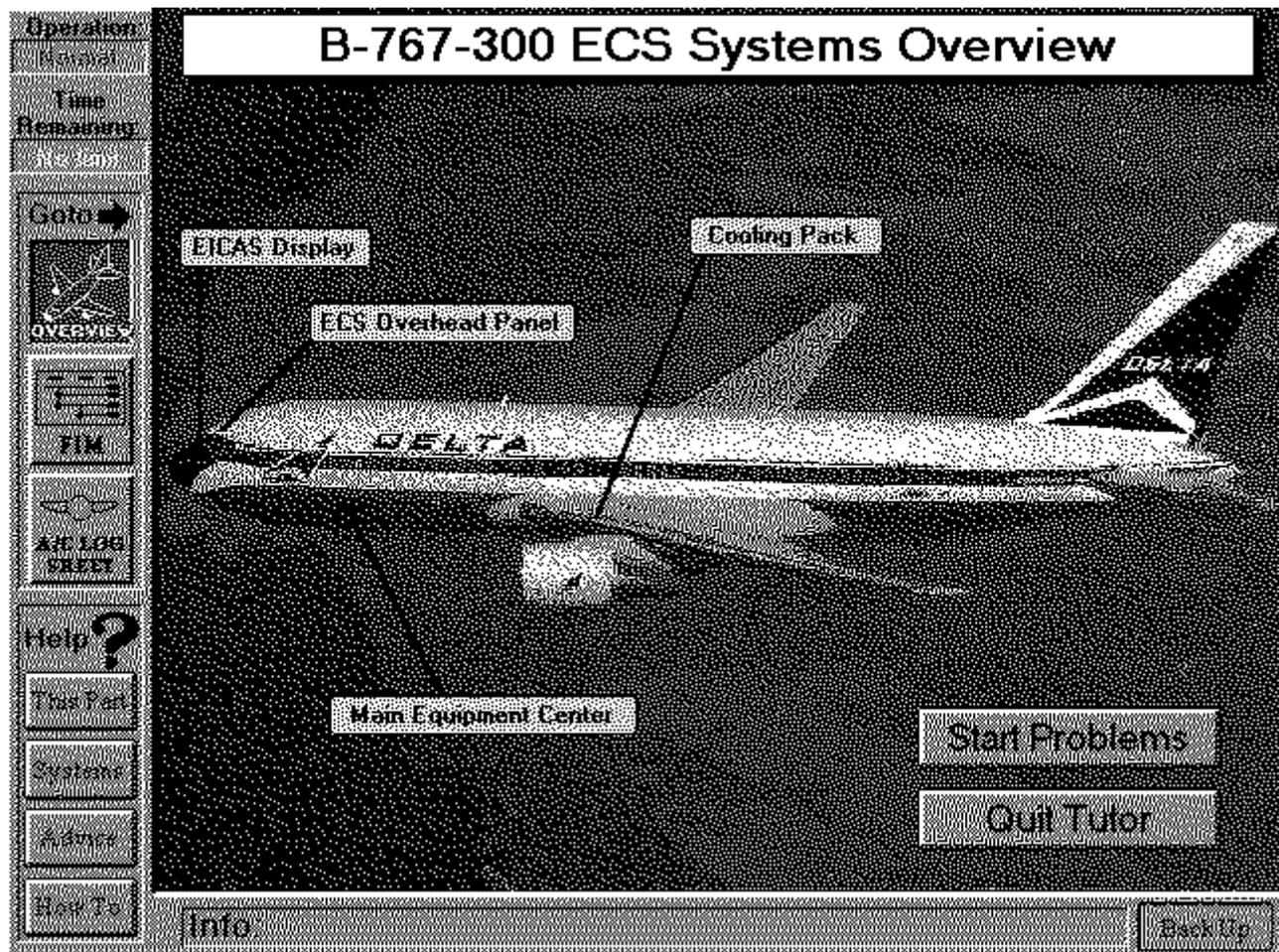


Figure 2.1 A Sample Screen from the ECS Tutor

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.

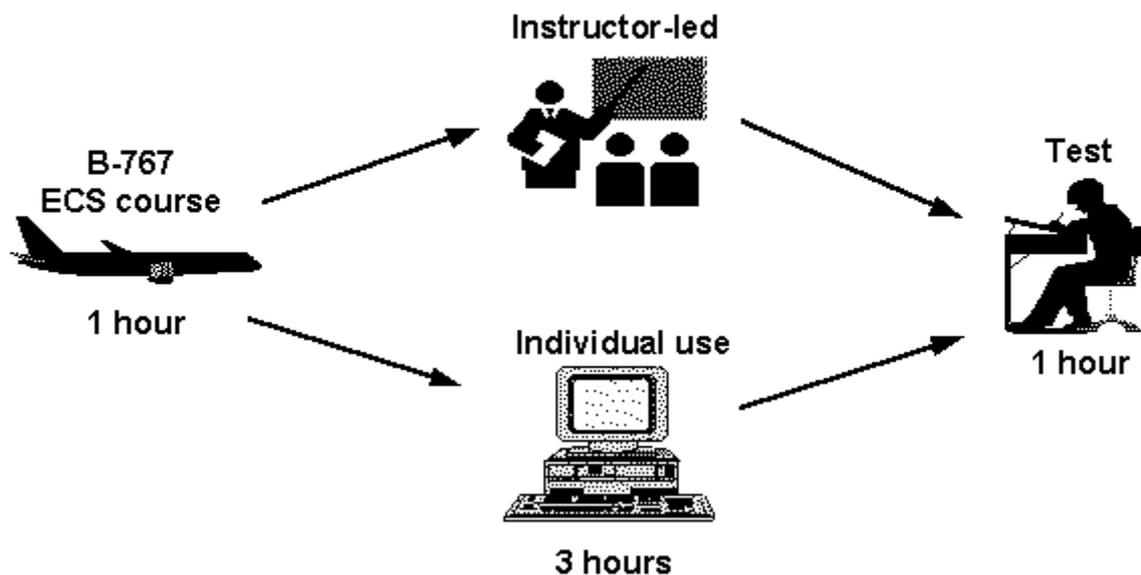


Figure 2.2 Design of the Experiment

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

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. [Figures 2.5a](#), [2.5b](#), [2.5c](#), and [2.5d](#) in the [Appendix A](#) shows the score distributions of the two groups for the four

examination sections.

	<u>Instructor-led</u>	<u>Individual-use</u>
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%

Table 2.1 Summary of Test Statistics

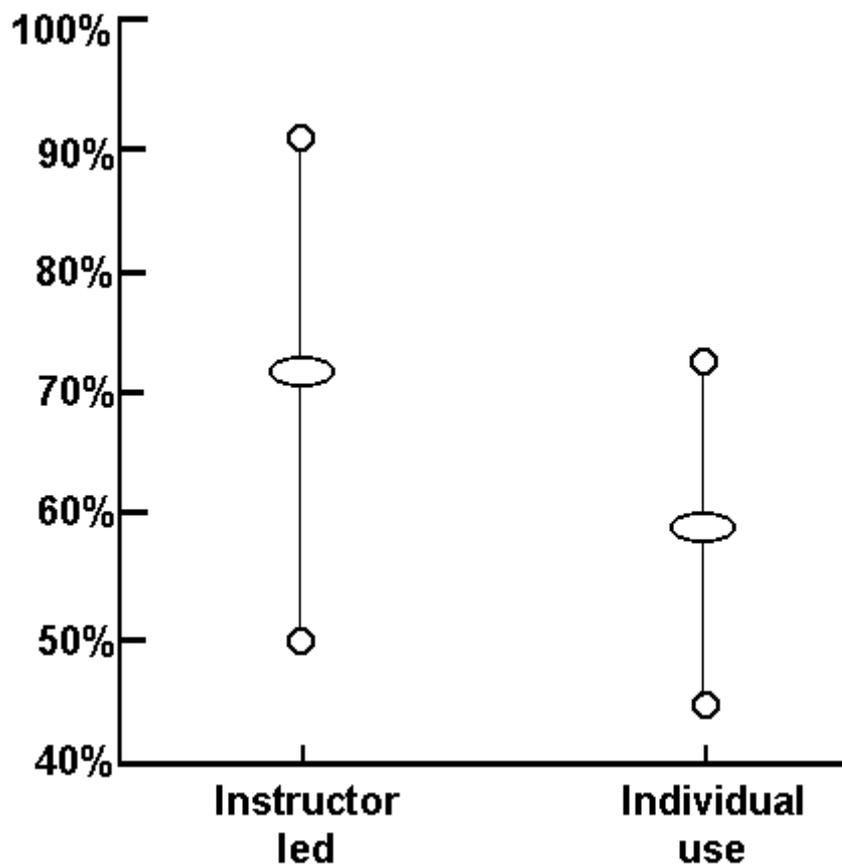


Figure 2.3 Overall Test Scores for Instructor-led (A) and Individual-use (B) Groups

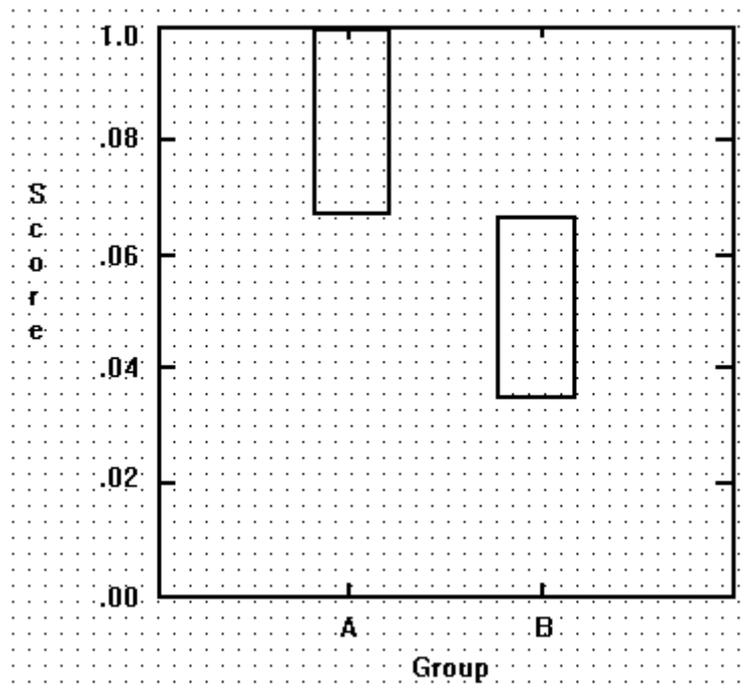


Figure 2.5a Distribution of Scores for Examination Sections A.

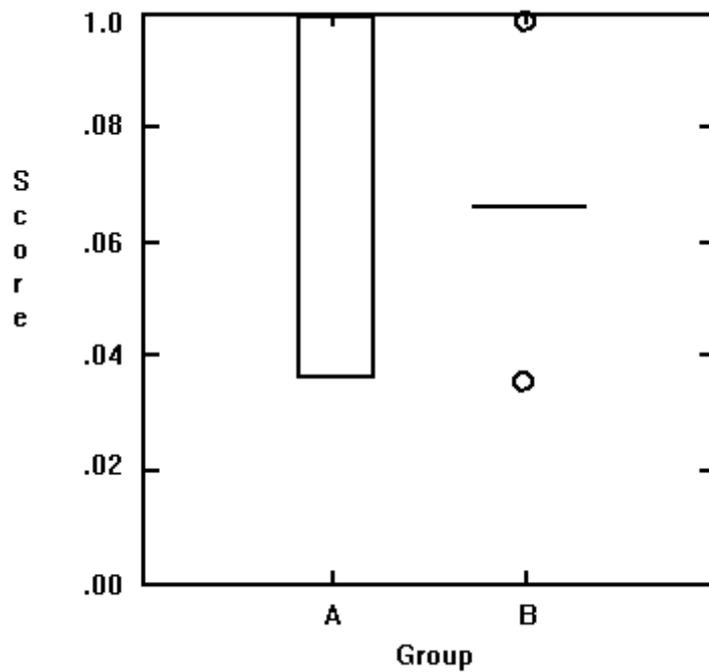


Figure 2.5b Distribution of Scores for Examination Sections B.

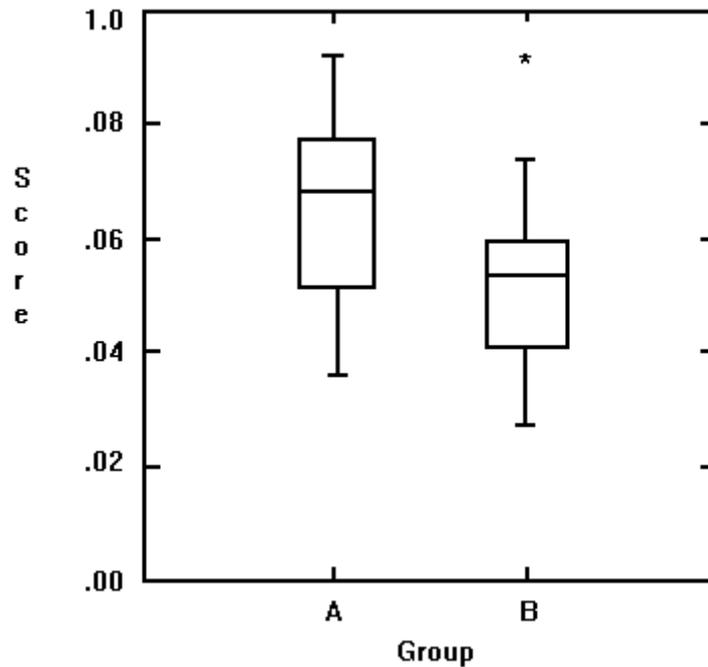


Figure 2.5c Distribution of Scores for Examination Sections C.

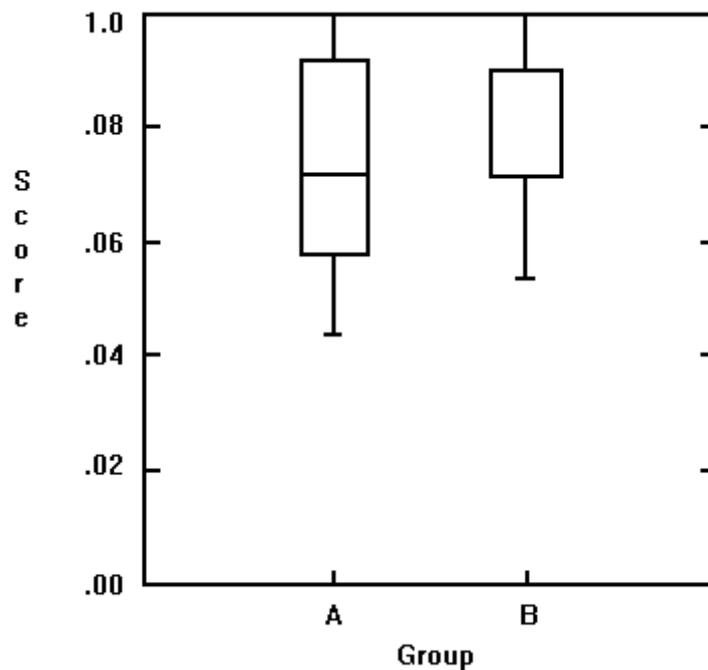


Figure 2.5d Distribution of Scores for Examination Sections D.

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.

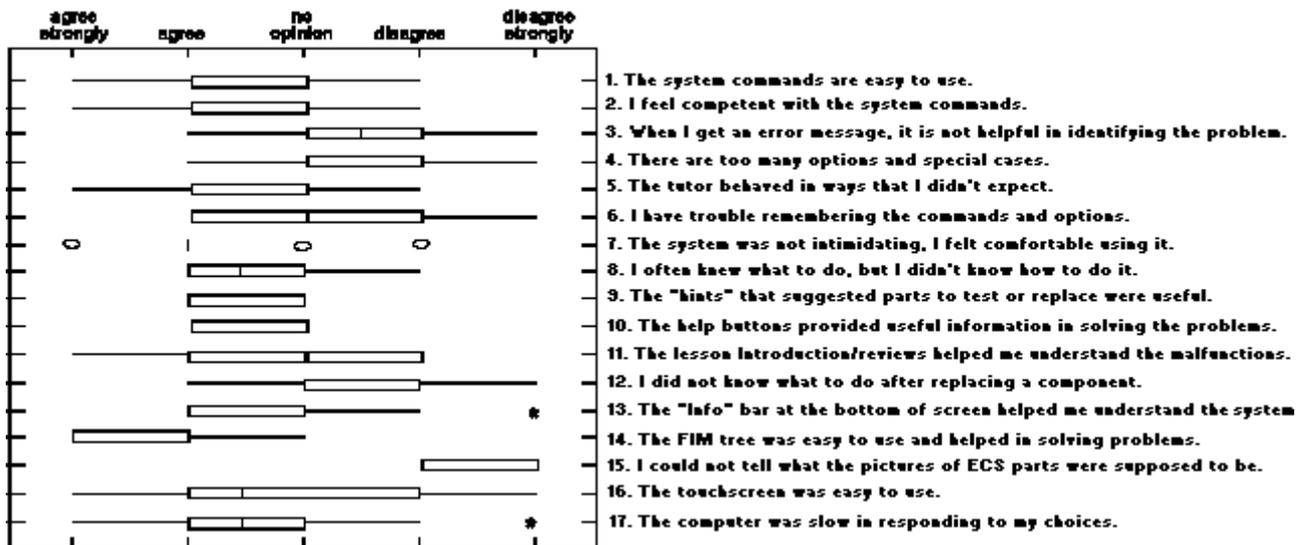


Figure 2.6 Examination Response Distribution for the Individual-use Group

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.

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).

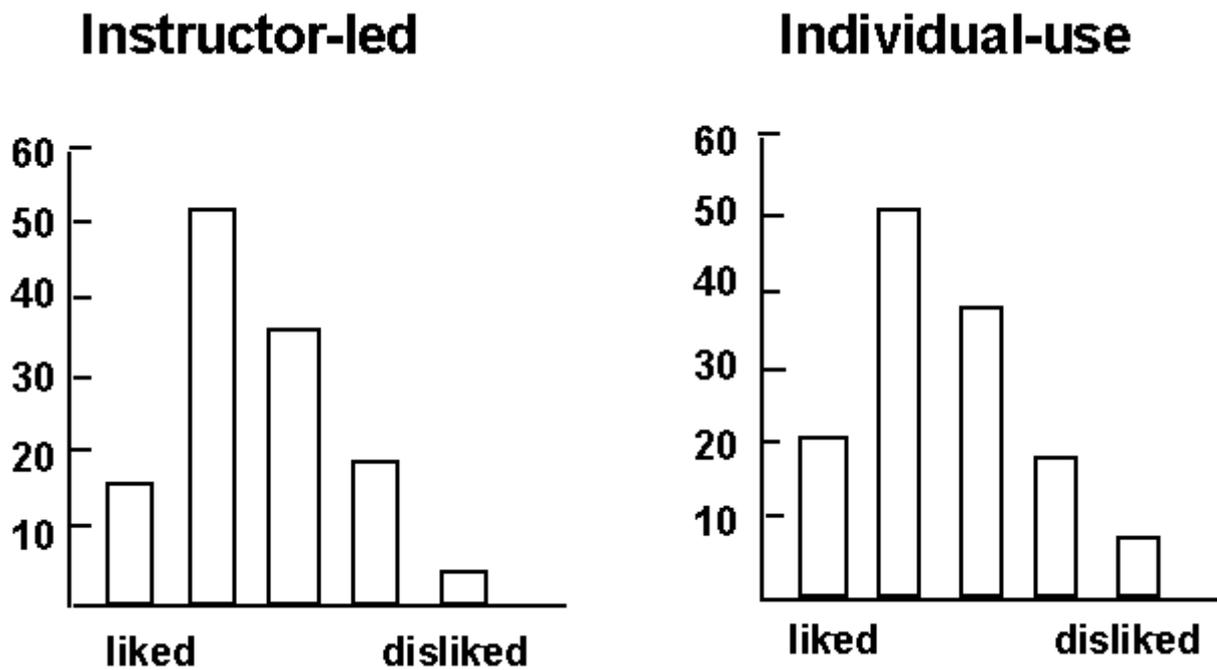


Figure 2.4 Tutor Preferences of the Two Groups, Measured by Poll Responses

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.

<p>Subject B-49: This training without an instructor is close to useless; it should only be used to enhance the class.</p>
<p>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.</p>
<p>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.</p>
<p>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.</p>

Table 2.2 Written Responses to Poll

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.

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 ten 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

Figures 2.5a, 2.5b, 2.5c, 2.5d: 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.

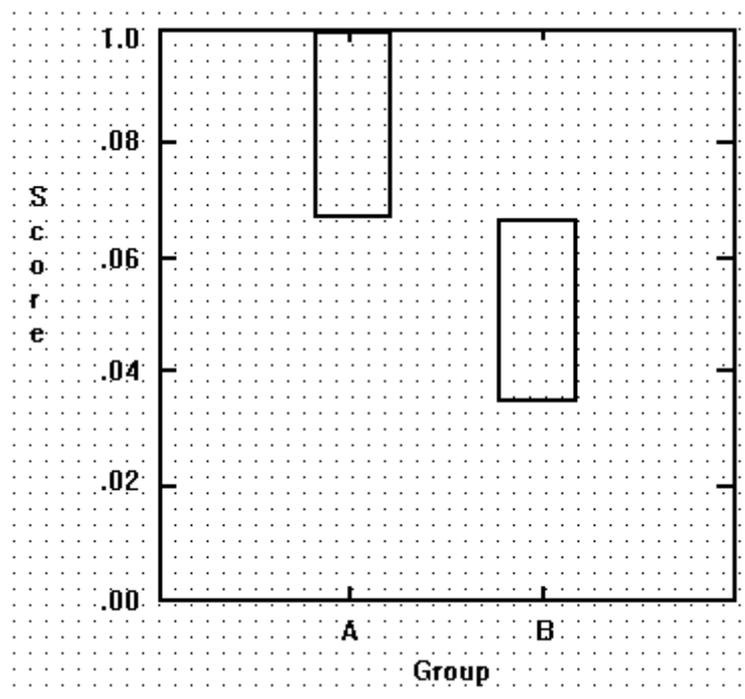


Figure 2.5a Distribution of Scores for Examination Sections A.

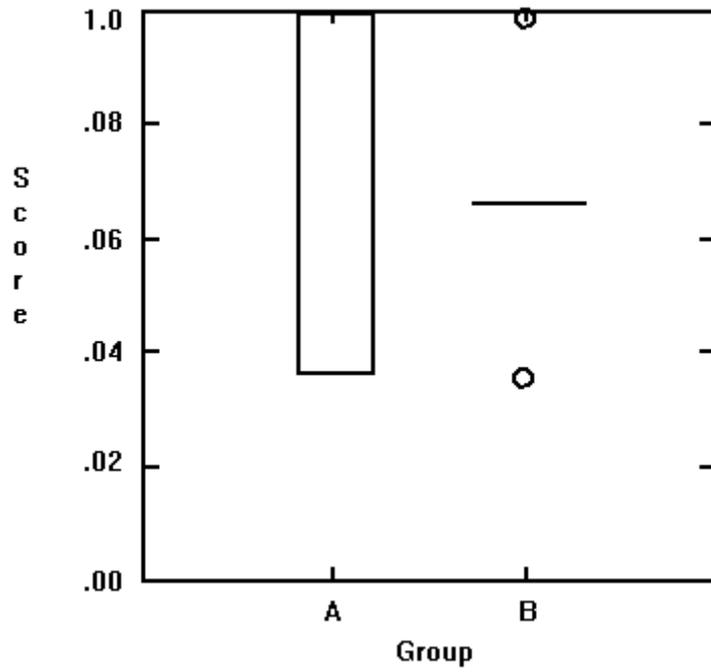


Figure 2.5b Distribution of Scores for Examination Sections B.

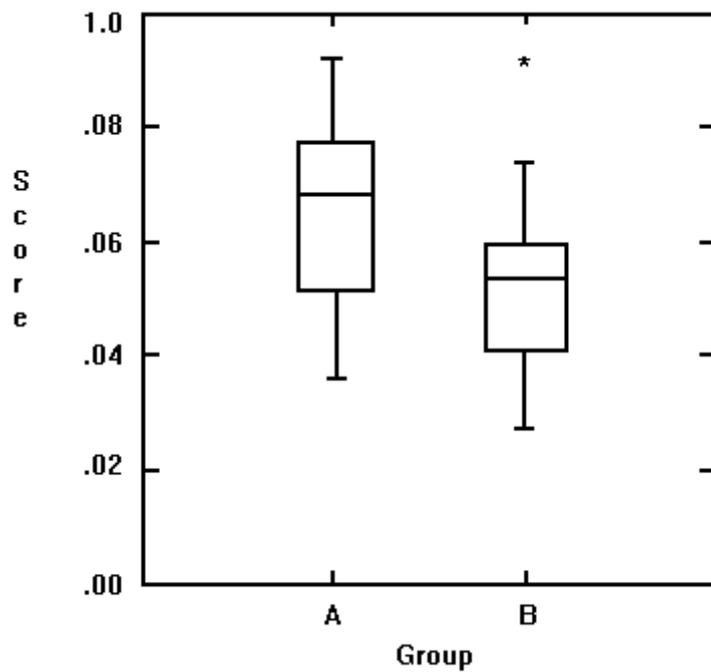


Figure 2.5c Distribution of Scores for Examination Sections C.

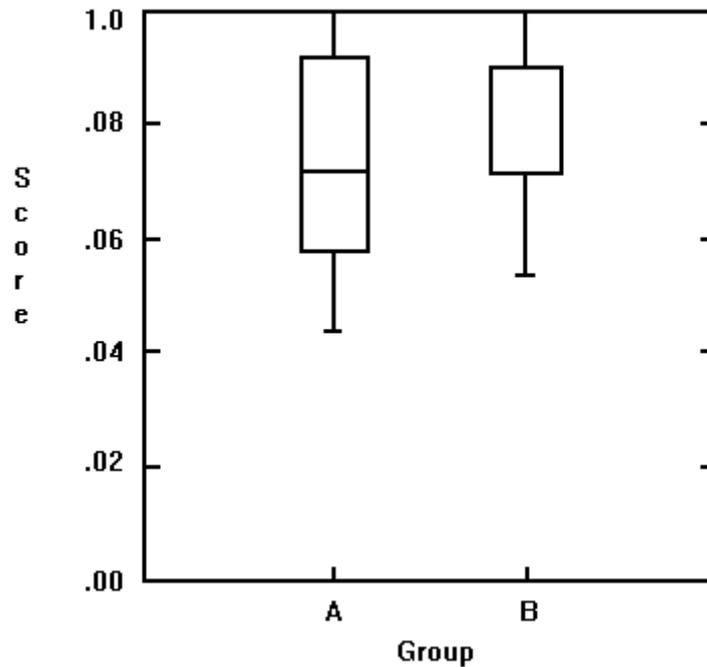


Figure 2.5d Distribution of Scores for Examination Sections D.

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](#)."

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 neutral disagree strongly
 agree 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.

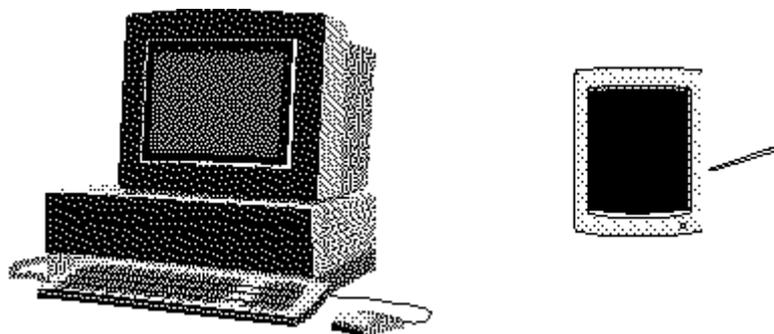


Figure 3.1 Typical Desktop and Pen Computers

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 [CPU](#)s, 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 [CPU](#)s. 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.

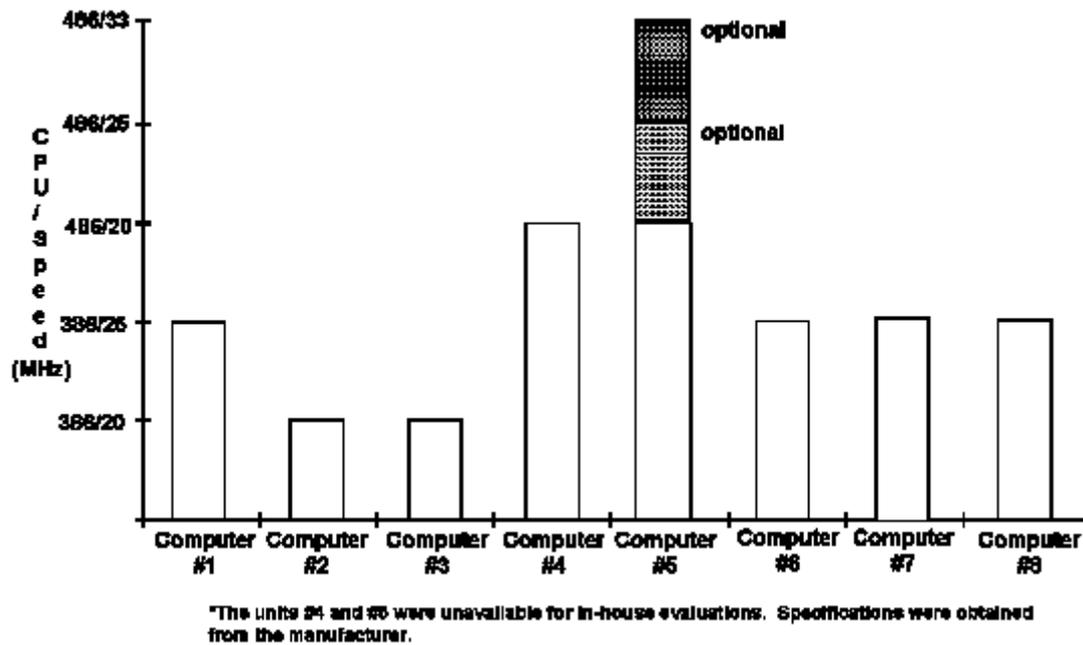


Figure 3.2 Comparison of CPU Type and Speed

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.

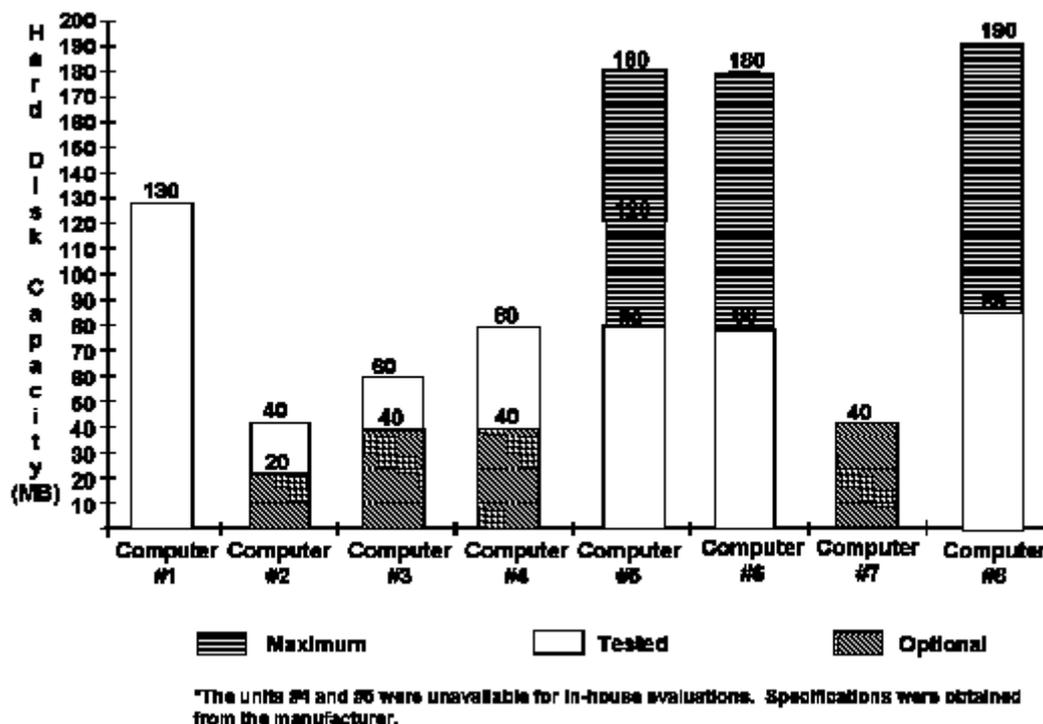


Figure 3.3 Comparison of Hard Disk Capacities

Display Type. 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.

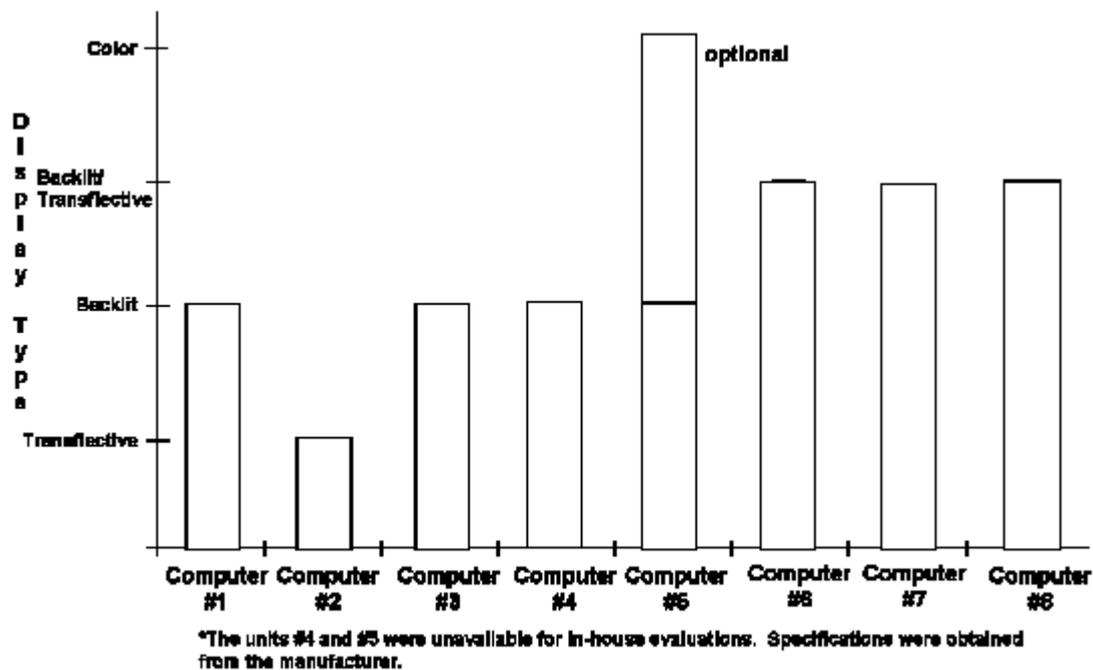


Figure 3.4 Comparison of Display Types

Weight. 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.

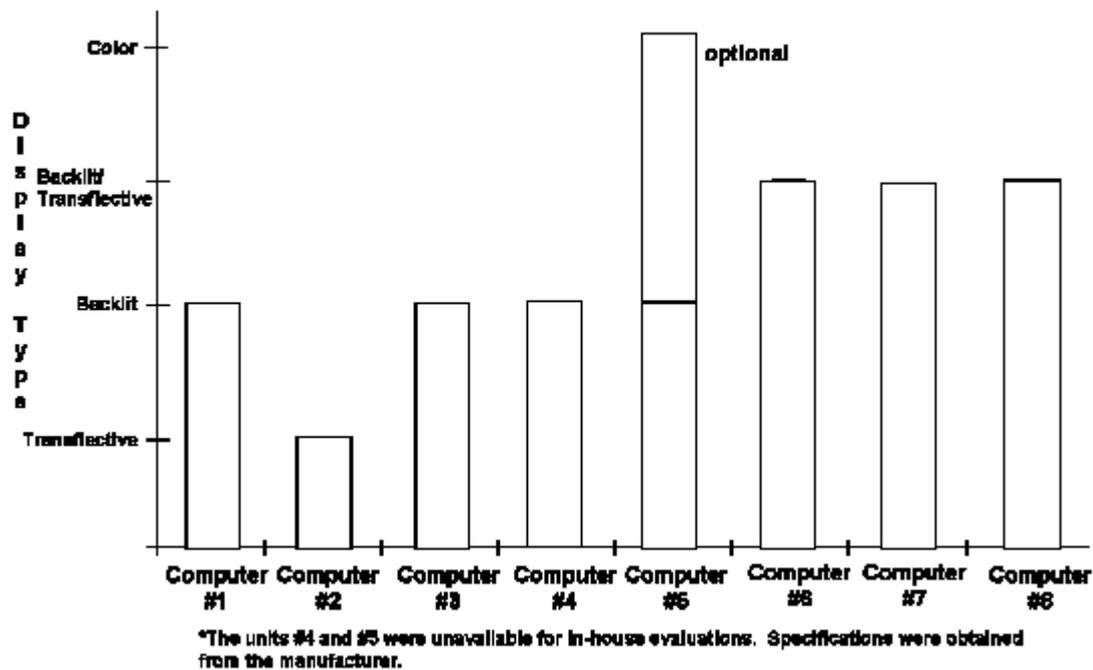


Figure 3.5 Comparison of Weights

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.

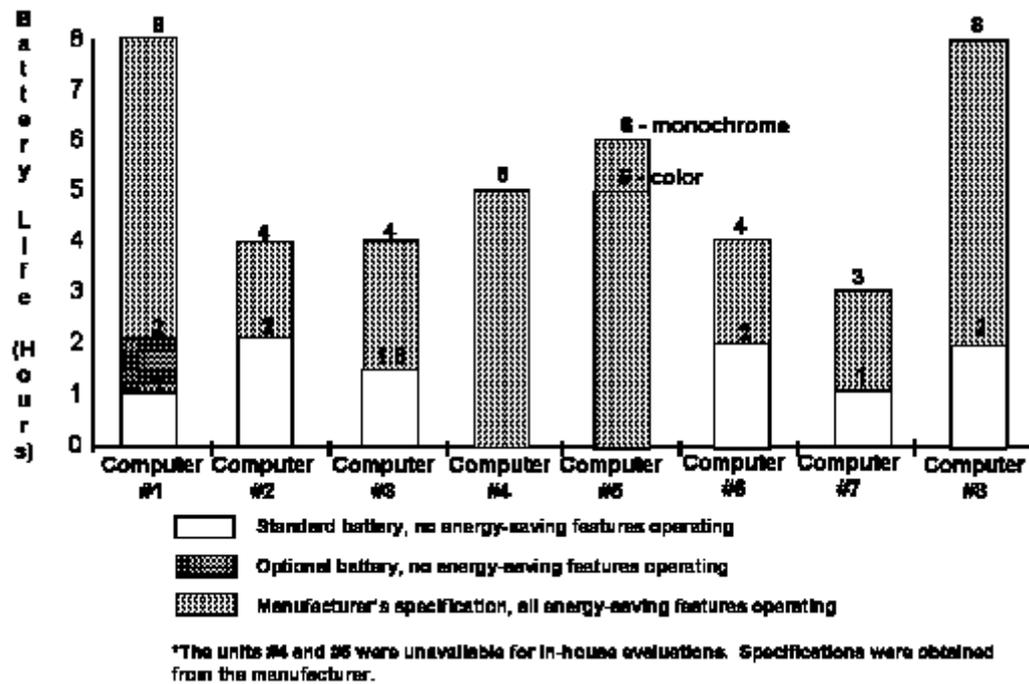


Figure 3.6 Comparison of Battery Life

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.

	Computer #1	Computer #2	Computer #3	Computer #4	Computer #5	Computer #6	Computer #7	Computer #8
Rugged	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

Table 3.1 Miscellaneous Pen Computer Characteristics

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](#).

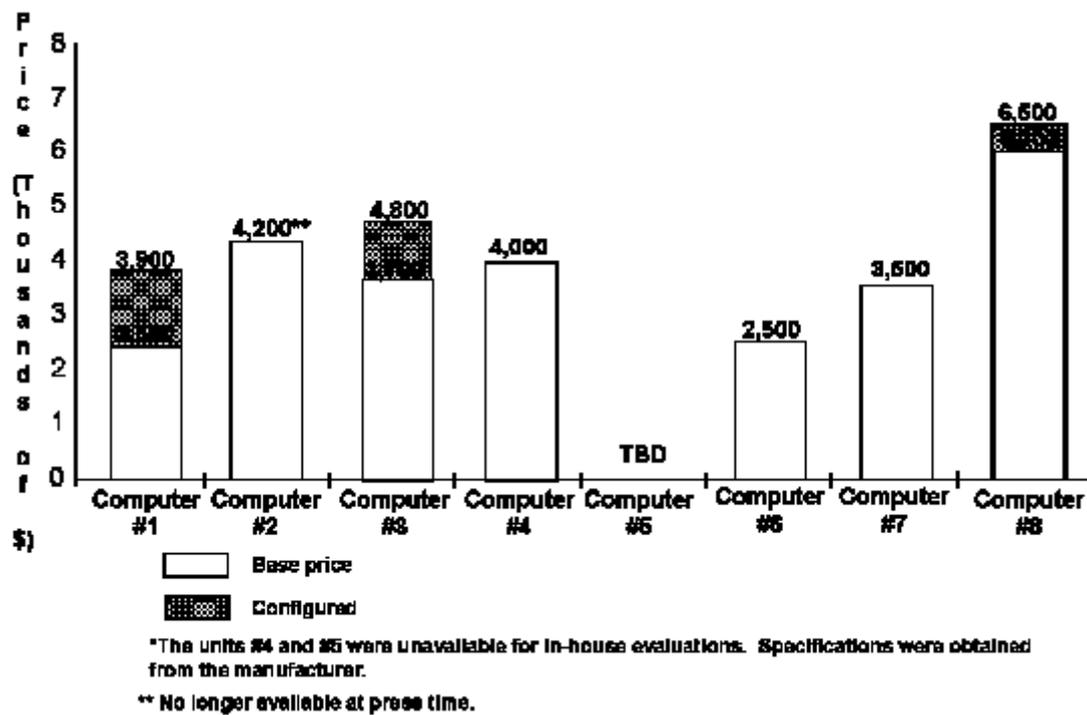


Figure 3.7 Comparison of Prices

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.

3.4.3 Common Limitations of Current Pen Computers

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.

<u>Suggested location:</u>	<u>Reasons:</u>
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.

<u>Suggested Computer:</u>	<u>Reasons:</u>
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 #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 (e.g., 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.

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 (e.g., 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 [CPU](#)s are used in most pen computers. However, 80486 and 80586 [CPU](#)s 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 [CPU](#) s, 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 (e.g., less than 4 lbs.) with fewer software tools (e.g., no word processors, spreadsheets, etc.) available or a heavier product with more tools?

Lightweight, Few tools Heavy, More tools

Would you prefer using a smaller, more lightweight product (e.g., less than 4 lbs.) if software tools available on the pen computer (e.g., 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 (e.g., 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 (e.g., less than 4 lbs.) with an operating environment (e.g., DOS) that was different from the office computer operating environment (e.g., 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 (e.g., pen, mouse, trackball)

Keyboard input device (either affixed, detachable, or "virtual")

Storage device (e.g., 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 (e.g., 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 Panhuse (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 50 percent probability of detection for a 1/2 inch 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.

Negative findings:

- 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 Wechsler 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/Distractibility 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

- Behraves, 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.

Panhuse, 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., Schurman, 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 FAA 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.

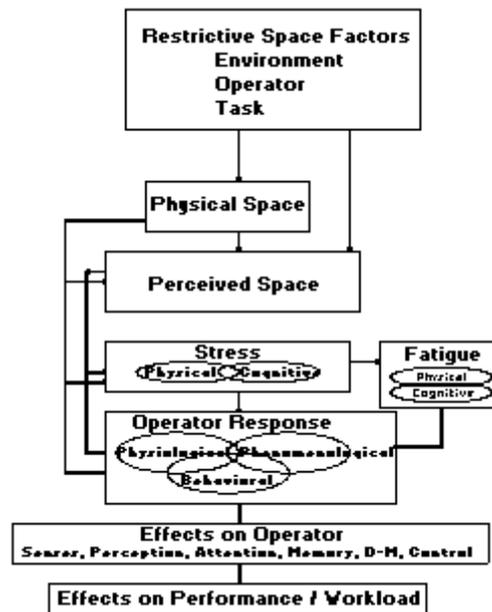


Figure 5.1 Restrictive Space Model

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.

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

Table 5.1 Restrictive Space Factors

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](#)).

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

Table 5.2

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 results 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 5.11.5](#)). 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. Boohar (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 I: 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. and Fatigue in the Inspection Environment

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.

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

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

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).

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.

Table 6.2 Commonly Used Lighting Sources

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.

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

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

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., cherry picker), 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](#).

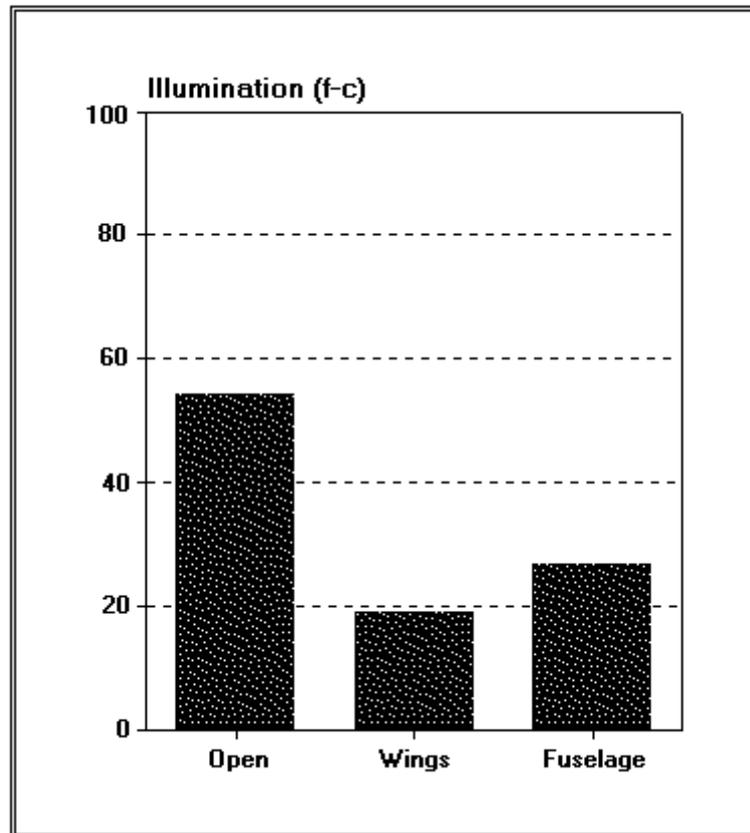


Figure 6.1 Illumination by Aircraft

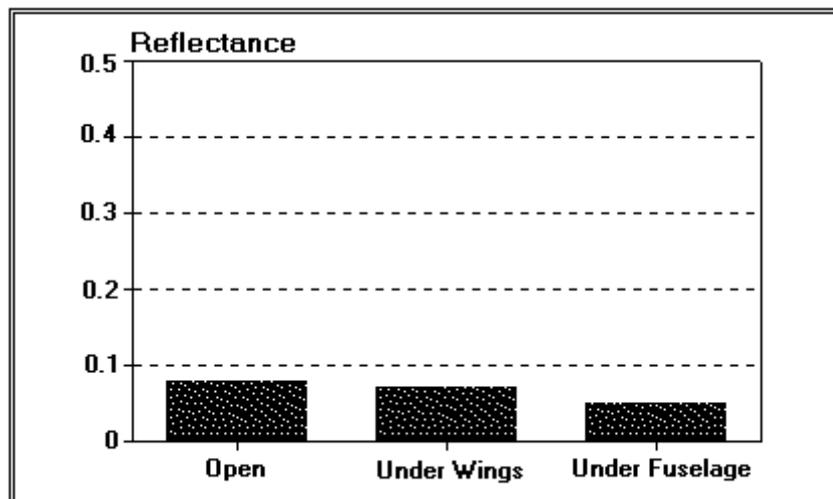


Figure 6.2 Perceived Glare by Aircraft Area

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.

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.

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

Table 6.4 Task Light Environment and Illumination by Task 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.

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 distribution was skewed towards inadequate in the lower exterior and interior areas ([Figure 6.6](#)).

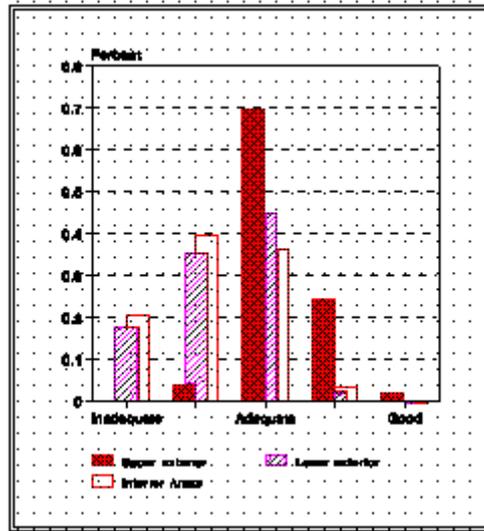


Figure 6.3 Perceived Light Level by Aircraft Area

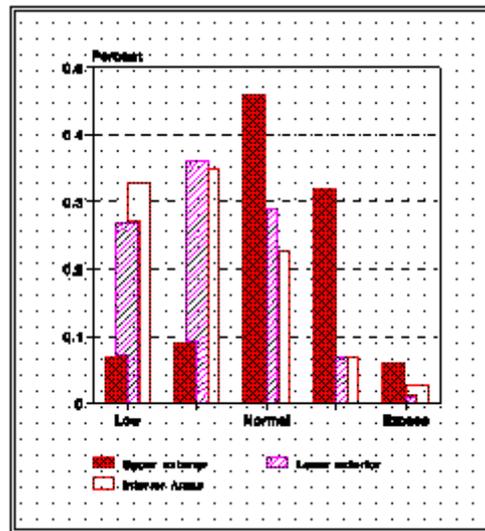


Figure 6.4 Perceived Contrast by Aircraft Area

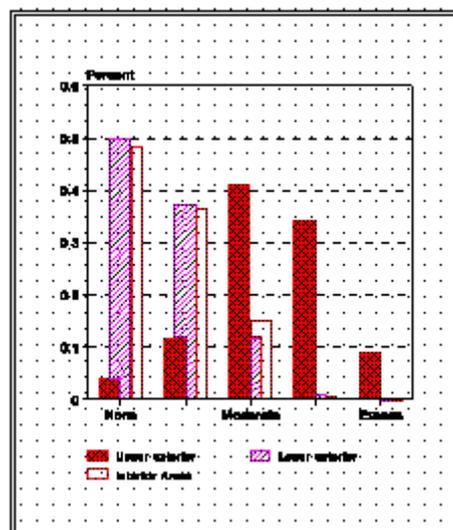


Figure 6.4 Perceived Glare by Aircraft Area

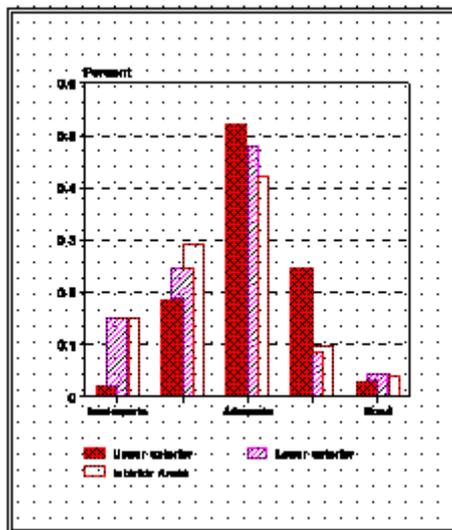


Figure 6.6 Perceived Color Rendering by Aircraft Area

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.

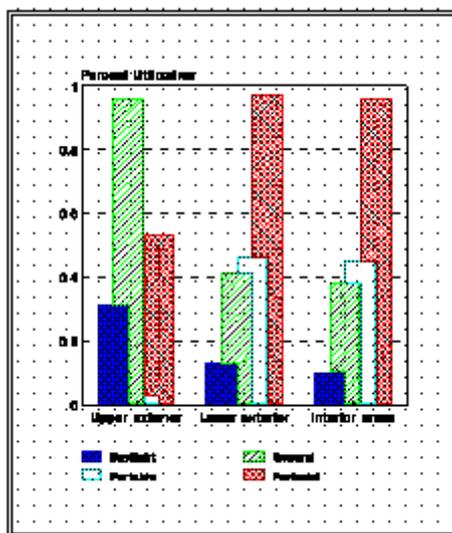


Figure 6.7 Lighting Utilization by Aircraft Area

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.

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. Power source	3. Set-up	3
			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		3
		Total	13
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

Table 6.5 Relevant Lighting Considerations Based Upon Inspector Perceptions

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.

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 be 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.

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

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*	1.3	Yes	Adequate	Battery	None indicated	Headband Battery clamp

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

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.

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 flashlight 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

Table 6.8 Lighting Recommendations for Various Tasks Based Upon the Task Demands, the Lighting Requirements, and the Size of the 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 [Tables 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 unrestricted, 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.

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.

CHARACTERISTICS	PERSONAL	PORTABLE
LIGHT	<ol style="list-style-type: none"> 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 	<ol style="list-style-type: none"> 1. Output/brightness 2. Glare/brightness control 3. Distribution/aim 4. Color rendering 5. Contrast 6. Alternative sources 7. Flicker
EASE OF HANDLING	<ol style="list-style-type: none"> 1. Weight/size 2. Accessories 3. Power source 	<ol style="list-style-type: none"> 1. Weight/size 2. Accessories 3. Set-up
DURABILITY	<ol style="list-style-type: none"> 1. General 2. Safety requirements 3. Bulb life 4. Battery life 	<ol style="list-style-type: none"> 1. General 2. Safety requirements 3. Bulb life
FLEXIBILITY	<ol style="list-style-type: none"> 1. Task demands 2. Fault types 	<ol style="list-style-type: none"> 1. Task demands 2. Fault types
OTHER ATTRIBUTES	<ol style="list-style-type: none"> 1. Cost 2. Space 3. Individual differences 	<ol style="list-style-type: none"> 1. Cost 2. Space 3. Individual differences

Table 6.9 Lighting Source Design Considerations

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.

CHARACTERISTICS	VISUAL ENVIRONMENT
Light	<ol style="list-style-type: none"> 1. Light level 2. Glare 3. Distribution 4. Color rendering 5. Contrast 6. Flicker
Work Shift	<ol style="list-style-type: none"> 1. Light (day/night) 2. Shiftwork
Maintenance	<ol style="list-style-type: none"> 1. Paint 2. Hangar cleanliness
Other Attributes	<ol style="list-style-type: none"> 1. Access devices 2. Availability of lighting sources

Table 6.10 General Visual Environment Design Considerations

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).

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 I: 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:

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

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 [Tables 7.1a](#), [7.1b](#), and [7.1c](#) attempt to summarize the issues brought out in the previous sections in the form of assertive and usable statements.

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: <ul style="list-style-type: none"> • Boundary conventions <ul style="list-style-type: none"> • initial capitalization • final punctuation marks • extra space • question mark at end of question • exclamation mark • Direct speech conventions <ul style="list-style-type: none"> • quotation marks • paragraphing for change of speaker
6.	Use of word conventions: <ul style="list-style-type: none"> • 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: <ul style="list-style-type: none"> • Use a typeface like Helvetica that has no redundant features. • Avoid using a generic dot-matrix typeface.

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

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 7.1b Guidelines for Design of Paper-based Documentation for Aircraft Inspection (cont'd)

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 writing in most positions.
49.	Provide standard writing tools (pens, pencils, etc.) that enable writing in all positions, even against vertical surfaces.

Table 7.1c Guidelines for Design of Paper-based Documentation for Aircraft Inspection (cont'd)

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](#).

POS	JOB DESCRIPTION	*MECH*	*INSP*
03FEB92 07.51 REV DATE: 09-04-90			
WORK CONTROL CARD MFG P/N: J07-3200-21-02-001 PROC, A/C, 3200, B727, ATA32,		M&E P/N: 07-521-1-0017 LANDING GEAR	
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 [Tables 7.2a](#) and [7.2b](#). 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.

1. INFORMATION READABILITY	
A. Typographic Layout	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • system unresponsive of graphics • spatial information conveyed through text, results in the use of complex and lengthy sentences which are difficult to comprehend

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

3. INFORMATION ORGANIZATION	
A. Information Classification	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • no layering of information • not conducive to expert as well as novice usage • difficulty in writing such unstructured information
C. Other organizational issues	<ul style="list-style-type: none"> • no use of naturally occurring page modules for fitting in information • improper sequencing of tasks
4. PHYSICAL HANDLING & ENVIR.	
	<ul style="list-style-type: none"> • 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 7.2 b A-Check Workcard: Issues Identified Within the Taxonomy (cont'd)

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.

Table 7.3 A-Check Workcard Usage: Interpretations of 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 [Tables 7.2a](#) and [7.2b](#) 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.

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

<p><i>2A/B LANDING GEAR (Nose & Main)</i></p> <p>Check systems condition</p> <p>A. Landing gear and wheel wells</p> <p>(1) <i>Check</i>: general condition damage and evidence of fluid leaks</p> <p>(2) <i>Check</i>: - wheel wells - landing gear assemblies - chafing of hydraulic lines for broken aldel clamps on hydraulic lines on the</p> <p><i>Note</i>: Discrepancies on the above item must be corrected prior to the dispatch of the aircraft.</p> <p>(3) <i>Check</i>: - Brake deboost valves for being within limits</p> <p><i>Note</i>: See attachment for PEX effected ACFT deboost valve check</p>	<p><i>Work Control Card: Landing Gear SCIC/AAC</i></p> <p><i>02Feb92 07:51</i> <i>Rev date 090490</i></p> <p><i>No. 02189 MFG PLAN 107</i> <i>3200-21-020-001</i> <i>Proc. A/C: 3200, B727, ATA32</i> <i>M&E PLAN:</i> <i>07-0521-1-0017</i></p>
--	---

Page 5

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

A design was proposed for the workcard holder using the issues of [Tables 7.2a](#) and [7.2b](#) 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](#).



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 [Tables 7.4a](#) and [7.4b](#). 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.

1. INFORMATION READABILITY	
A. Typographic Layout	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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 interchangeably) • 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

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

3. INFORMATION ORGANIZATION	
A. Information Classification	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • no layering of information • not conducive to expert as well as novice usage • difficulty in writing such unstructured information
C. Other organizational issues	<ul style="list-style-type: none"> • 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	
<ul style="list-style-type: none"> • 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 7.4b C-Check Workcard: Issues Identified within the Taxonomy (cont'd)

Q. No	Interpretation
1.	- 48% of the inspectors find the present workcard a useful source of information.
2.	- 33% 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.	- 38% 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.

Table 7.5 C-Check Workcard: Interpretations of Inspector Responses

7.9.4 Workcard for C-check: Proposed Design

The issues highlighted by the inspector responses and those identified in [Tables 7.4a](#) and [7.4b](#) 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.

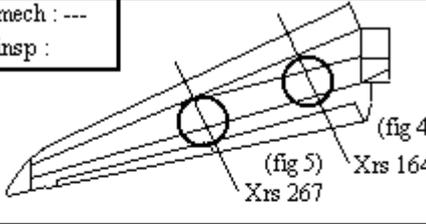
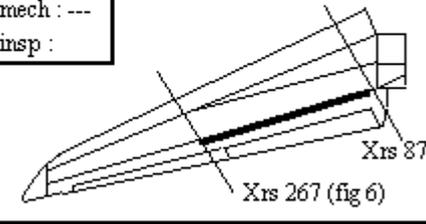
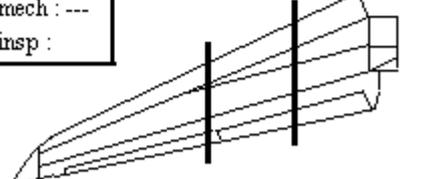
<p>LEFT WING INSPECTION</p> <p>2. Check: - upper cap vertical leg - visible portion of the lower spar cap vertical leg below the bottom edge of the rear spar web. for evidence of cracks Refer: Figures 4 & 5 for known locations of possible cracks (T05-5710-1-0001)</p> <p>3. Check: - lower cap AFT horizontal leg for evidence of cracks Refer: - Figure 6. for known locations of possible cracks (T05-5710-1-0001) Method: use visual aids - boroscope - extension mirror - flash light</p> <p>D. Inspect</p> <p>1. Check: - fences - vortilons for cracks, loose fasteners, security and structural irregularities.</p>	<p>C- Check Work Card - Page 5 of 10</p> <p>mech : --- insp :</p>  <p>mech : --- insp :</p>  <p>mech : --- insp :</p> 
---	--

Figure 7.4a C-Check Workcard: Proposed Design

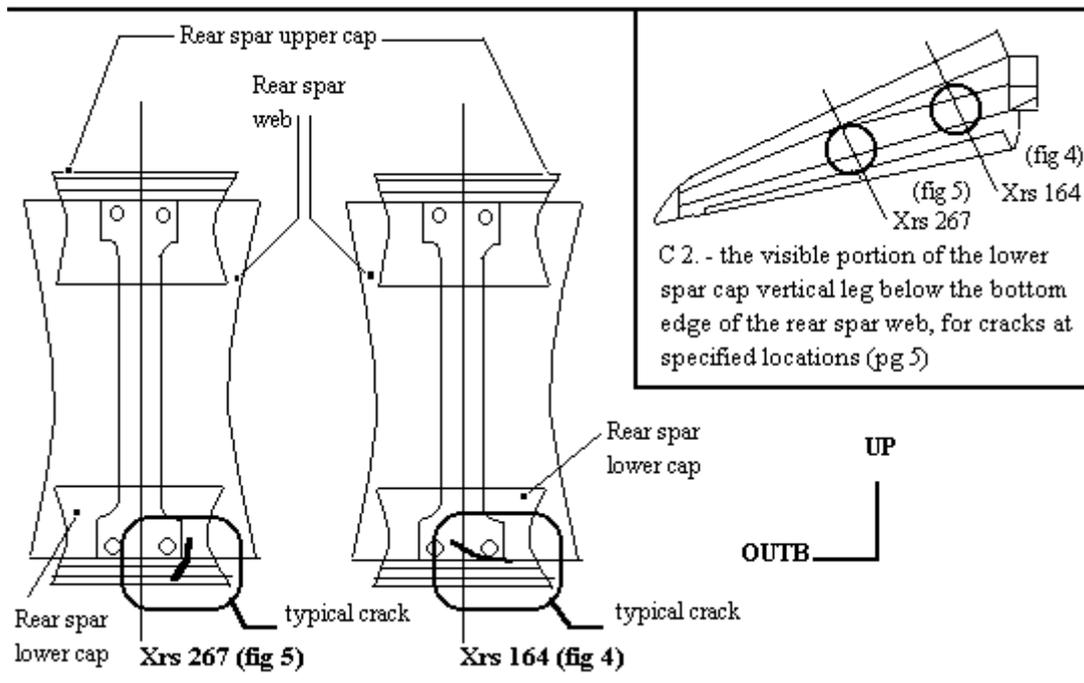


Figure 7.4b C-Check Workcard: Proposed Design - Graphics Attachments

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.

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.

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.

Inspector #	Task 1	Task 2
11	Current W/C - Left Wing	Proposed W/C - Right Wing
12	Current W/C - Left Wing	Proposed W/C - Right Wing
13	Current W/C - Right Wing	Proposed W/C - Left Wing
14	Current W/C - Right Wing	Proposed W/C - Left Wing
15	Proposed W/C - Right Wing	Current W/C - Left Wing
16	Proposed W/C - Right Wing	Current W/C - Left Wing
17	Proposed W/C - Left Wing	Current W/C - Right Wing
18	Proposed W/C - Left Wing	Current W/C - Right Wing

Table 7.6 Experimental Design

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.

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 (1.3)	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 sing 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

Table 7.7 Summary of the Results

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.

7.12 REFERENCES

- Campbell, A. J., Marchetti, F. M. and Mewhort, D. J. K. (1981). Reading speed and text production: A note on night-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:

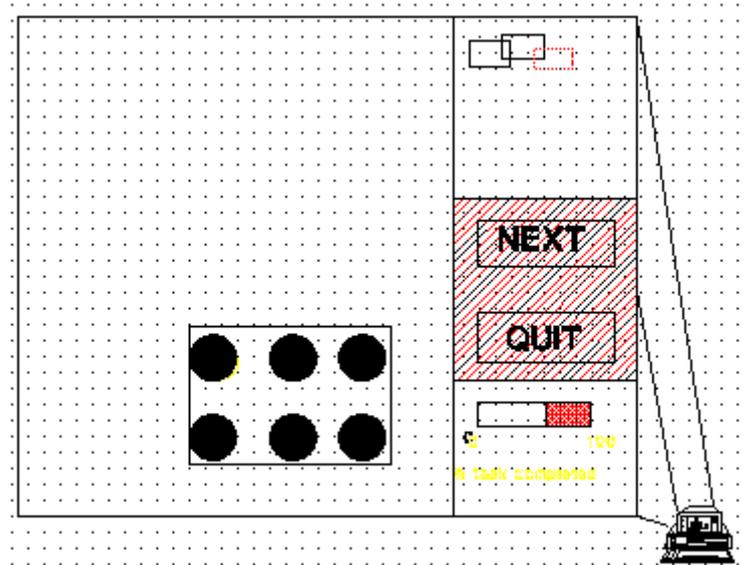


Figure 8.1 Screen of Visual Simulation

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.

Group	Rivet Faults		Area Faults	
	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

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

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 on 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.

	CONTROL GROUP	PROCESS GROUP	VISUAL GROUP	PERFORMANCE GROUP
Number of Fixations	+8.6	-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

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

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.

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.

	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

Table 8.3 Percentage Performance Changes by Group for Feedback Experiment

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.

	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

Table 8.4 Percentage Performance Changes by Group for Feedforward Experiment

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](#).

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

Table 8.5 Percent Correct Decisions for Attribute Perceptions

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.

	PASSIVE TRAINING	ACTIVE TRAINING
Percent Correct, Decision	77.4	87.8
Percent Correct, Inspection	73.5	88.3
Information Transmitted, Decision	0.76	1.06
Information Transmitted, Inspection	0.74	0.99

Table 8.6 Performance by Group, Integration Experiment

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.

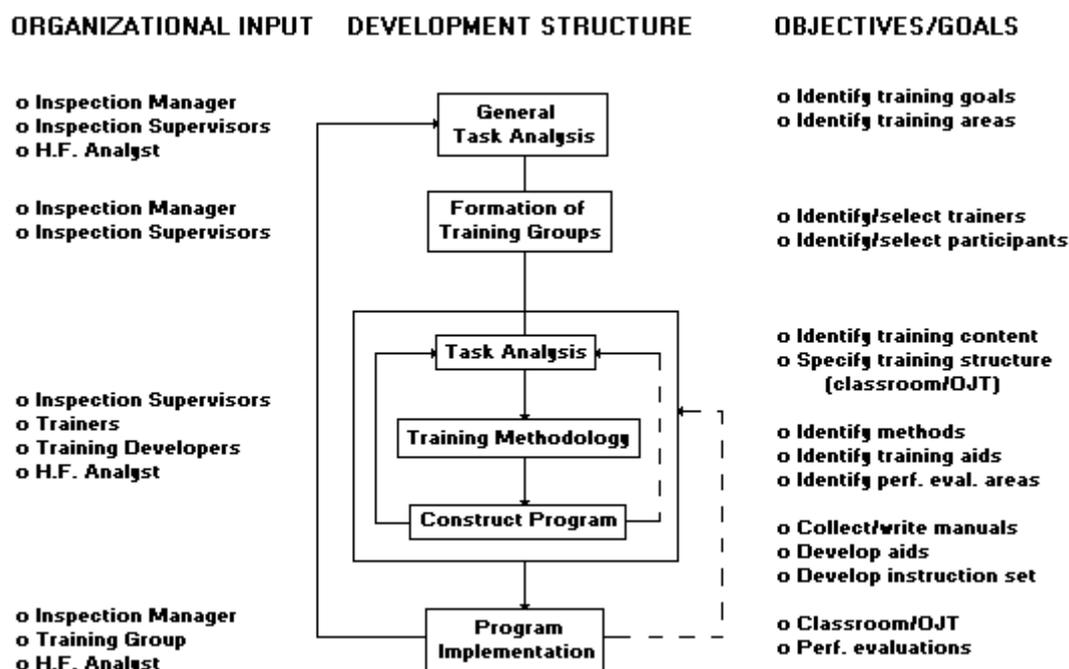


Figure 8.2 Model for Training Program Development in Commercial Aviation

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.

WHEEL WELL, DOORS, ADJACENT COMPONENTS		NOSE GEAR ASSEMBLY & INSTALLATION		NOSE GEAR TIRES & WHEEL ASSEMBLY	
Structure	Defects	Structure	Defects	Structure	Defects
1. Wheel well hydraulic lubing conduits	- Condition - Corrosion - Fluid Leakage	1. NLG shock strut, bracestrut, torque arm, ground sensing mechanism, cables, actuating cylinder, linkages, spring	- corrosion - visual damage - nicks & dings** - fluid leaks - security	1. Wheel hubs valves, 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 NL 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 notify 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				

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

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 collect 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.

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.