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	• Amount of information • Experts versus novices • Use of cues/indications/checklists	• Levels of information • Accessibility • Flexibility of use
2. Layout of information	• Spatial layout • Grouping of information • Chunking	
3. Presentation of Text	• Visual organization • Visual density • Letter case	• Word spacing • Line spacing • Line length
4. Presentation of Graphics	• Visual organization • Spatial location w.r.t. text	• Contrast • Labeling
5. Language Constraints	• Minimal number of words • Standardized nomenclature • Appropriate abbreviations	• Fixed syntax • Concise wording
6. Physical Implementation	• Completeness of information • Consistency across cards	• 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	• 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	• 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	• 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	• 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	• 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	• 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 workcard. • 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., cherrypicker).

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:

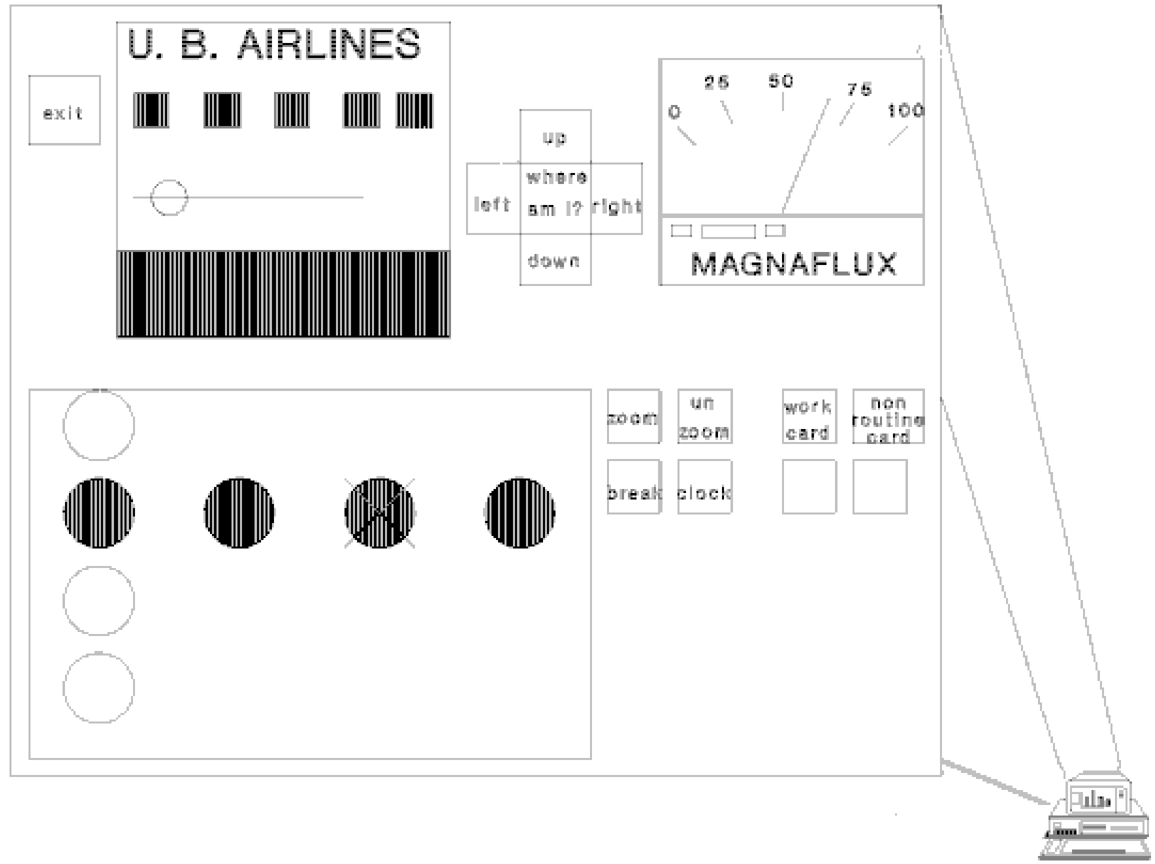


Figure 5.1 NDI Inspection Task Simulation

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:

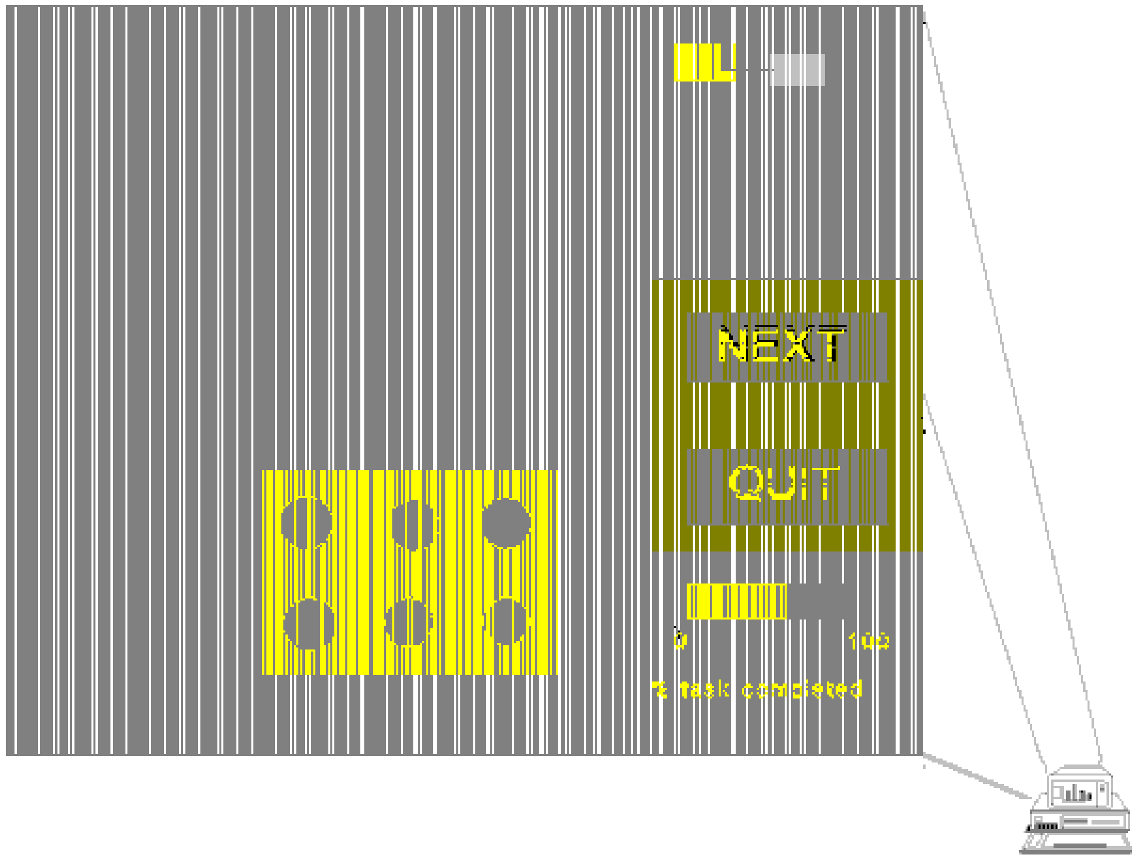


Figure 5.2 Layout of the Simulated Inspection Task

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 ultra-sonic 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:

• *Knowledge-based:*	Search for information and hypothesis testings in novel situations may lead to acts which are judged as errors after the fact;
• *Rule-based:*	The law of least effort may lead to underspecified cues;
• *Skill-based:*	Optimization of motor skill needs feedback from boundaries of acceptable performance (speed-accuracy tradeoff);

INTERFERENCE AMONG COMPETING CONTROL STRUCTURES:

• *Knowledge-based:*	False analogies; interference in means-end hierarchy;
• *Rule-based:*	Functional fixation; adherence to familiar rules;
• *Skill-based:*	Capture by frequently used motor schemata;

LACK OF RESOURCES:

• *Knowledge-based:*	Limitations of linear reasoning in causal networks; insufficient knowledge, time, force, etc.;
• *Rule-based:*	Inadequate memory for rules;
• *Skill-based:*	Lack of speed, precision, force;

STOCHASTIC VARIABILITY:

• *Knowledge-based:*	Slips of memory in mental models;
• *Rule-based:*	Erroneous recall of data or parameters related to rules;
• *Skill-based:*	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	▪ Excessive demands on memory.
Incomplete Knowledge	▪ Knowledge lacking on system dynamics, parameters, side effects, etc.
Delayed Feedback	▪ Delays in feedback concerning the execution of decisions.
Bounded Rationality	▪ Limited cognitive capacity relative to problem size leads to satisficing behavior.
Memory Cueing	▪ Familiar information in LTM is cued by problem content.
Insufficient Consideration of Process	▪ Past history disregarded when handling dynamic current events.
Causal Series Vs. Nets	▪ Oversimplification of causality, thinking in terms of immediate goals.
Attentional Limitations	▪ Finite resources of attentional processes.
Confirmation Bias	▪ Tendency to maintain current hypothesis facing contradictory evidence.
Vagabonding	▪ Moving amongst issues without detailed consideration of any one.
Overconfidence	▪ Excessive belief in the correctness of one's own knowledge.
Memory Slip	▪ Memory failure due to faulty activation of schemata.
Selectivity	▪ Selectively process information. Attention to wrong features of task can result.
Biased Reviewing	▪ Error in reviewing planned course of action.
Illusory Correlation	▪ Failure to detect correlation or understand the logic of covariation.
Lack of Resources	▪ Lack of mental capacity for causal reasoning, insufficient knowledge.
Complexity Problems	▪ Problems in understanding system due to its complex nature.

RULE-BASED ERROR SHAPING FACTORS

RULE-BASED ERROR SHAPING FACTORS	
Availability	▪ Tendency to use intuitive rules or use rules that readily come to mind.
Countersigns	▪ These are input indications that the more general rule is inapplicable.
Rigidity	▪ Mindset (cognitive conservatism) results in refusal to change familiar procedure.
Encoding Deficiency	▪ Encode inaccurately or fail to encode the properties of the problem space.
Inadvisable Rules	▪ Rules that satisfy immediate goals but can cause errors due to side effects.
Wrong Rules	▪ Rules that are wrong for the current situation.
Inelegant Rules	▪ Rules that achieve the goal but are inefficient.
First Exceptions	▪ Errors caused on first occasion that is an exception to the general rule.
SKILL-BASED ERROR SHAPING FACTORS	
Omissions	▪ Omission of actions/action sequences needed to achieve a specified goal.
Perceptual Confusion	▪ A familiar match is accepted instead of the correct match.
SATO	▪ Speed accuracy tradeoff.
Motor Schemata Capture	▪ Focussed attention absence leads to takeover by frequently used schemata.
Stochastic Variability	▪ Variability in control of movements.
Reduced Intentionality	▪ Delay between intention and action.
Repetitions	▪ Actions that are unnecessarily repeated.
Reversals	▪ Unintentionally reversing an action just committed.
Interference	▪ 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.

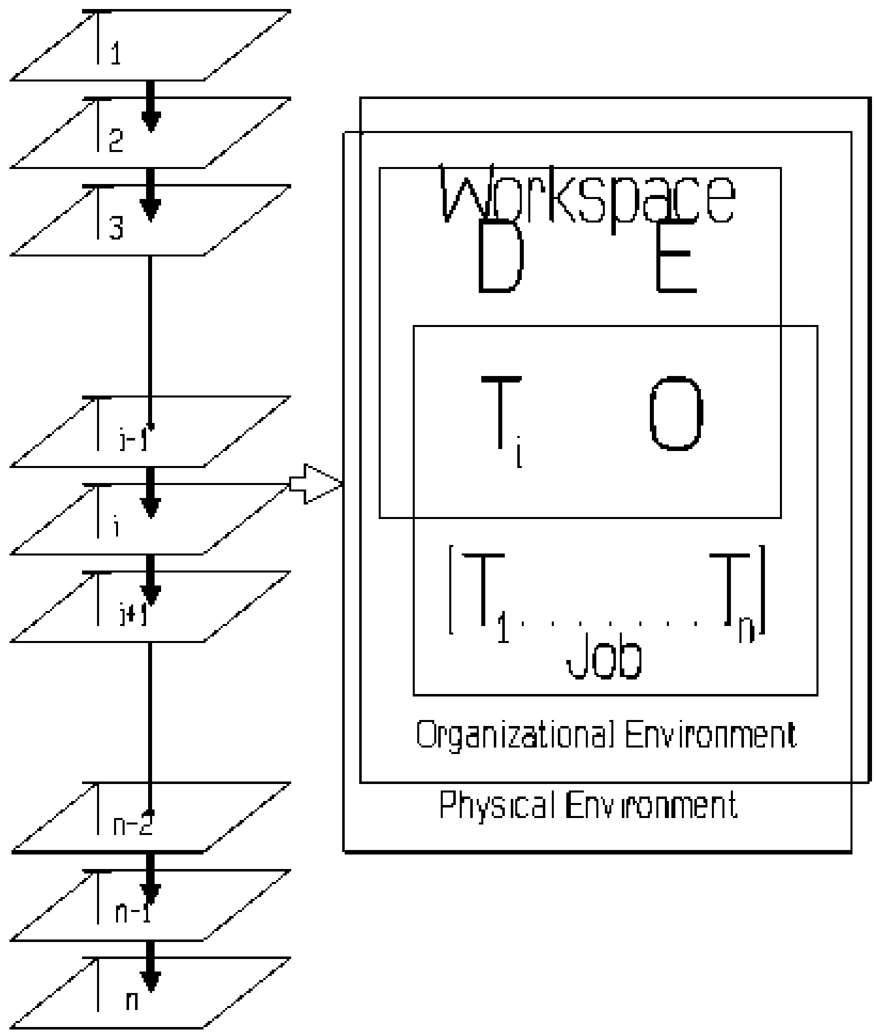


Figure 5.3 System Mode

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	Correct instructions written.
	1.2	Correct equipment procured.
	1.3	Inspector gets instructions.
	1.4	Inspector reads instructions.
	1.5	Inspector understands instructions.
	1.6	Correct equipment available.
	1.7	Inspector gets equipment.
	1.8	Inspector checks/calibrates equipment.
Task 2 - ACCESS	2.1	Locate area to inspect.
	2.2	Area to inspect.
	2.3	Access area to inspect.
Task 3 - SEARCH	3.1	Move to next lobe.
	3.2	Enhance lobe (e.g. illuminate, magnify for vision, use dye penetrant,

Task 3 - SEARCH	3.1	Move to next lobe.
	3.2	Enhance lobe (e.g. illuminate, magnify for vision, use dye penetrant, tap for auditory inspection).
	3.3	Examine lobe.
	3.4	Sense indication in lobe.
	3.5	Match indication against list.
	3.6	Remember matched indication.
	3.7	Remember lobe location.
	3.8	Remember access area location.
	3.9	Move to next access area.
Task 4 - DECISION	4.1	Interpret indication.
	4.2	Access comparison standard.
	4.3	Access measuring equipment.
	4.4	Decide on if it is a fault.
	4.5	Decide on action.
	4.6	Remember decision/action.
Task 5 - RESPOND	5.1	Mark fault on aircraft.
	5.2	Record fault.
	5.3	Write repair action.
Task 6 - REPAIR	6.1	Repair fault.
Task 7 - BUY-BACK	7.1	Initiate.
	7.2	Access.
	7.3	Search.
	7.4	Decision.
	7.5	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,	7.2 Noise
Appropriateness, Veridical)	7.3 Temperature/Ventilation
3.5 Legibility (Physical)	7.4 Chemical Hazards
4.0 TASK	7.5 Vibration
4.1 Physical Requirements	7.6 Electrical Shock Hazards
4.2 Informational Requirements	8.0 WORKSPACE
4.3 Characteristics	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.

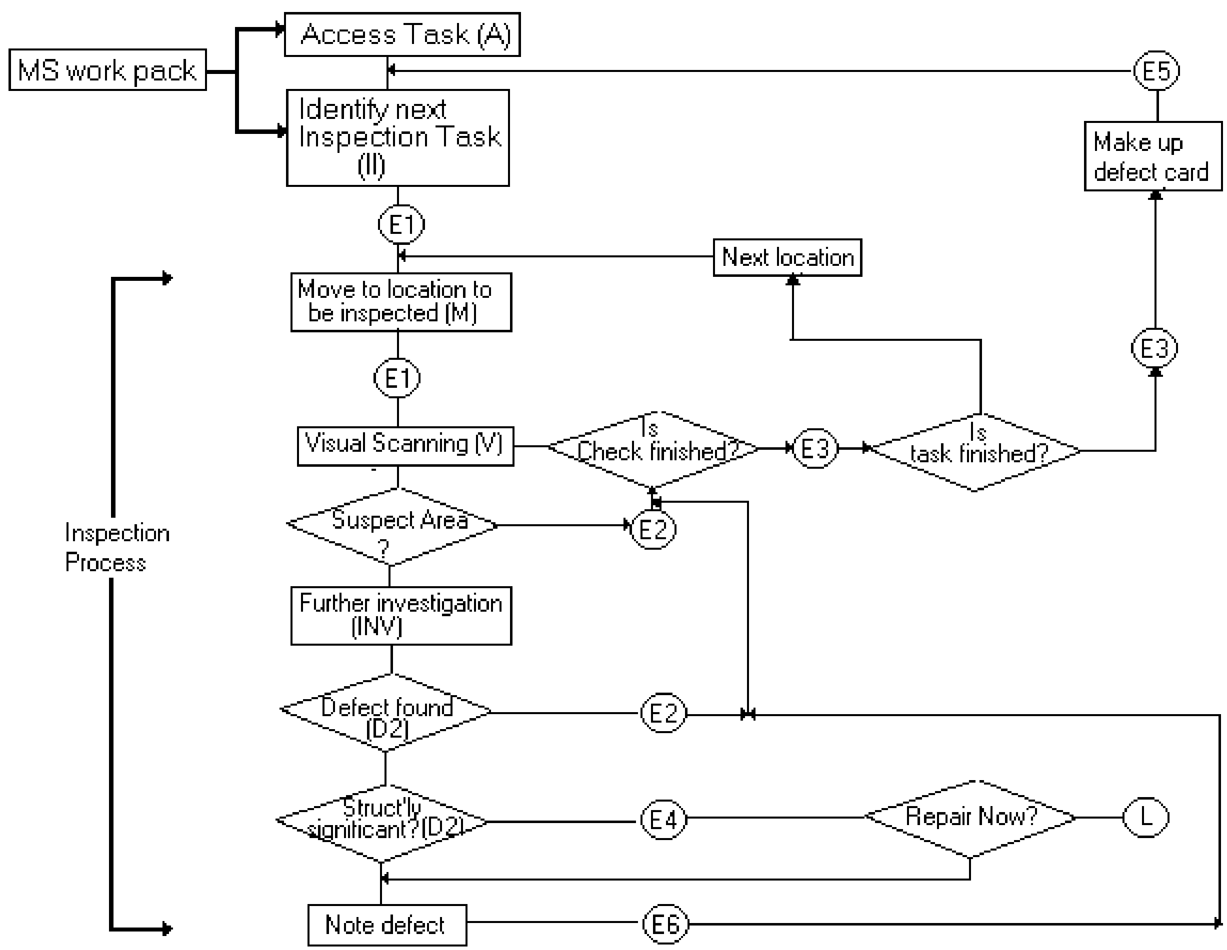


Figure 5.4 Inspection Model Flowchart (Lock and Strutt, 1985)

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.

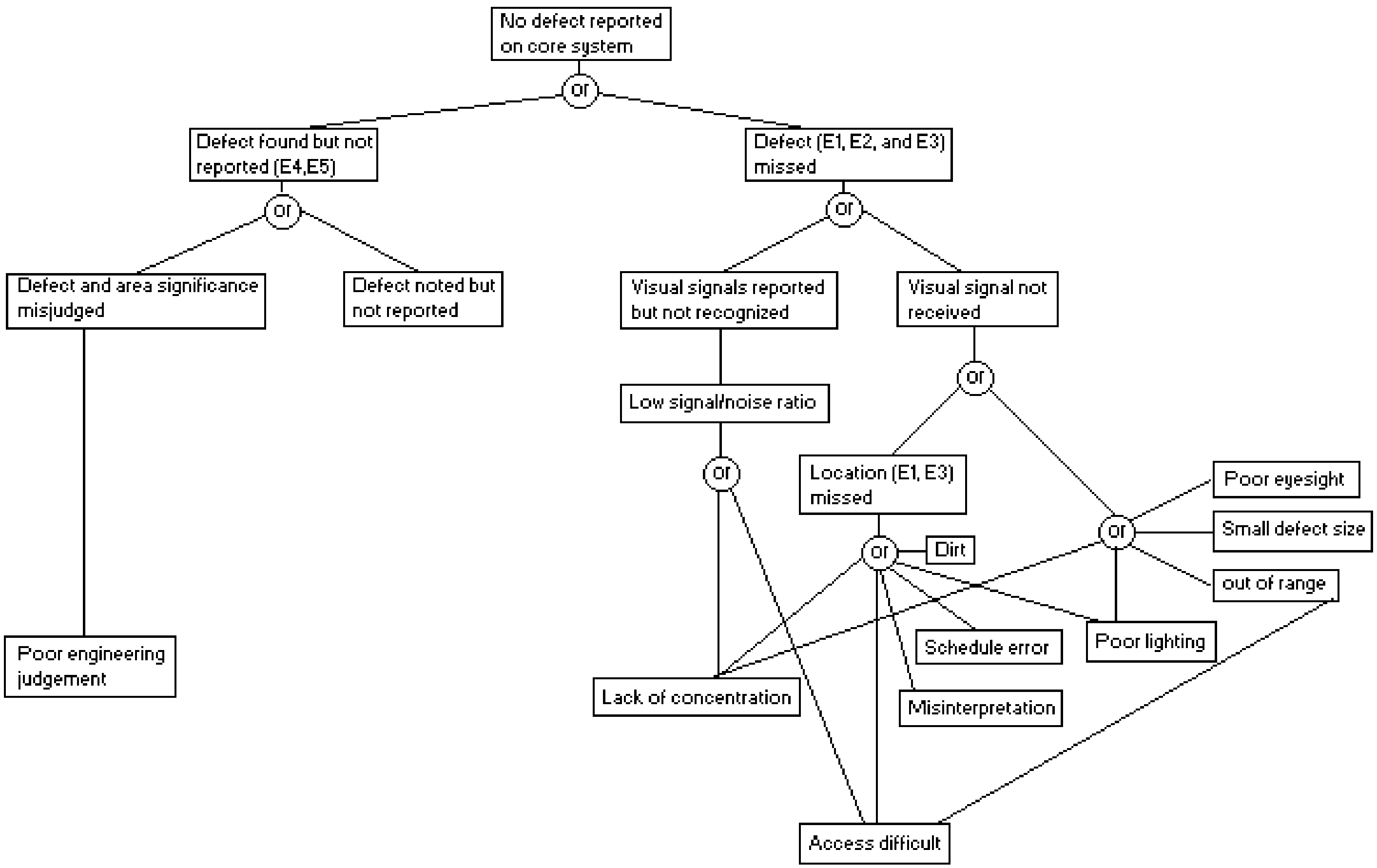


Figure 5.5 Inspection Error Fault Tree (Lock and Strutt, 1985)

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	KNOWLEDGE
Learning and Adaptation	6, 11, 12, 20, 22, 24, 25	1, 2, 11, 20, 22, 24, 28	1, 13, 16, 17, 18, 19, 20, 22, 24, 26, 27, 29
Interference Among Competing Control Structures	12, 14, 23, 24, 30	3, 11, 20, 22, 23, 24, 31	1, 3, 15, 16, 17, 19, 20, 23, 24, 32
Lack of Resources	33	33, 34	1, 3, 4, 13, 16, 17, 19, 20, 21, 22, 23, 24, 33, 34
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, 20, 21, 22, 35

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

		SYSTEM ELEMENTS							
Systemic Errors	Levels of Cognitive Control	Task	Personnel	Job	Workspace	Equipment	Documentation	Physical Environ.	Organiza Environ.
LEARNING	Skill	6,11,12	20		6	6,22,24,25	6,24	25	
	Rule	11,28	20			11,22,24,28	1,2,24,28		
	Knowledge	13,19, 26	16,17,18,20	16,17, 26,29		13,19,22, 24,26,27,29	1,16,19,24		17,18,29
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	1,16,19,24,32		15,17,32
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	1,16,19,24,33,34		17
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	8	
	Rule	11	16,17,20, 21,35			11,22,24,35	2,24,35		
	Knowledge		16,17,20, 21,35	16,17		4,22,35	16,35		17

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.

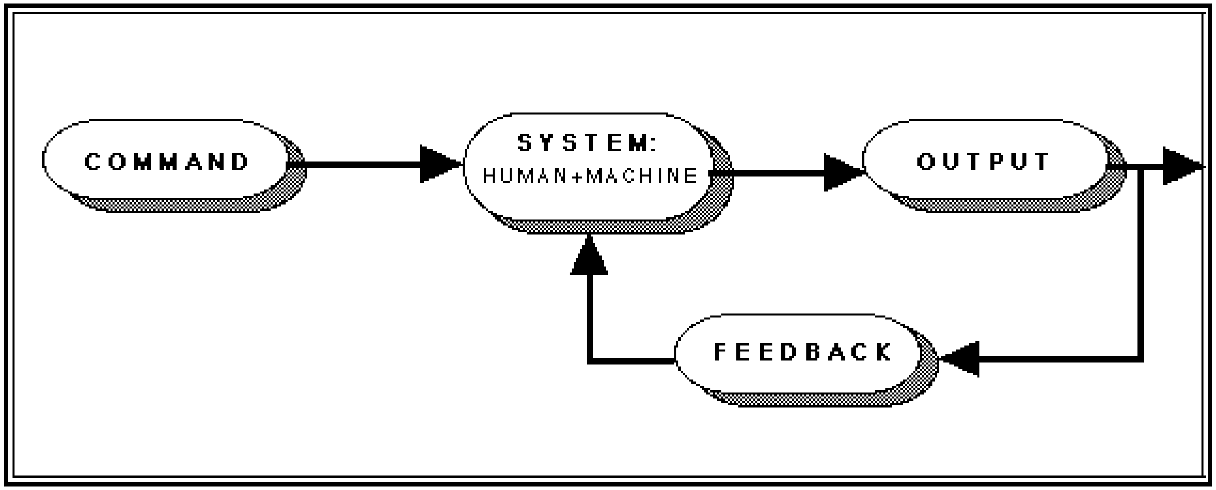


Figure 5.6 Closed-Loop Control

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.

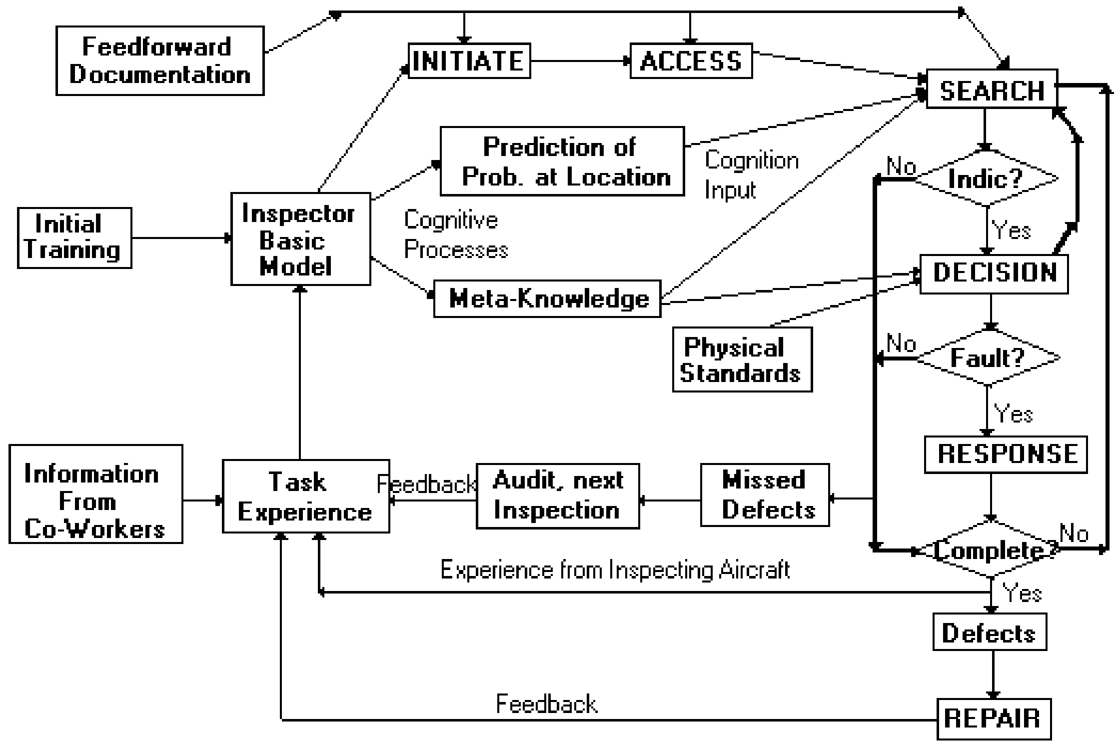


Figure 5.7 Model of Information Flow in Aircraft Maintenance and Inspection (Drury and Prabhu, 1991)

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	▪ Missing rivet ▪ Hole in skin		
FOVEAL (PURE SEARCH)	▪ Missing rivet ▪ Hole in skin ▪ Deep dents ▪ Large cracks ▪ Prominent corrosion		
FOVEAL DECISION		▪ Borderline corrosion ▪ Slight wear ▪ Dished rivets ▪ Ripples in skin ▪ Small cracks	
EXTRA-FOVEAL SEARCH	▪ Chipped paint in periphery leads to next fixation		
DECISION-MAKING (OUTSIDE OF SEARCH)		▪ Streaks around rivets trigger inspection for loose rivets ▪ Powdery contamination triggers search for corrosion ▪ Borderline defects	▪ 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 conscioususe 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 5.20shows 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 for the Decision Making Component of Aircraft Inspection

BEHAVIOR TYPE	FAILURE MODE (PHENOTYPES)	MECHANISMS OF HUMAN MALFUNCTION AND ERROR SHAPING FACTORS (GENOTYPES)	INFORMATION REQUIREMENTS
KNOWLEDGE-BASED	4.1.1 Classify as wrong fault. 4.4.1 False alarm. 4.4.2 Missed fault.	o Selectivity o Incomplete mental model o Lack of knowledge o Confirmation bias	Training in fault causation mechanisms
	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.	o Overconfidence o Incomplete knowledge o Biased reviewing	Crew resource management training Outcome and cognitive feedback
	4.2.5 Measuring equipment wrongly calibrated.	o Overconfidence o Incomplete knowledge	Training and aiding for understanding NDI system physics
RULE-BASED	4.2.4 Fail to calibrate measuring equipment.	o Recall error o Memory slip	Checklists Feedback in NDI instrument
	4.2.6 Fail to use measuring equipment. 4.3.5 Fail to use comparison standards.	o Recall error o Overconfidence o Memory slip o Mindset or rigidity	Checklist
	4.6.2 Fail to record decision.	o Recall error, memory slip o Availability o Mindset o Wrong rules	More field-usable fault-recording devices
	4.2.1 Choose wrong measuring equipment. 4.3.1 Choose wrong comparison standards. 4.5.1 Choose wrong action.	o Stereotype takeover o Memory slip o First exceptions o Availability	Checklist Computer entry and checking of equipment and standard ID's for feedback
	4.2.5 Measuring equipment wrongly calibrated.	o Familiar shortcut o Misinterpretation	Checklist, enforcement of rules
	4.1.1 Classify as wrong fault. 4.4.1 False alarm. 4.4.2 Missed Fault.	o Encoding deficiency o Familiar pattern not recognized	Training in fault classification Cognitive feedback
SKILL-BASED	4.2.1, 4.2.4, 4.3.1, 4.3.5, 4.2.6, 4.5.1, 4.6.2	o Omissions o Recency and frequency of use o Speed accuracy tradeoff	Skills training Skills maintenance for infrequently-used behaviors

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 (Xc). 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 5.21summarizes 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 Xc. 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_c)	---	---	---	---	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

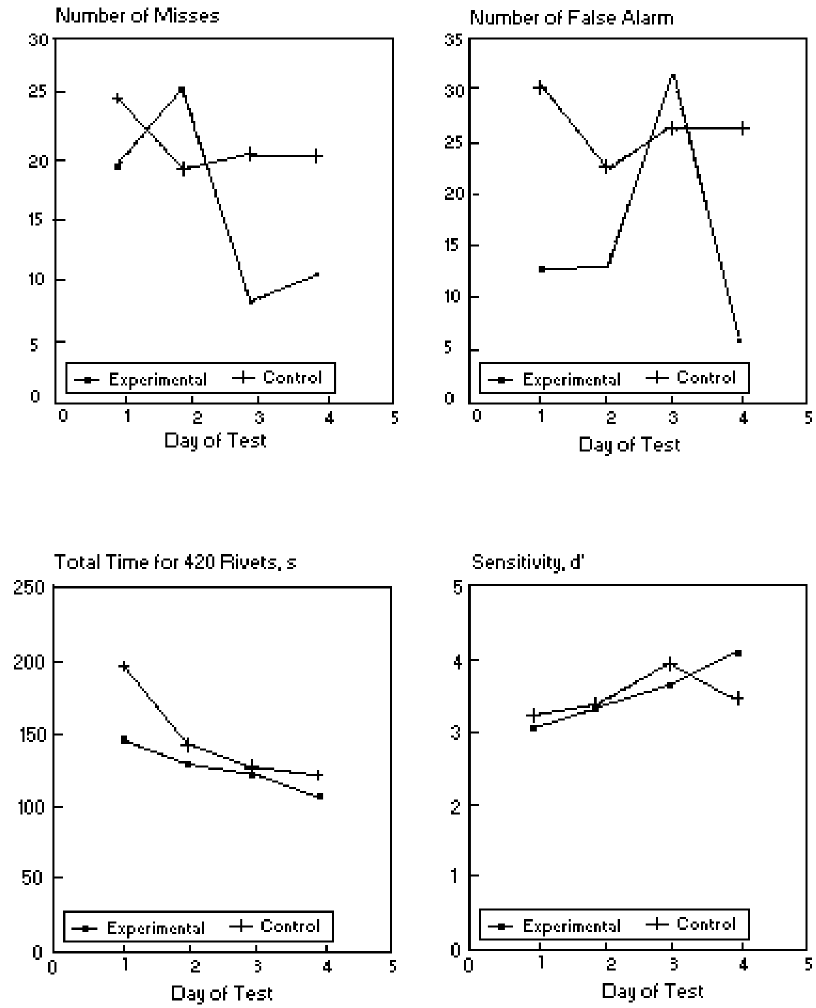


Figure 5.8 Day X Group Interactions for Off-Line Feedback Experiment

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.

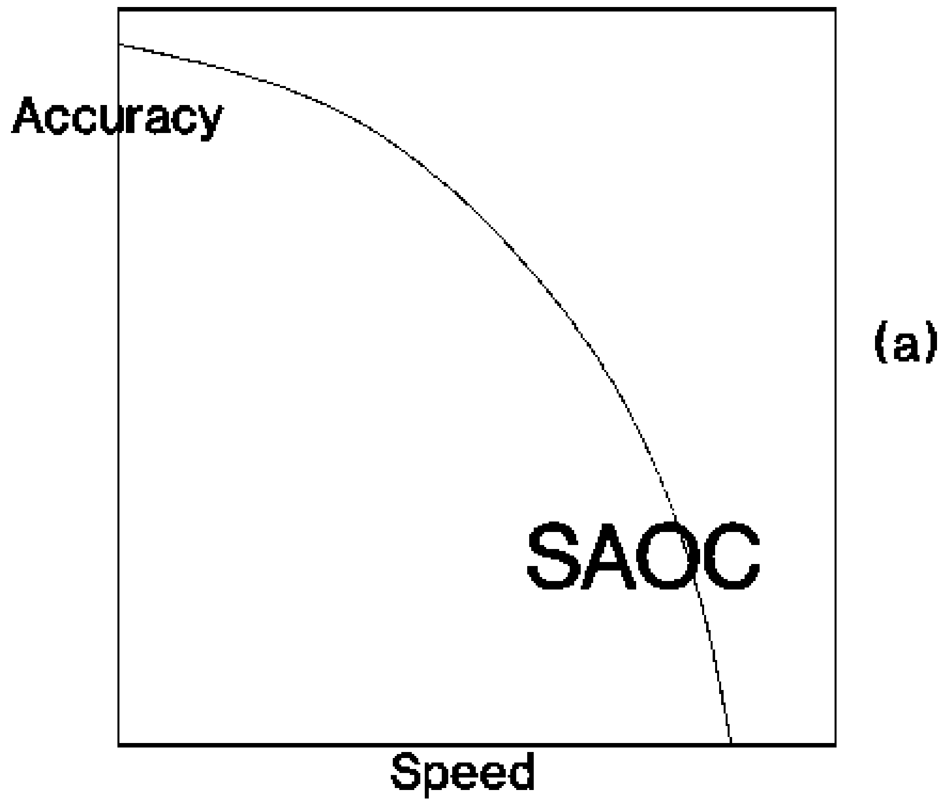


Figure 5.9a Generalized Speed/Accuracy Operating Characteristics (SAOC)

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).

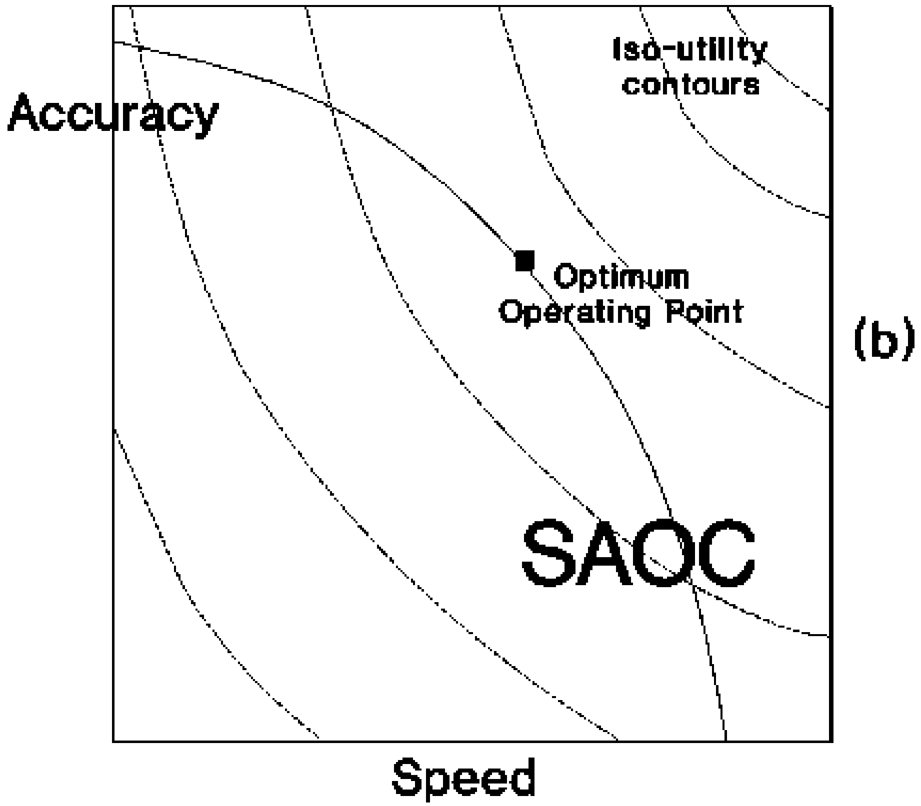


Figure 5.9b Generalized Speed/Accuracy Operating Characteristics (SAOC)

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.

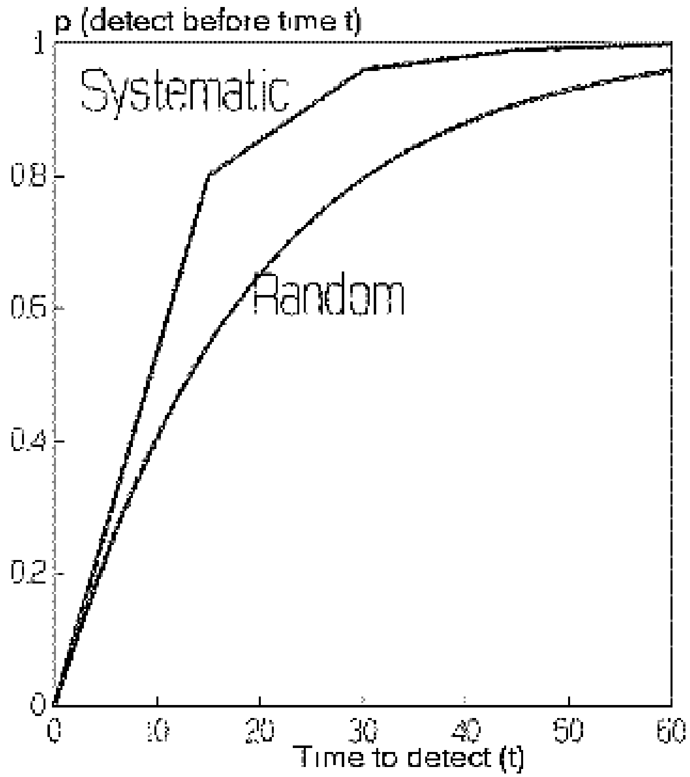


Figure 5.10 Typical Cumulative Search Time Distribution. Giving the SAOC for Visual Search

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 ***(Xc)*** 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 ***T1, T2***, 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 factional design with six subjects nested within each group. The following performance measures were collected:

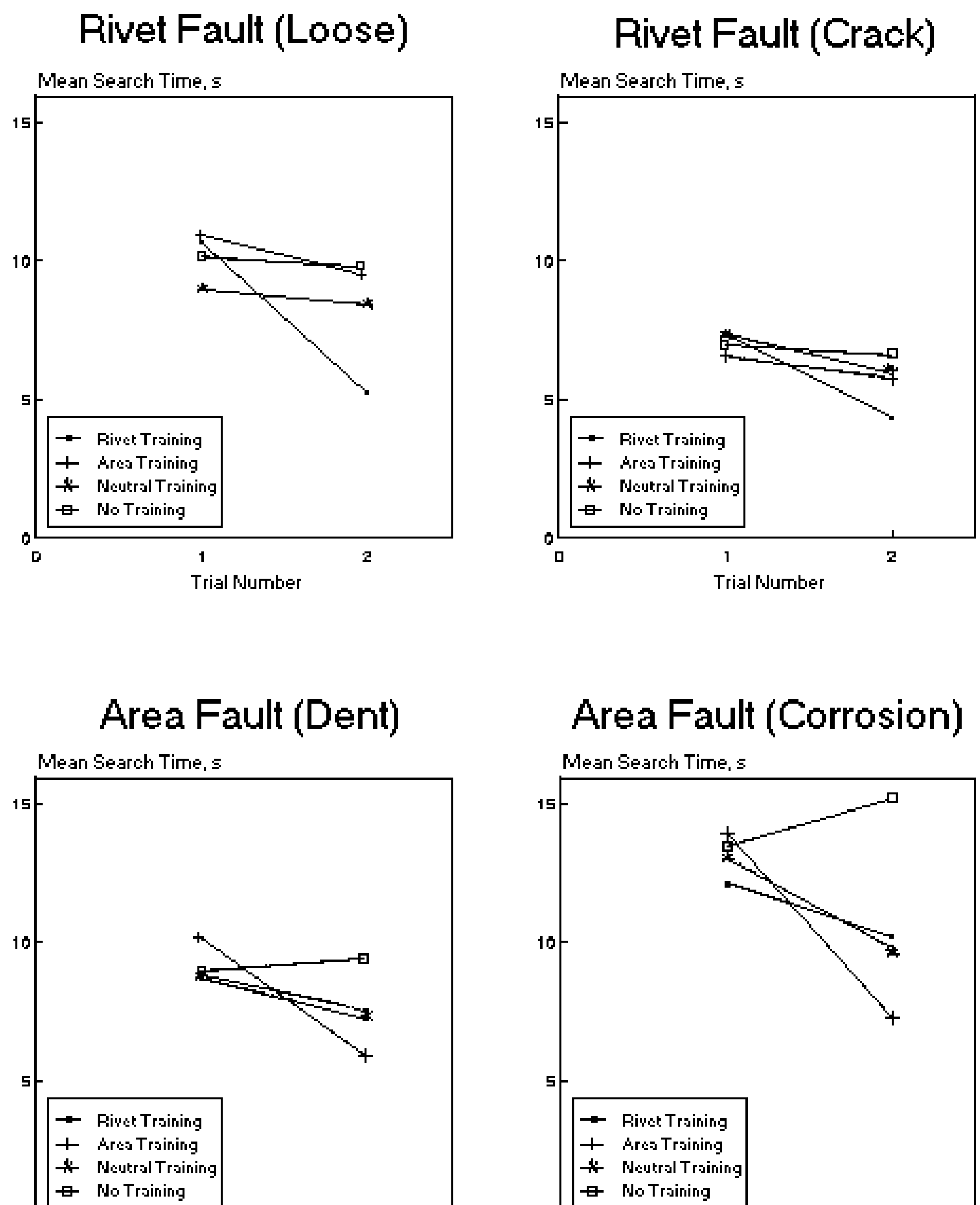
1. Number of correct responses for each of the six fault positions in the visual lobe task.
2. Time to detect a fault in each screen for the visual search task.

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 analyzes 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



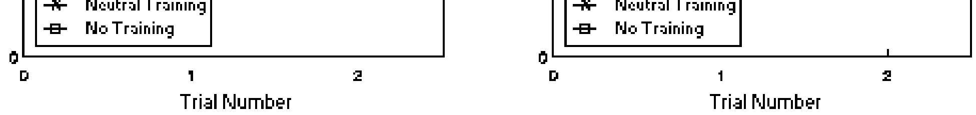


Figure 5.11 Search Performance Before and After Visual Lobe Training

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 5.23relates 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 Johannson, 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 Guttmann, 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.