GOOD PRACTICES IN VISUAL INSPECTION

Colin G. Drury
Applied Ergonomics Group Inc.
98 Meadowbrook Road
Willimasville, NY 14221-5029

Jean Watson
Federal Aviation Administration
Flight Standards Service

May 2002

Table of Contents

1.0 EXECUTIVE SUMMARY

Visual Inspection is the single most frequently-used aircraft inspection technique, but is still error-prone. This project follows previous reports on fluorescent penetrant inspection (FPI) and borescope inspection in deriving good practices to increase the reliability of NDI processes through generation of good practices based on analysis of the human role in the inspection system.

Inspection in aviation is mainly visual, comprising 80% of all inspection by some estimates, and accounting for over 60% of AD notices in a 2000 study. It is usually more rapid than other NDI techniques, and has considerable flexibility. Although it is usually defined with reference to the eyes and visible spectrum, in fact Visual Inspection includes most other non-machine-enhanced methods, such as feel or even sound. It is perhaps best characterized as using the inspectors' senses with only simple job aids such as magnifying loupes or mirrors. As such, Visual Inspection forms a vital part of many other NDI techniques where the inspector must visually assess an image of the area inspected, e.g. in FPI or radiography. An important characteristic of Visual Inspection is its flexibility, for example in being able to inspect at different intensities from walk-around to detailed inspection. From a variety of industries, including aviation, we know that when the reliability of visual inspection is measured, it is less than perfect. Visual inspectors, like other NDI inspectors, make errors of both missing a defect and calling a non-defect (misses and false alarms respectively).

This report used a Hierarchical Task Analysis (HTA) technique to break the task of Visual Inspection into five major functions: Initiate, Access, Search, Decision and Response. Visits to repair facilities and data collected in previous projects were used to refine these analyses. The HTA analysis was continued to greater depth to find points at which the demands of the task were ill-matched to the capabilities of human inspectors. These are points where error potential is high. For each of these points, Human Factors Good Practices were derived. Overall, 58 such Good Practices were developed, both from industry sources and human factors analyses. For each of these Good Practices, a specific set of reasons were produced to show why the practice was important and why it would be helpful.

Across the whole analysis, a number of major factors emerged where knowledge of human performance can assist design of Visual Inspection tasks. These were characterized as:

Time limits on continuous insepction performance
The visual environment
Posture and visual inspection performance
The effect of speed of working on inspection accuracy
Training and selection of inspectors
Documentation design for error reduction

Each is covered in some detail, as the principles apply across a variety of inspection tasks including visual inspection, and across many of the functions within each inspection task.

Overall, these 58 specific Good Practices and six broad factors help inspection departments to design inspection jobs to minimize error rates. Many can be applied directly to the "reading" function of other NDI techniques such as FPI or radiography.

2.0 OBJECTIVES AND SIGNIFICANCE

This study was commissioned by the Federal Aviation Administration (FAA), Office of Aviation Medicine for the following reasons:

2.1 Objectives

Objective 1. To perform a detailed human factors analysis of visual inspection.
Objective 2. To use the analysis to provide Human Factors guidance (best practices) to improve the overall reliability of visual inspection.

2.2 Significance

Visual inspection comprises the majority of the inspection activities for aircraft structures, power plants and systems. Like all inspection methods, visual inspection is not perfect, whether performed by human, by automated devices or by hybrid human/ automation systems. While some inspection probability of detection (PoD) data is available for visual inspection most recommendations for visual inspection improvement are based on unquantified anecdotes or even opinion data. This report uses data from various non-aviation inspection tasks to help quantify some of the factors affecting visual inspection performance. The human factors analysis brings detailed data on human characteristics to the solution of inspection reliability problems. As a result of this research, a series of best practices are available for implementation. These can be used in improved training schemes, procedures, design of equipment and the inspection environment so as to reduce the overall incidence of inspection error in visual inspection tasks for critical components.

3.0 INTRODUCTION

Visual inspection is the most often specified technique for airframes, power plants and systems in aviation. The FAA's Advisory Circular 43-204 (1997)[1] on Visual Inspection for Aircraft quotes Goranson and Rogers (1983)[2] to the effect that over 80% of inspections on large transport category aircraft are visual inspections (page 1). A recent analysis of Airworthiness Directives issued by the FAA from 1995 to 1999 (McIntire and Moore, 1993)[3] found that 561 out of 901 inspection ADs (62%) specified visual inspection. In fact, when these numbers are broken down by category, only 54% of ADs are visual inspection for large transport aircraft, versus 75% for the other categories (small transport, general aviation, rotorcraft).

3.1 Visual Inspection Defined

There are a number of definitions of visual inspection in the aircraft maintenance domain. For example, in its AC-43-204,[1] the FAA uses the following definition:

> "Visual inspection is defined as the process of using the unaided eye, alone or in conjunction with various aids, as the sensing mechanism from which judgments may be made about the condition of a unit to be inspected."

The ASNT's Non-Destructive Testing Handbook, Volume 8 (McIntire and Moore, 1993)[3] has a number of partial definitions in different chapters. Under Section 1, Part 1, *Description of Visual and Optical Tests* (page 2), it defines:

> "…. Visual and optical tests are those that use probing energy from the visible portion of the electromagnetic spectrum. Changes in the light's properties after contact with the test object may be detected by human or machine vision. Detection may be enhanced or made possible by mirrors, borescopes or other vision-enhancing accessories."

More specifically for aircraft inspection, on page 292 in Section 10, Part 2, for optically-aided visual testing of aircraft structure, visual inspection is defined by what it can *do* rather than what it *is*:

> "visual testing is the primary method used in aircraft maintenance and such tests can reveal a variety of discontinuities. Generally, these tests cover a broad area of the aircraft structure. More detailed (small area) tests are conducted using optically aided visual methods. Such tests include the use of magnifiers and borescopes."

However, there is more to visual inspection than just visual information processing.

3.2 Characteristics of Visual Inspection

As used in aviation, visual inspection goes beyond "visual," i.e. beyond the electromagnetic spectrum of visible wavelengths. In a sense, it is the default inspection technique: if an inspection is not one of the specific NDI techniques (eddy current, X-ray, thermography, etc.) then it is usually classified as visual inspection. Thus, other senses can be used in addition to the visual sense. For example, visual inspection of fasteners typically includes the action of feeling for fastener/structure relative movement. This involves active attempts, using the fingers, to move the fastener. In human factors, this would be classified as tactile or more generally haptic inspection. A different example is checking control cables for fraying by sliding a rag along the cable to see whether it snags. Other examples include the sense of smell (fluid leakage, overheated control pivots), noise (in bearings or door hinges) and feel of backlash (in engine blades, also in hinges and bearings). The point is that "visual" inspection is only partially defined by the visual sense, even though vision is its main focus.

Visual inspection is of the greatest importance to aviation reliability, for airframes, power plants and systems. It can indeed detect a variety of defects, from cracks and corrosion to loose fasteners, ill-fitting doors, wear and stretching in control runs and missing components. It is ubiquitous throughout aircraft inspection, so that few inspectors will perform a specialized NDI task without at least a "general visual inspection" of the area specified. Visual inspection also has the ability to find defects in assembled structures as well as components. With remote sensing, e.g. borescopes and mirrors, this *insitu* characteristic can be extended considerably. Visual inspection is the oldest inspection technique, in use from the pioneer days of aviation, and it can be argued that all other NDI techniques are enhancements of visual inspection. Radiographic and D-sight inspection are obvious extensions of visual inspection, as they give an image that is a one-to-one veridical representation of the original structure, in a way not different in principle to the enhancement provided by a mirror or a magnifying lens. Thus, understanding visual inspection is in many ways the key to understanding other inspection techniques. The previous reports in this series were obvious examples: FPI and borescope inspection. Almost all the other NDI techniques (with the exception of some eddy-current and ultrasonic systems, and tap tests for composites) have an element of visual inspection. Often the sensing systems have their signals processed in such a way as to provide a one-to-one mapping of the output onto the structure being examined. In this way they provide a most natural representation of the structure and help prevent errors associated with inspector disorientation. Examples would be thermography and radiographic images. Indeed Section 11, Part 1, of McIntine and Moore (1993)[3] lists specifically the visual testing aspects of leak testing, liquid penetrant, radiography, electromagnetic, magnetic particle, and ultrasonic testing to show the pervasiveness of visual inspection.

If visual inspection is important and ubiquitous, it is also flexible. First, visual inspection can often be orders of magnitude more rapid than NDI techniques. If all inspections were via specialist NDI techniques, aircraft would spend little time earning revenue. The ingenuity of NDI personnel and applied physicists has often been used to speed inspection, e.g. in inaccessible areas thus avoiding disassembly, but these innovations are for carefully pre-specified defects in pre-specified locations. The defining characteristic of visual inspection is its ability to detect a wide range of defect types and severities across a wide range of structures.

Clearly, NDI techniques extend the range of human perception of defects, even to hidden structures, but they are slower and more focused. For example, an eddy current examination of a component is designed to find a particular subset of indications (e.g. cracks) at particular pre-defined locations and orientations. Thus, for radius cracks, it is highly reliable and sensitive, but it may not detect cracks around fastener holes without a change to the probe or procedure. We can contrast the flexibility of visual inspection, i.e. range of defect types, severities, locations, orientations, with the specificity of other NDI techniques. Visual inspection is intended to detect literally any deviation from a correct structure, but it may only do so for a fairly large severity of indication. NDI techniques focus on a small subset of defect characteristics, but are usually more sensitive (and perhaps more reliable) for this limited subset.

One final aspect of flexibility for visual inspection is its ability to be implemented at many different levels. Visual inspection can range in level from the pilot's walk-around before departure to the detailed examination of one section of floor structure for concealed cracks using a mirror and magnifier. The FAA's AC-43-204[1] defines four levels of visual inspection as follows:

1. **Level 1.** Walkaround. The walkaround inspection is a general check conducted from ground level to detect discrepancies and to determine general condition and security.
2. **Level 2.** General. A general inspection is made of an exterior with selected hatches and openings open or an interior, when called for, to detect damage, failure, or irregularity.
3. **Level 3**. Detailed. A detailed visual inspection is an intensive visual examination of a specific area, system, or assembly to detect damage failure or irregularity. Available inspection aids should be used. Surface preparation and elaborate access procedures may be required.
4. **Level 4**. Special Detailed. A special detailed inspection is an intensive examination of a specific item, installation, or assembly to detect damage, failure, or irregularity. It is likely to make use of specialized techniques and equipment. Intricate disassembly and cleaning may be required.

However, other organizations and individuals have somewhat different labels and definitions. The ATA's Specification 100[4] defines a General Visual Inspection as:

> "…. a check which is a thorough examination of a zone, system, subsystem, component or part, *to a level defined by the manufacturer*, to detect structural failure, deterioration or damage and to determine the need for corrective maintenance." (my italics)

This aspect of leaving the definition to the manufacturer introduces another level of (possibly subjective) judgment into the decision. For example, one manufacturer of large transport aircraft defines a General Visual Inspection as:

> "A visual check of exposed areas of wing lower surface, lower fuselage, door and door cutouts and landing gear bays."

This same manufacturer defines Surveillance Inspection as:

> "..A visual examination of defined interval or external structural areas."

Wenner (2000)[5] notes that one manufacturer of regional transport aircraft categorizes inspection levels as:

Light service
Light visual
Heavy visual
Special

…. adding to the potential confusion. The point to be made is that level of inspection adds flexibility of inspection intensity, but at the price of conflicting and subjective definitions. This issue will be discussed later in light of research by Wenner (2000)[5] on how practicing inspectors interpret some of these levels.

In summary, visual inspection, while perhaps rather loosely defined, is ubiquitous, forms an essential part of many more specialized NDI techniques, and is flexible as regards the number and types of indication it can find and the level at which it is implemented. In order to apply human factors principles to improving visual inspection reliability, we need to consider the technical backgrounds of both inspection reliability and human factors.

Human factors has been a source of concern to the NDI community as seen in, for example, the NDE Capabilities Data Book (1997).[6] This project is a systematic application of human factors principles to the one NDI technique most used throughout the inspection and maintenance process.

4.0 TECHNICAL BACKGROUND: NDI RELIABILITY AND HUMAN FACTORS

There are two bodies of scientific knowledge that must be brought together in this project: quantitative NDI reliability and human factors in inspection. These are reviewed in turn for their applicability to visual inspection. This section is closely based on the two previous technique specific reports (Drury, 1999,[7] 2000[8]), with some mathematical extensions to the search and decision models that reflect their importance in visual inspection.

4.1 NDI Reliability

Over the past two decades there have been many studies of human reliability in aircraft structural inspection. Almost all of these to date have examined the reliability of Nondestructive Inspection (NDI) techniques, such as eddy current or ultrasonic technologies. There has been very little application of NDI reliability techniques to visual inspection. Indeed, neither the Non-Destructive Testing Handbook, Volume 8 (McIntire and Moore, 1993)[3] nor the FAA's Advisory Circular 43-204 (1997)[1] on Visual Inspection for Aircraft list either "reliability" or "probability of detection (PoD)" in their indices or glossaries.

From NDI reliability studies have come human/machine system detection performance data, typically expressed as a Probability of Detection (PoD) curve, e.g. (Rummel, 1998).[9]

This curve expresses the reliability of the detection process (PoD) as a function of a variable of structural interest, usually crack length, providing in effect a psychophysical curve as a function of a single parameter. Sophisticated statistical methods (e.g. Hovey and Berens, 1988)[10] have been developed to derive usable PoD curves from relatively sparse data. Because NDI techniques are designed specifically for a single fault type (usually cracks), much of the variance in PoD can be described by just crack length so that the PoD is a realistic reliability measure. It also provides the planning and life management processes with exactly the data required, as structural integrity is largely a function of crack length.

A recent issue of ASNT's technical journal, *Materials Evaluation* (Volume 9.7, July 2001)[11] is devoted to NDI reliability and contains useful current papers and historical summaries. Please note, however, that "human factors" is treated in some of these papers (as in many similar papers) in a non-quantitative and anecdotal manner. The exception is the paper by Spencer (Spencer, 2001)[12] which treats the topic of inter-inspector variability in a rigorous manner.

A typical PoD curve has low values for small cracks, a steeply rising section around the crack detection threshold, and level section with a PoD value close to 1.0 at large crack sizes. It is often maintained (e.g. Panhuise, 1989)[13] that the ideal detection system would have a step-function PoD: zero detection below threshold and perfect detection above. In practice, the PoD is a smooth curve, with the 50% detection value representing mean performance and the slope of the curve inversely related to detection variability. The aim is, of course, for a low mean and low variability. In fact, a traditional measure of inspection reliability is the "90/95" point. This is the crack size which will be detected 90% of the time with 95% confidence, and thus is sensitive to both the mean and variability of the PoD curve.

Two examples may be given of PoD curves for visual inspection to illustrate the quantitative aspects of reliability analysis. The first, shown in Figure 1, is taken from the *NDE Capabilities Data Book* (1997)[6] and shows the results of visual inspection of bolt holes in J-85 sixth stage disks using an optical microscope. Each point plotted as an "X" could only be an accept or reject, so that it must be plotted at either PoD = 0 (accept) or PoD = 1.0 (reject). The curve was fitted using probit regression, shown by Spencer (2001) to be an appropriate statistical model. The 90% PoD point 0.593 (15.1 mm) that corresponds to the 90/95 point is larger at 0.395 inches (10.0 mm), reflecting the fact that to be 95% certain that the 0.90 level of PoD has been reached, we need a crack length of about 10 mm rather than about 7 mm.

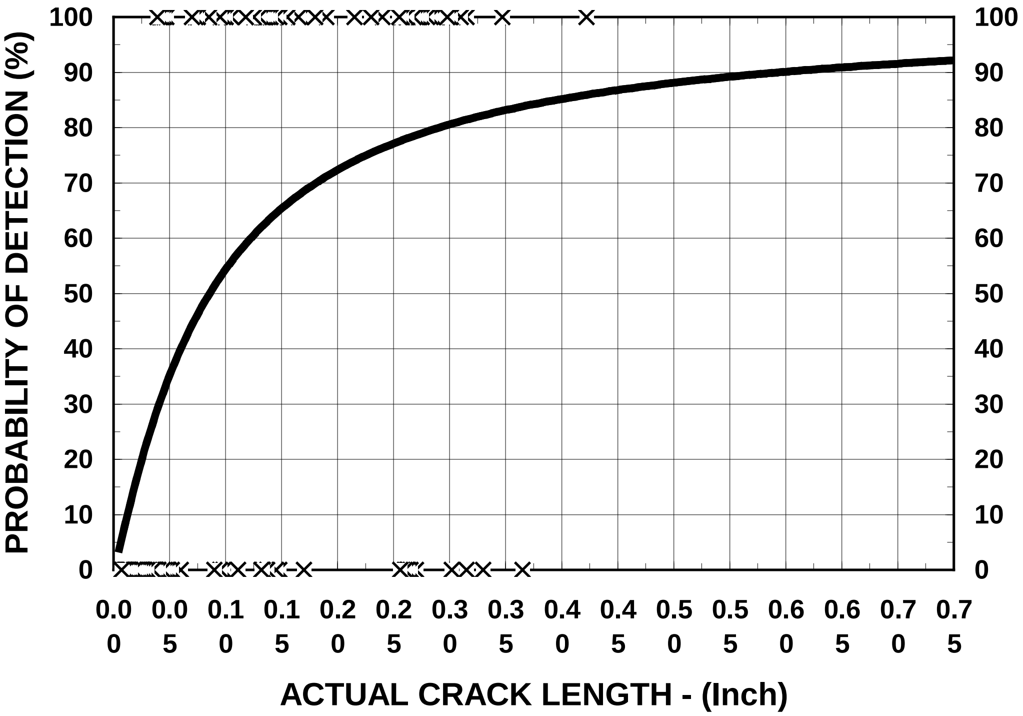


Figure 1. PoD curve of etched cracks in Inconel and Haynes 188 at 30X magnification.

The second example is from a Benchmark study of visual inspection by Spencer, Schurman and Drury (1996).[14] Here, ten inspectors inspected fuselage areas of an out-of-service B-737 for mainly cracks and corrosion. The overall PoD curve for known cracks is shown in Figure 2. A number of points about this curve are important to understanding the reliability of visual inspection. First, this was an on-site inspection using practicing inspectors, rather than a test of isolated specimens under laboratory conditions. Hence, the absolute magnitude of the crack lengths are larger than those in Figure 1. Second, the PoD curve does not appear to asymptote at a PoD of 1.0 for very large cracks. This implies that there is a finite probability of an inspector missing even very large cracks. Third, the variability about the curve means that crack length is not the only variable affecting detection performance. From our knowledge of human inspection performance (Section 4.2) we can see that crack width and contrast should affect PoD, as well as factors such as crack accessibility (Spencer and Schurman, 1995).[15]

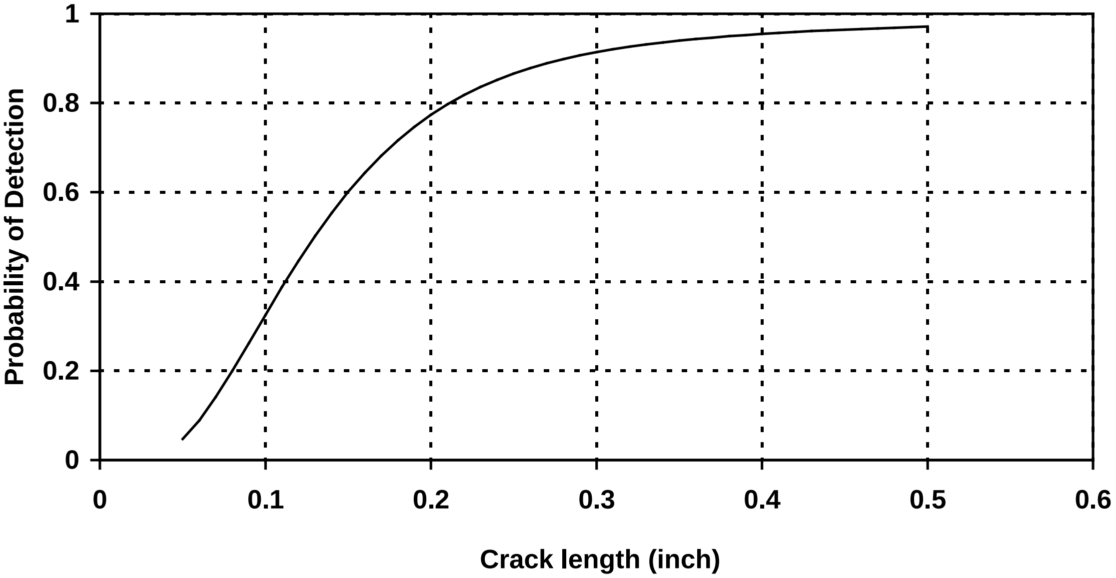


Figure 2. Mean PoD for visual inspection of known cracks in VIRP Benchmark study

In NDI reliability assessment one very useful model is that of detecting a signal in noise. Other models of the process exist (Drury, 1992)[16] and have been used in particular circumstances. The signal and noise model assumes that the probability distribution of the detector's response can be modeled as two similar distributions, one for signal-plus-noise (usually referred to as the signal distribution), and one for noise alone. (This "Signal Detection Theory" has also been used as a model of the human inspector, see Section 4.2). For given signal and noise characteristics, the difficulty of detection will depend upon the amount of overlap between these distributions. If there is no overlap at all, a detector response level can be chosen which completely separates signal from noise. If the actual detector response is less than the criterion or "signal" and if it exceeds criterion, this "criterion" level is used by the inspector to respond "no signal." For non-overlapping distributions, perfect performance is possible, i.e. all signals receive the response "signal" for 100% defect detection, and all noise signals receive the response "no signal" for 0% false alarms. More typically, the noise and signal distributions overlap, leading to less than perfect performance, i.e. both missed signals and false alarms.

The distance between the two distributions divided by their (assumed equal) standard deviation gives the signal detection theory measure of discriminability. A discriminability of 0 to 2 gives relatively poor reliability while discriminabilities beyond 3 are considered good. The criterion choice determines the balance between misses and false alarms. Setting a low criterion gives very few misses but large numbers of false alarms. A high criterion gives the opposite effect. In fact, a plot of hits (1 – misses) against false alarms gives a curve known as the Relative Operating Characteristic (or

ROC) curve which traces the effect of criterion changes for a given discriminability (see Rummel, Hardy and Cooper, 1989).[17]

The NDE Capabilities Data Book (1997)[6] defines inspection outcomes as:

		Flaw Presence	
		Positive	**Negative**
NDE Signal	**Positive**	True Positive No Error	False Positive Type 2 Error
	Negative	False Negative Type 1 Error	True Negative No Error

And defines

PoD = Probability of Detection = $\dfrac{TruePositives}{TruePositives + FalseNegatives}$

PoFA = Probability of False Alarm = $\dfrac{FalsePositives}{TrueNegatives + FalsePositives}$

The ROC curve traditionally plots PoD against (1 – PoFA). Note that in most inspection tasks, and particularly for any task on commercial aircraft, the outcomes have very unequal consequences. A failure to detect (1 – PoD) can lead to structural or engine failure, while a false alarm can lead only to increased costs of needless repeated inspection or needless removal from service.

This background can be applied to any inspection process, and provides the basis of standardized process testing. It is also used as the basis for inspection policy setting throughout aviation. The size of crack reliably detected (e.g. 90/95 criterion), the initial flaw size distribution at manufacture and crack growth rate over time can be combined to determine an interval between inspections which achieves a known balance between inspection cost and probability of component failure.

The PoD and ROC curves differ between different techniques of NDI (including visual inspection) so that the technique specified has a large affect on probability of component failure. The techniques of ROC and PoD analysis can also be applied to changing the inspection configuration, for example the quantitative study of multiple FPI of engine disks by Yang and Donath (1983).[18] Probability of detection is not just a function of crack size, or even of NDI technique. Other factors can assume great importance, particularly in visual-based inspection techniques. This points to the requirement to examine closely <u>all</u> of the steps necessary to inspect an item, and not just those involving the inspector.

If we are to examine all of the steps in visual inspection, a task breakdown is needed as a starting point. Later, in Section 4.2, we provide a generic task breakdown from human factors principles, but first we use aircraft-inspection-specific breakdowns AC-43-204 provides one such breakdown as follows:

1. Basis for the inspection
2. Preparation for the inspection
3. Implementation of the inspection
4. Evaluation

As an example of the considerations forming each of these four steps, their application to level 2 (General Inspection) is quoted from AC-43-204[1]:

1. Basis for Inspection. When a specific problem is suspected, the general inspection is carried out to identify, if possible, the difficulty. General inspections are also routinely used when panels are open for normal maintenance.

2. Preparation for the Inspection. Ensure cleanliness of the aircraft. The necessary tools and equipment required may include flashlight, mirror, notebook, droplight, rolling stool, tools for removal of panels, ladder stands, or platforms. Other aids such as jacking of the aircraft may or may not be discretionary; knowledge of a specific aircraft may be essential; and common problems may require information, even if not on the inspection card.

3. Implementation. General looking is not enough. As the inspector, you should continually ask "What is wrong with this picture?" Be inquisitive. Question whether you have seen this before. Move, shake, pull, twist, and push all parts possible. Apply weight to load bearing components. Compare one side to the other if applicable. Be aware of other systems in the inspection area. Look for abnormalities in the area, even if not related to this inspection. Adjusting the source of illumination, view items under inspection from different angles. Is the area pressurized? If so, does this affect any part of the inspection? Inspect all structural components, all moveable parts, all attach points, and brackets. Check all cables, conduits, and hoses for condition and clearance. Check condition and security of load and stress points. Look for chafing and fretting corrosion. Observe proximity of one part to another. Look for loose or missing fasteners, use of proper sealants, noticeable cracks, indications of corrosion, and debris in closed areas. Observe that cables, conduits, and hoses are properly routed. Observe that there is sufficient strain relief. Observe rivets for damage. Look for smoked rivets and discoloration of paint. (Localized chipping of paint, cracked paint on sealant, or fretting corrosion are indicative of movement.)

4. Findings. Transfer all information relating to discrepancies from your notebook. Record discrepancies as a work order. Discoveries during the inspection may indicate the need for a more detailed inspection. Depending on the findings, this may be either a Level 3 or Level 4.

The *Non-Destructive Testing Handbook Visual and Optical Testing*[3] has many lists pertaining to visual inspection, but these typically list the components to be inspected, the techniques available, and the defect types to be inspected for (e.g. p 156 – 158 weld inspection and p 262-269 for pressure vessels). All agree on the importance of “human factors” but see this as mainly training and motivation. As such, on appropriate intervention would be certification (e.g. p. 181-187). Rather than following this example, we shall consider the human inspector’s role in some technical detail in the following section as a prelude to performing a detailed task analysis for visual inspection, as the basis for deriving good practices.

4.2 Human Factors in Inspection

Note: There have been a number of recent book chapters covering this area, which will be referenced here rather than using the original research sources.

Human factors studies of industrial inspection go back to the 1950’s when psychologists attempted to understand and improve this notoriously error-prone activity. From this activity came literature of increasing depth focusing an analysis and modeling of inspection performance, which complemented the quality control literature by showing how defect detection could be improved. Two early books brought much of this accumulated knowledge to practitioners: Harris and Chaney (1969)[19] and Drury and Fox (1975).[20] Much of the practical focus at that time was on enhanced inspection techniques or job aids, while the scientific focus was on application of psychological constructs, such as vigilance and signal detection theory, to modeling of the inspection task.

As a way of providing a relevant context, we use the generic functions which comprise all inspection tasks whether manual, automated or hybrid. Table 1 shows these functions, with the specific application to visual inspection. We can go further by taking each function and listing its correct outcome, from which we can logically derive the possible errors (Table 2). Later in the report we will progressively expand each generic function to derive human factors good practices in visual inspection.

Humans can operate at several different levels in each function depending upon the requirements. Thus, in Search, the operator functions as a low-level detector of indications, but also as a high-level cognitive component when choosing and modifying a search pattern. It is this ability that makes humans uniquely useful as self-reprogramming devices, but equally it leads to more error possibilities. As a framework for examining inspection functions at different levels the skills/rules/knowledge classification of Rasmussen (1983)[21] will be used. Within this system, decisions are made at the lowest possible level, with progression to higher levels only being invoked when no decision is possible at the lower level.

For most of the functions, operation at all levels is possible. Access to an item for inspection is an almost purely mechanical function, so that only skill-based behavior is appropriate. The response function is also typically skill-based, unless complex diagnosis of the defect is required beyond mere detection and reporting. Such complex diagnosis is often shared with others, e.g. engineers or managers, if the decision involves expensive procedures such as changing components or delaying flight departure.

Function	Visual Inspection Description
1. Initiate	All processes up to accessing the component. Get and read workcard. Assemble and calibrate required equipment.
2. Access	Locate and access inspection area. Be able to see the area to be inspected at a close enough level to ensure reliable detection.
3. Search	Move field of view across component to ensure adequate coverage. Carefully scan field of view using a good strategy. Stop search if an indication is found.
4. Decision	Identify indication type. Compare indication to standards for that indication type.
5. Response	If indication confirmed, then record location and details. Complete paperwork procedures. Remove equipment and other job aids from work area and return to storage. If indication not confirmed, continue search (3).

Table 1. Generic function description and application to visual inspection

Function	Correct Outcome	Logical Errors
Initiate	Inspection equipment functional, correctly calibrated and capable.	1.1 Incorrect equipment 1.2 Non-working equipment 1.3 Incorrect calibration 1.4 Incorrect or inadequate system knowledge
Access	Item presented to inspection system	2.1 Wrong item presented 2.2 Item mis-presented 2.3 Item damaged by presentation
Search	Indications of all possible non-conformities detected, located	3.1 Indication missed 3.2 False indication detected 3.3 Indication mis-located 3.4. Indication forgotten before decision
Decision	All indications located by Search correctly measured and classified, correct outcome decision reached	4.1 Indication incorrectly measured/confirmed 4.2 Indication incorrectly classified 4.3 Wrong outcome decision 4.4 Indication not processed
Response	Action specified by outcome decision taken correctly	5.1 Non-conforming action taken on conforming item 5.2 Conforming action taken on non-conforming item 5.3 Action incomplete

Table 2. Generic functions and errors for visual inspection

4.2.1 Critical Functions: search and decision

The functions of search and decision are the most error-prone in general, although for much of inspection, especially NDI, setup can cause its own unique errors. Search and decision have been the subjects of considerable mathematical modeling in the human factors community, with direct relevance to visual inspection. The sections on search and decision are adapted from Drury (1999).[7]

Search: In visual inspection, and in other tasks such as X-ray inspection, the inspector must move his/her eyes around the item to be inspected to ensure that any defect will eventually appear within an area around the line of sight in which it is possible to achieve detection. This area, called the visual lobe, varies in size depending upon target and background characteristics, illumination and the individual inspector’s peripheral visual acuity. As successive fixations of the visual lobe on different points occur at about three per second, it is possible to determine how many fixations are required for complete coverage of the area to be searched.

Eye movement studies of inspectors show that they do not follow a simple pattern in searching an object. Some tasks have very random appearing search patterns (e.g., circuit boards), whereas others show some systematic search components in addition to this random pattern (e.g., aircraft structures). However, all who have studied eye movements agree that performance, measured by the probability of detecting an imperfection in a given time, is predictable assuming a random search model. The equation relating probability (p_t) of detection of a single imperfection in a time (t) to that time is

$$p_t = 1 - \exp\left(-\frac{t}{\bar{t}}\right)$$

where $\bar{t}$ is the mean search time. Further, it can be shown that this mean search time can be expressed as

$$\bar{t} = \frac{t_o A}{apn}$$

where

t_o = average time for one fixation
A = area of object searched
a = area of the visual lobe
p = probability that an imperfection will be detected if it is fixated. (This depends on how the lobe (a) is defined. It is often defined such that p = ½. This is an area with a 50% chance of detecting an imperfection.

From these equations we can deduce that the time taken to search an area is extremely important in determining search success. Thus, there is a speed/accuracy tradeoff (SATO) in visual search, so that if insufficient time is spent in search, defects may be missed. We can also determine what factors affect search performance, and modify them accordingly.

Thus, the area to be searched (A) is a direct driver of mean search time. Anything we can do to reduce this area, e.g. by instructions about which parts of an object not to search, will help performance. Visual lobe area needs to be maximized to reduce mean search time, or alternatively to increase detection for a given search time. Visual lobe size can be increased by enhancing target background contrast (e.g. using the correct lighting) and by decreasing background clutter. It can also be increased by choosing operators with higher peripheral visual acuity and by training operators specifically in visual search or lobe size improvement. Research has shown that there is little to be gained by reducing the time for each fixation, t_o , as it is not a valid selection criterion, and cannot easily be trained.

We can extend the equations above to the more realistic case of multiple targets present on an area or item searched (Morawski, Drury and Karwan, 1980).[22] If there are (n) targets then the time to locate the *first* target is also exponential, but with $\bar{t}$ for (n) identical targets related to $\bar{t}$ for 1 target by

$$\bar{t}_n = \frac{1}{n} \bar{t}_1$$

That is, the more targets that are present, the faster the first one will be found. This formulation can be extended to (n) different targets (Morawski, Drury and Karwan, 1980)[22] and to the time to find <u>each</u> of the targets (Drury and Hong, 2001[23]; Hong and Drury, 2002).[24]

Of course, when the search is part of an inspection task, there may be zero targets present, i.e. the item or area may be defect free. Under these circumstances, the inspector must make a decision on when to stop searching and move on to another item or area. This decision produces a stopping time for zero defects in contrast to a search time when at least one defect is found. A stopping time also applies when the inspector's search process fails even though defects are present. It is possible to use optimization techniques to determine what the stopping time <u>should</u> be, given the probability of a defect being present the cost of the inspector's time, and the cost of missing a defect. This procedure has been used for both random and systematic search models (Morawski and Karwan and Drury, 1992[25]; Karwan, Morawski, and Drury, 1995)[26]. In the simplest case of a single target for a random search model we take the probabilities and costs for the three outcomes shown in Table 3 and sum the (cost x probability) of each outcome.

Outcome	**Probability**	**Cost**
1. No defect present	$1 - p'$	$- k t$
2. Defect present, not detected	$p' (\exp(-t / \bar{t}))$	$- k t$
3. Defect present, detected	$p' (1 - \exp(-t / \bar{t}))$	$V - k t$

Table 3. Probabilities and costs for inspection outcomes for a prior probability of defect = p'

Note that if there is no defect present or if the defect is not detected, the "value" is just minus the cost of the inspector's time at \$k per hour. If a defect is present and detected, there is a positive value \$V, usually a large number. (We could equally well use the cost of missing a defect instead of the value of finding a defect: the math is the same.) We can find the long-term expected value of the inspection process by summing (probability X value) across all three outcomes. This gives:

$$E\ (\text{value}) = -\,k\,t\,(1 - p') - k\,t\,p' \exp(-t / \bar{t}\,) + (V - k\,t\,)\,p'\,(\,1 - \exp(-t / \bar{t}\,)\,)$$

This expected value can be maximized by some particular stopping time t* , which we can find by equating the first derivative of the equation to 0.0. This gives:

$$t^* = \bar{t}\ \log_e [\ V\ p' / k]$$

Note that t* increases when p' is high, V is high and k is low. Thus, a longer time should be spent inspecting each area where

- There is a greater prior probability of a defect
- There is a greater value to finding a defect
- There is a lower cost of the inspection.

In fact, when people perform inspection tasks, they tend to choose stopping times in the same way that this simple model implies (Chi and Drury, 1998[27]; Baveja, Drury and Malone, 1996[28]). This is important in practice as it shows the factors affecting the Speed / Accuracy Trade Off (SATO) for the search function of inspection. Note that we are not implying that we should find the cost of a missed defect and make a rather unethical calculation of the costs of an aircraft catastrophe compared to the costs of paying an inspector. That is not how the MSG-3 process works. But analyses such as the derivation of optimal stopping time t* allow us to define in a quantitative manner the pressures on inspectors, and hence, derive good practices for helping inspectors improve their effectiveness. Note also that the analysis above represents only visual search (hence there are no decision errors such as false alarms), that it only covers the simplest simulation of one possible defect with a known prior probability, and that it assumes that a rather naïve economic maximization is the ultimate goal of the inspection system. These limitations can be removed with more complex models, e.g. Chi and Drury (2001)[29].

The equation given for search performance assumed random search, which is always less efficient than systematic search. Human search strategy has proven to be quite difficult to train, but recently Wang, Lin and Drury (1997)[30] showed that people can be trained to perform more systematic visual search. Also, Gramopadhye, Drury and Sharit (1997)[31] showed that particular forms of feedback can make search more systematic.

Decision: Decision-making is the second key function in inspection. An inspection decision can have four outcomes, as shown in Table 4. These outcomes have associated probabilities, for example the probability of detection is the fraction of all nonconforming items that are rejected by the inspector shown as p_2 in Table 4.

	True State of Item	
Decision of Inspector	**Conforming**	***Nonconforming***
Accept	Correct accept, p_1	Miss, (1 - p_2)
Reject	False alarm, (1 - p_1)	Hit, p_2

Table 4. Four outcomes of inspection decisions

Just as the four outcomes of a decision-making inspection can have probabilities associated with them, they can have costs and rewards also: costs for errors and rewards for correct decisions. Table 5 shows a general cost and reward structure, usually called a "payoff matrix," in which rewards are positive and costs negative. A rational economic maximizer would multiply the probabilities of Table 4 by the corresponding payoffs in Table 5 and sum them over the four outcomes to obtain the expected payoff. He or she would then adjust those factors under his or her control. Basically, SDT states that p_1 and p_2 vary in two ways. First, if the inspector and task are kept constant, then as p_1 increases, p_2 decreases, where the balance between p_1 and p_2 is defined mathematically. p_1 and p_2 can be changed together by changing the discriminability for the inspector between acceptable and rejectable objects. The most often tested set of assumptions comes from a body of knowledge known as the theory of signal detection, or SDT (McNichol, 1972).[32] This theory has been used for numerous studies of inspection, for example, sheet glass, electrical components, and ceramic gas igniters, and has been found to be a useful way of measuring and predicting performance. It can be used in a rather general nonparametric form (preferable) but is often seen in a more restrictive parametric form in earlier papers (Drury and Addison, 1963).[33] McNichol is a good source for details of both forms.

	True State of Item	
Decision of Inspector	**Conforming**	***Nonconforming***
Accept	a	-b
Reject	-c	d

Table 5. Four payoff values of inspection decisions

The objective in improving decision-making is to reduce decision errors. There can arise directly from forgetting imperfections or standards in complex inspection tasks or indirectly from making an incorrect judgment about an imperfection's severity with respect to a standard. Ideally, the search process should be designed to improve the conspicuity of rejectable imperfections (nonconformities) only, but often the measures taken to improve conspicuity apply equally to nonrejectable imperfections. Reducing decision errors usually reduces to improving the discriminability between imperfection and a standard.

Decision performance can be improved by providing job aids and training that increase the size of the apparent difference between the imperfections and the standard (i.e. increasing discriminability). One example is the provision of limit standards well-integrated into the inspector's view of the item inspected. Limit standards change the decision-making task from one of absolute judgment to the more accurate one of comparative judgment. Harris and Chaney (1969)[19] showed that limit standards for solder joints gave a 100% performance improvement in inspector consistency for near-borderline cases.

One area of human decision-making that has received much attention is the vigilance phenomenon. It has been known for half a century that as time on task increases, then the probability of detecting perceptually-difficult events decreases. This has been called the vigilance decrement and is a robust phenomenon to demonstrate in the laboratory. Detection performance decreases rapidly over the first 20-30 minutes of a vigilance task, and remains at a lower level as time or task increases. Note that there is not a period of good performance followed by a sudden drop: performance gradually worsens until it reaches a steady low level. Vigilance decrements are worse for rare events, for difficult detection tasks, when no feedback of performance is given, where the task is highly repetitive and where the person is in social isolation. All of these factors are present to some extent in visual inspection of engines (e.g. the repetitive nature of inspecting a whole row of rivets or similar structural elements, so that prolonged vigilance is potentially important here.

A difficulty arises when this body of knowledge is applied to inspection tasks in practice. There is no guarantee that vigilance tasks are good models of inspection tasks, so that the validity of drawing conclusions about vigilance decrements in inspection must be empirically tested. Unfortunately, the evidence for inspection decrements is largely negative. Discussion of the vigilance decrement and associated good practices will be covered more thoroughly in Section 7.

From the above analysis, it is clear that inspection is not merely the decision function. The use of models such as signal detection theory to apply to the whole inspection process is misleading in that it ignores the search function. For example, if the search is poor, then many defects will not be located. At the overall level of the inspection task, this means that PoD decreases, but this decrease has nothing to do with setting the wrong decision criteria. Even such devices as ROC curves should only be applied to the decision function of inspection, not to the overall process unless search failure can be ruled out on logical grounds.

This can be illustrated from the data on visual inspection of lap joints for rivet cracks (Spencer, Schurman and Drury, 1996).[14] In the Benchmark evaluation of inspection performance noted earlier, one task was a standardized one of inspecting several of the panels with (grown) cracks starting at rivet holes. These were the panels used in the earlier ECRIRE study of eddy current inspection (Spencer and Schurman, 1995).[15] By analyzing video tapes of the inspectors performing this inspection task, it was possible to find out whether the inspection process at each rivet had been only search or search-plus-decision. Decisions could be seen from the inspectors interrupting their search to change the angle of their flashlight, or more their head for a different viewing angle or even feel the rivet area. Thus, search failure (i.e. never locating an indication) could be distinguished from decision failure (either failing to report an indication as a defect (miss), or reporting a defect where none existed (false alarm)). Figure 3 and 4 show the distributions across inspectors of search and decision success, respectively (from Drury, 1999).[7]

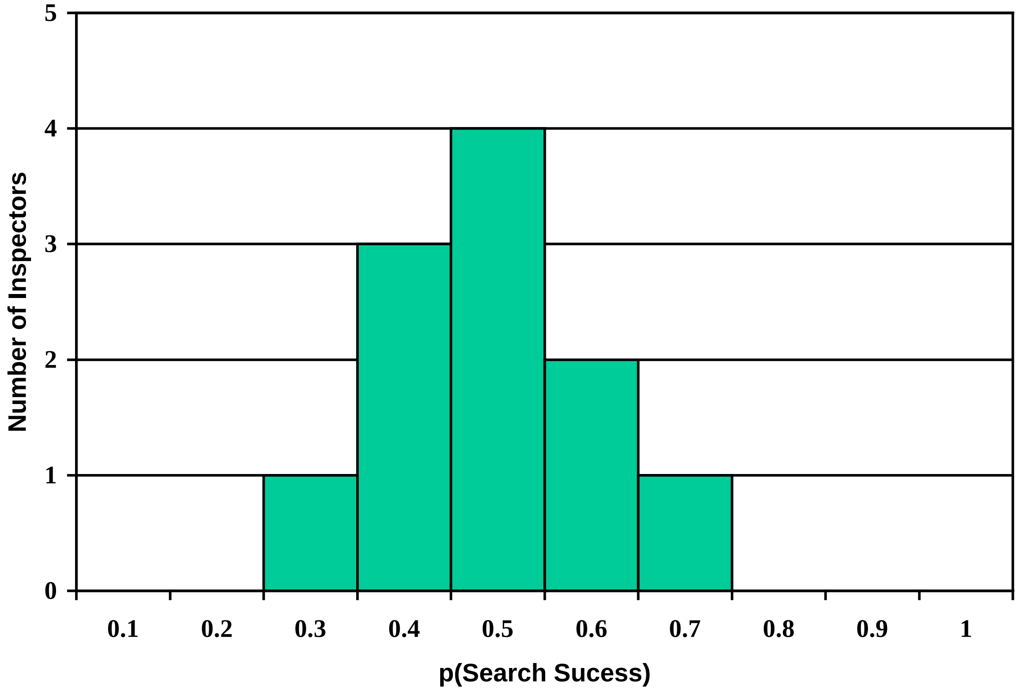


Figure 3. Distribution of search performance for 11 visual inspectors.

Note that probability of search success is quite narrowly grouped around a mean of 0.6. This shows that most of the lack of defect detection was due to a consistent, but not impressive search performance. Figure 4 shows a ROC plot in that it plots the two aspects of decision performance against each other. In this figure, we have used the positive aspects of performance (hits, correct acceptance) for the two axes, so that better performance is indicated by increases along both axes. Most ROC curves plot hits against false alarms to give better performance towards the upper left corner of the graph, which makes interpretation non-intuitive.

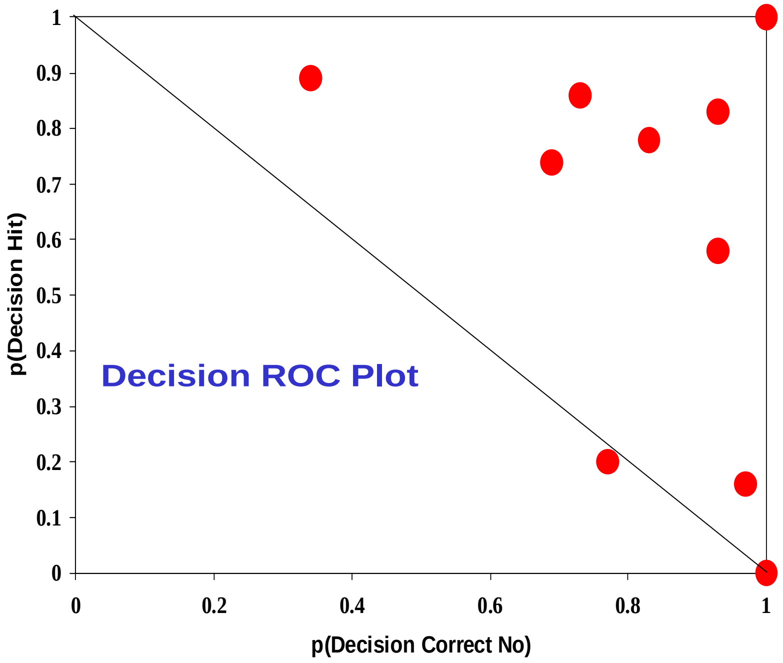


Figure 4. ROC curve showing distribution of decision performance for 11 visual inspectors

This section has shown a generic function description of visual inspection and used this to highlight the two functions most likely to cause errors: search and decision. Key variables affecting the reliability of search and decision have been derived from the respective models, and will become the basis for deriving human factors good practices in visual inspection.

5.0 RESEARCH OBJECTIVES

1. Review the literature on (a) NDI reliability and (b) human factors in inspection.
2. Apply human factors principles to the use of visual inspection, so as to derive a set of recommendations for human factors good practices.

6.0 METHODOLOGY

While there are PoD curves available for visual inspection in aircraft inspection (See Section 4.2 for two examples), they rarely show the effects of variables other than defect size and (perhaps) inspector-to-inspector variability. Thus, they are not a good source for determining which factors affect visual inspection performance, and hence for deriving human factors good practices. Three sources were used for deriving good practices:

1. Reference information on aircraft visual inspection, such as ACs and reports of the Visual Inspection Research Program at Sandia National Laboratories' Aging Aircraft NDI Validation Center (AANC).

2. Reference to the extensive literature on factors affecting visual inspection outside of the aviation industry.

3. Observation and task analysis of aircraft inspection tasks covering a wide range of different activities. These ranged from pre-flight inspection and overnight inspection (neither of which is carried out by people with a job title of inspector) to letter checks (B, C, D or equivalent) performed at airline and third party maintenance facilities.

In fact, one observation and task analysis was used, as in the earlier FPI and borescope reports, to structure the remainder of this report. The author has worked with many visual inspectors over the past decade. These interactions have been at airline and third party maintenance facilities and some in the course of taking part in the AANC's Visual Inspection Research Programs. Examples of inspection tasks observed can be found in the original reports of task analyses of 21 visual inspection tasks published in Phase 1 of the FAA/OAM initiative on Human Factors in Aviation Maintenance and Inspection. These were in the form of a task description and analysis in terms of the human subsystem involved and human factors relevant observations, as shown in Table 6. This gives part of a Honeycomb Panel Inspection and is typical of the whole set of visual inspection task analysis.

From this set of visual inspection task analyses, a list of 108 human factors issues was obtained to form part of the basis for the human factors good practices given later. These issues represent both current good practices and error opportunities. The main tool in deriving good practices was to progressively redefine the generic function analysis of Table 1 until it became a detailed list of both task steps and potential errors to be avoided in implementing these steps. This progress redescription is known as Hierarchical Task Analysis. .

6.1 Hierarchical Task Analysis

As noted above, the function analysis of Table 1 was progressively refined to produce a detailed task description of the visual inspection process. Because each function and process is composed of tasks, which are in turn composed of subtasks, a more useful representation of the task description was needed. A method that has become standard in human factors, Hierarchical Task Analysis (HTA) was used. In HTA, each function and task is broken down into sub-tasks using the technique of progressive re-description. At each breakdown point there is a plan, showing the decision rules for performing the sub-tasks. Often the plan is a simple list ("Do 3.1 to 3.5 in order") but at times there are choices and branches.

Task: HONEYCOMB PANEL INSPECTION
Location: Left Wing **Control:** Continuous, Discrete
Attention: Number of time-share tasks **Perception:**
Memory: STSS, Working, Long-Term **Decision:** Sensitivity, Criterion, Timing
Senses: Visual, Tactile, Auditory **Posture:** Reaching, Forces, Balance, Extreme Angles
Feedback: Quality, Amount, Timing

	Task Analysis									
	Sub-Systems									
Task Description	**A**	**S**	**P**	**D**	**M**	**C**	**F**	**P**	**O**	**Observations**
INITIATE										
1.0 Collect workcard from supervisor										Workcard specified area to be inspection: left and right wings and drawing of trailing and leading edge honeycomb panels.
2.0 Read workcard.			X		X					Workcard specified key points: areas susceptible to cracks.
3.0 Determine key areas.		X	X							Object required for tapping not specified in the workcard.
ACCESS										
1.0 Go to the aircraft. 1.1 Assure that wing flap lowered.										
2.0 Enter the aircraft fuselage through the entry door. (Scaffolding built around the aircraft.)										
3.0 Get on to the top of the left wing under the middle exit fuselage door for performing the inspection on top of the wing surface.						X		X	X	Top surface could be wet and slippery. This could be dangerous especially at the wing edges.
4.0 Get on the wing and use the platform built underneath to perform the inspection under the wings. 4.1 If the platform does not extend all the way procure a moving platform and continue inspection.						X		X	X	Reaching edges of the wing is dangerous because it is difficult to get a proper foothold.

SEARCH										
1.0 Auditory inspection: top wing surface										
1.1 Tap wing surface using a stick.										
1.2 Start from fuselage side moving towards wing up.										Systematic tapping pattern: This ensures that the entire area has been covered. There was a lot of intermittent interfering noise in the background. This could affect the auditory judgment needed in this inspection.
1.3 Listen for unusual auditory signal.	X	X								
2.0 Visual search on top surface key area: area just below off-wing slides are highly susceptible to cracks.			X		X					Similar pattern may not be adopted by all inspectors.
2.1 Hold flashlight perpendicular to the surface to look for cracks.			X		X					Possibility of missing area while tapping if a systematic pattern is not adopted by all the inspectors.
2.2 Hold flashlight at a grazing incident to look for bulges, ripples and delaminations on wing surface.	X		X							Inspectors adopt a very casual attitude while performing the entire task.
3.0 Use the flashlight under the wings to look for cracks.									X	Poor lighting under the wings.
3.1Hold the flashlight perpendicular to the wing surface.									X	The platform under the wings does not cover the entire area and the inspector has to procure a moving platform (which is not always available) to complete the inspection.
3.2 Tap the surface under the wings similar to the top surface and listen for unusual auditory signals.					X					The above mentioned activity could disrupt the search process.

Table 6. Example of Early Visual Inspection Task Analysis (Drury et al, 1990)

The HTA applied to visual inspection of airframes and engines can be found in Appendix 1. The overall level is broken into its branches each of which is then carried further in a tabular form to provide the link to human factors knowledge. The tabular form of each branch is given in Appendix 1. What this shows is a more complete task description of each sub-task under “Task Description”. The final column, headed “Task Analysis” shows the human factors and other system reliability issues in the form of questions that must be asked in order to ensure reliable human and system performance. Essentially, this column gives the human factors issues arising from the task, making the link between the human factors literature in Section 4.2 and the original Function level description in Table 1.

7.0 RESULTS

7.1 Detailed Good Practices

As in previous studies in this series, the direct presentation of human factors good practices is found in Appendix 2. It is given as Appendix 2 because it is so lengthy, with 58 entries. It is organized process-by-process following the HTA in Appendix 1. For each good practice, there are three columns:

> **1. Process:** Which of the seven major processes is being addressed?
> **2. Good Practice:** What is a recommended good practice within each process? Each good practice uses prescriptive data where appropriate, e.g. for time ob task. Good practices are written for practicing engineers and managers, rather than as a basis for constructing legally-enforceable rules and standards.
> **3. Why?** The logical link between each good practice and the errors it can help prevent. Without the "why" column, managers and engineers would be asked to develop their own rationales for each good practice. The addition of this column helps to train users in applying human factors concepts, and also provides help in justifying any additional resources.

There is no efficient way of summarizing the 58 detailed good practices in Appendix 2: the reader can only appreciate them by reading them. It is recommended that one process, e.g. Decision, is selected first and examined in detail. The good practices should then be checked in turn with each inspector performing the job to find out whether they are actually met. Again, the question is not whether a practice is included in the operating procedures, but whether it is followed for all visual inspections by all inspectors. The good practices in Appendix 2 can even be separated and used as individual check items. These can the be sorted into, for example, those which are currently fully implemented, those which can be undertaken immediately, and those which will take longer to implement.

7.2 Control Mechanisms

Some issues, and their resulting good practices, are not simple prescriptions for action, but are pervasive throughout the visual inspection system. Note that this report does not go into depth on the background of each control mechanism, as background material is readily available on each. *The Human Factors Guide for Aviation Maintenance 3.0*[35] is one readily accessible source of more information. This is available at the HFAMI web site: http://hfskyway.faa.gov. An additional more general source is the *ATA Spec 113 Human Factors Programs*,[35] available on the ATA's web site: www.air-transport.org

7.2.1 Time Limits in Inspection performance

The whole issue of sustained performance on inspection tasks is applicable to many types of aviation inspection, such as visual inspection and NDI. The following section provides more detail on the issue than was given in Section 4.2, and is adapted from a response requested of the author by the FAA and NTSB during 2001. Failure of both airframe inspection (Aloha incident) and engine inspection (Sioux City incident, Pensacola incident) has highlighted the potential impact of human limitations in inspection performance. A common thread in all three incidents was that inspection failure occurred during inspection tasks of normal working duration, i.e. a working shift with typical breaks. This vigilance decrement phenomenon is characterized by detection performance decreasing rapidly over the first 20-30 minutes of a vigilance task, and remaining at a lower level as time on task increases. As noted earlier, while this is easy to demonstrate in the laboratory, there is considerable argument in the human factors community over how much transfers to jobs such as aircraft inspection, or even inspection in a broader industrial context.

We begin by examining the laboratory evidence for vigilance tasks, noting the conditions under which decrement is most and least likely. Then we examine inspection tasks such as those implicated in the accidents above to determine what features are shared with laboratory vigilance tasks. Finally, we review studies that have attempted to measure vigilance decrement in field conditions, both for aviation inspection and for broader industrial inspection tasks. Human Factors Good Practices are drawn concerning time limits for inspection task performance.

The Vigilance Literature: A watch keeper's ability to maintain sustained attention first came under experimental scrutiny in World War II. The research was driven by the finding that trained observers in anti-submarine patrol aircraft reported less detections as their watch progressed (Mackworth, 1948).[37] The task was simulated in the laboratory with an apparatus that produced regular visible events, most of which were benign, but occasional ones that were defined as "signals" or "targets." Using naval personnel as participants, Mackworth found that detection performance decreased in every half-hour interval of the task. He labeled this the "vigilance decrement." Because he used half-hour time periods in the Navy's standard four-hour watch for collecting his data, his results are often interpreted as vigilance declining after 30 minutes of time on task. This is something of a misconception, as in fact about half of the loss is found in the first 15 minutes, and performance does not get much worse beyond 30 minutes (Teichner, 1974,[38] quoted in Huey and Wickens, 1993).[39] Indeed, Figure 5 shows the time course of detection performance in one study by (Craig, 1977),[40] when the initial fall in performance can be seen clearly.

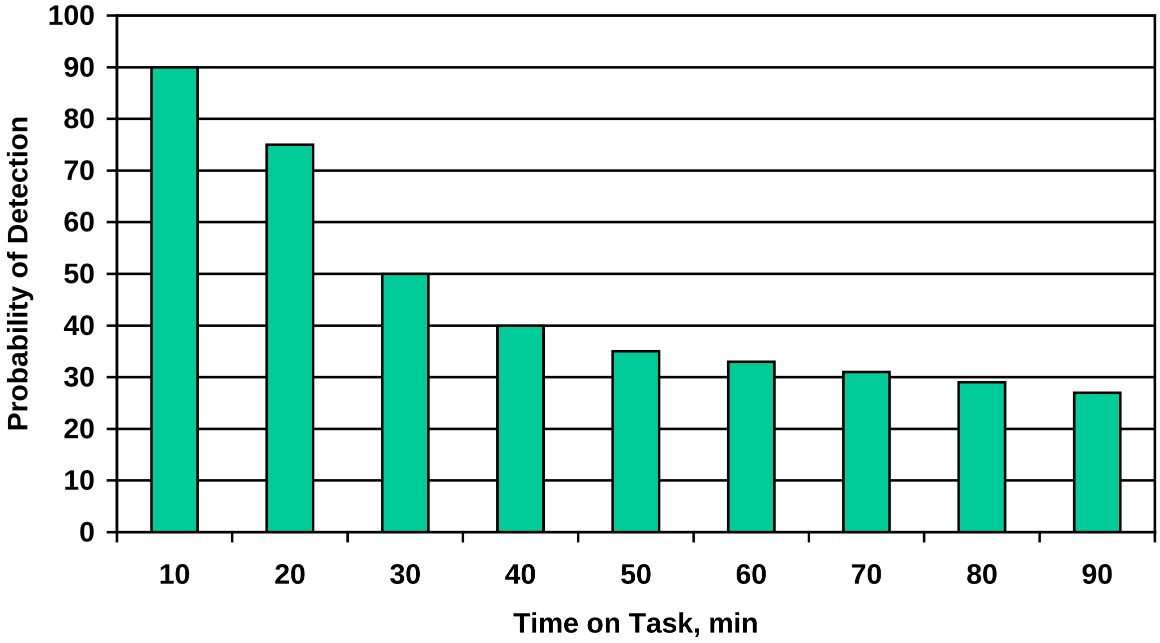


Figure 5. Time course of probability of detection in a typical vigilance task.

Since the early studies, a considerable body of knowledge has been accumulated on vigilance tasks, with thousands of experiments in many laboratories. A laboratory vigilance task has participants attempting to detect relatively rare signals but important in a continuous task that has the participant's full attention. Performance is measured by:

Hit Rate = probability of detecting a true signal
False Alarm Rate = probability of responding "signal" to a non-signal event

The general finding is that hit rate decreases with time on task, sometimes accompanied by a reduction in false alarm rate. This can be interpreted in terms of the Signal Detection Theory (SDT) model of decision making given in Section 4.2. If hit rate decreases while false alarm rate remains constant, this is a true performance decrement, as the participant's ability to distinguish between targets and non-target events has been impaired. It is known as a "sensitivity decrement." Conversely, if hit rate or false alarm rate both decrease, then there has been a change in the participant's willingness to report anything, signal or not. This is known as a "bias change," or because of the way bias is typically measured, a "bias increment" (Wickens and Hollands, 2000, page 37).[41] In fact, in SDT terms a bias increment is an optimal response to very infrequent signals: observers who try to be correct as often as possible *should* decrease their response rate, which will reduce both hits and false alarms.

Vigilance decrements have generally been found using relatively unskilled participants, in often an abstract task, with no social interaction and no external interruptions. The factors known to affect vigilance performance have been classified (Wickens and Hollands, 2000)[41] into those that contribute to the **Sensitivity Decrement**:

1. Low signal strength, i.e. targets not easily distinguished from background.
2. Time or location uncertainly, i.e. targets do not appear at regular intervals, or at specific locations.
3. Higher memory load, i.e. having to remember what a signal looks like rather than having a typical signal permanently in view.
4. Observers who are not highly practiced, i.e. the task is not automatic.

Other factors contribute to **Bias Increment**:

1. Low probability that an event is a signal, i.e. many events, few of which should be responded to.
2. Low levels of feedback, i.e. observers rarely find out whether or not they missed a true signal. (Feedback is more generally a part of the payoff system associated with the defined rewards of a vigilance task. However, these rewards are defined as they constitute feedback.)
3.

Overall, it appears that sensitivity loss comes from sustained high levels of cognitive demand. The task may be boring subjectively, but it is not easy. A relatively high level of mental resources must be applied to the task over prolonged periods.

In contrast. bias changes are thought to arise from the decreased expectancy of a signal. As observers expect less signals, they report less signals. For example, if the training provides 5 signals in 50 events, then the observer will be expecting about 10% of events to be a signal. If signal rate is much lower (1%, 0.1% or even less for many aviation inspection tasks) then less responses will result over time as the observer adapts to the reality beyond the training.

Clearly, inspection tasks can often be characterized as attempting to detect rare (even extremely rare) but important signals over long periods of time. Thus, a priori, vigilance and inspection tasks have features in common, namely sustained attention. But equally, vigilance and inspection tasks may be quite different. Inspection often occurs in the noisy and social environment of the hangar rather than in the sound proofed isolation of a laboratory. Table 7 has been compiled to give a direct comparison between features known to lead to poor vigilance performance (column 1) and equivalent features of inspection tasks (column 2). For a more theoretical description of vigilance studies, see Huey and Wickens (1993),[40] Parasuraman, Warm and Dember (1987),[42] Warm and Dember (1998),[43] Molloy and Parasuraman (1996).[44]

VIGILANCE TASK ATTRIBUTE	INSPECTION TASK ATTRIBUTE
Important Signals	Cracks or other defects that can have direct safety consequences.
Rare Signals	Defects can range from quite common, e.g. corrosive areas on older aircraft, to extremely rare (e.g. cracks in jet engine titanium hubs). However, under most circumstances far less than 1 out of 10 inspected components will contain a reportable defect.
Low Signal Strength	Most defects are perceptually difficult to detect, often occurring within a background of non-defects, e.g. cracks among dirt marks and scratches.
Long Time on Task	Time on task can vary from a few minutes to about 2 hours without a break. Scheduled breaks are typically four per shift, but many tasks are self-paced so that inspectors can break early or continue beyond scheduled time to complete an area or component.
High Memory Load	Prototypical defects are usually stored in the inspector's memory, rather than being presented as part of the task. Sometimes typical defects are illustrated on workcards, but workcards are often poorly integrated into the inspection task.
Low Observer Practice	Inspectors are highly skilled and practiced, after 3-10 years as an AMT before becoming an inspector. However, for some rare defects, even experienced inspectors may literally never have seen one in their working lifetime.
Sustained Attention on One Task	Inspectors may have some tasks where just one defect type is the target, but these are often interspersed with other tasks (e.g. different components) where different defects, often less rare defects, are the target.
Time Uncertainty	Defect occurrence is rarely predictable although inspectors often return to the same area of the same aircraft or engine and attempt to predict when defects are likely.
Spatial Uncertainty	While the actual occurrence of defects at specific places on specific components may be unpredictable, the inspector can have much useful information to guide the inspection process. Training, service bulletins and shared experiences can help point inspectors to specific locations where defects are more likely.
Low Feedback	Aircraft inspectors do not get good feedback, mainly because there is no easy way to find what truly is a signal, especially a missed signal. Feedback on missed defects only comes when one is found at a subsequent inspection, or when an operational incident occurs. Even feedback on false alarms is sporadic. Feedback of both Misses and False Alarms is at best severely delayed and therefore of little use to the inspector.
Unrealistic Expectations	For more common defects, expectations from training can translate relatively faithfully into practice. However, for very rare defects, expectation may still be unrealistically high after considerable practice.
Isolated Inspection Environment	The hangar and even the shop inspection environment are typically noisy, social and distracting. Both noise and social interaction and even some forms of distraction have been found to improve vigilance performance in laboratory tasks.

Table 7. Comparison between attributes of vigilance tasks and aircraft inspection tasks

Field Studies of Vigilance: In applying vigilance data and models to aviation inspection tasks, we should start with realistic inspection tasks and ask whether a vigilance decrement was observed. Later we can broaden our consideration to simulations of aircraft inspection tasks, and then to other non-aviation inspection tasks.

Two studies of eddy current inspection under realistic conditions measured the effects of time on task on inspection performance, and are relevant to visual inspection. Spencer and Schurman (1995)[15] used 45 experienced eddy-current inspectors (including a four two-inspector teams) at nine hangar worksites in USA. The task was to inspect 36 panels, each containing a row of 20 rivets, plus nine panels with a row of 75 rivets. These simulated B-737 fuselage lap splices, with a fairly high signal rate of one crack for about seven rivets. The task was self-paced and lasted about 4 hours. Inspectors took breaks as they needed, often after 30-120 minutes of task performance. There was no significant difference in either hit rate or false alarm rate between the first and second halves of the task. Murgatroyd, Worrall and Waites (1994)[43] simulated a similar eddy current task with about one crack per 150 rivets, using 12 experienced inspectors. Work was performed in either 30 minute or 90 minute periods for six days over all shifts. Hit rate was very high, over 99%, but no difference in either hit rate or false alarm rate was found between 30 minute and 90 minute inspection periods.

In contrast, two laboratory tasks simulating eddy current inspection of rivet rows with non-inspector participants both showed significant vigilance decrements. Thackray (1994)[46] and Gramopadhye (1992)[47] both simulated the lap splice task on a computer screen using 60 and 90 minute sessions. Small decrements (1% to 5% decrease in hit rate) between the first and second halves of the session were statistically significant.

Moving further from aviation, very few studies have measured the effect of time on task on inspection performance. An early study of the inspection of chicken carcasses on a processing line (Chapman and Sinclair, 1975)[48] tested two experienced inspectors over 85 minute inspection sessions. There was a small warm-up effect over the first 25 minutes (hit rate increased from 67% to 70%) followed by a slow performance decline, reaching about 57% hit rate after 85 minutes. These differences were significant statistically. Fox (1977)[49] reports a study on inspection of rubber seals for automotive applications, using ten experienced inspectors for 30 min periods with a defect rate of about one per hundred seals. He found a 27% decrement in hit rate from the first to second 15 min period. This was decreased to a 18% decrement when "lively music" was played from the 15th to 20th minute of each sessions, again a statistically significant decrement. In a final study Hartley et al (1989)[50] measured the search for noxious weeds while flying over the Australian outback in a helicopter at about 5 mph. They compared detection performance for half-day and full day work periods, using ten experienced farmers, with significant results. For half day sessions, hit rates are about 90% in the mornings but only 54% in the afternoons. For full day sessions the equivalent figures were 48% and 14%.

Taken together, these studies all found some decrement in hit rate associated with increased time on task. Note that none measured false alarms so that we cannot classify the effects as changes in sensitivity or bias. Note also that the decrement periods ranged from 30 minutes to a whole day. Finally, note the wide range of decrements observed, from about 13% to 45%.

Conclusions on Vigilance Effects: From this review, we have established that reduced inspection hit rates can appear with time on task under some circumstances. Fortunately, for public safety, but unfortunately from the standpoint of valid conclusions, the evidence is at its worst for aviation inspection tasks. In a laboratory, a vigilance decrement is typically shown, but there are considerable differences between this task and aviation inspection (Table 1). In non-aviation inspection tasks, some quite far removed from aviation inspection, vigilance decrements have been reported. One should exercise caution in over-interpreting these studies, however, as it is notoriously easier to publish significant than non-significant results. Even in the small number of published practical inspection studies using experienced inspectors, there is no simple prescription for a maximum time on task. Comparisons have typically been between two periods (e.g. first *vs* second 15 minutes, morning *vs* afternoon) rather than a more detailed measure of the time course of any vigilance decrement.

It would be safest to assume that some vigilance decrement potentially exists for all inspection tasks, but that many conditions can prevent this from affecting actual performance. Such conditions can include good feedback, social interaction, high signal rate or accurate briefing on defect expectations. Where all conditions are unfavorable (e.g. Table 1) it would be prudent to limit the period of continuous inspection. Whether such a limit should be 5, 15, 30 or 60 minutes is still unclear from the published data, but 20 to 30 min appears appropriate where the consequences of error are so high.

7.2.2 The Visual Environment

In no inspection technology is lighting and the visual environment a more obvious factor than in visual inspection. In the FPI report,[7] the main visual environment issue was dark adaptation in the reading booth, but the issues for visual inspection are far broader, reflecting the broader scope of visual inspection as a conglomerate of tasks. It is a truism that without adequate lighting, a defect cannot be visually detected, but the converse is not true. Good lighting does not ensure good detection performance. No is there any guarantee that improving the lighting will improve performance in visual inspection. Thus it appears that adequate lighting, or visual environment more generally, is a necessary but not sufficient condition for effective visual inspection.

An example of the difficulties in making visual environment improvements in inspection, consider two studies. The first was not in the aviation field, but had many similarities. Schmidt, Schmidt and Jarabek (1976)[51] studied the inspection of alloy castings for visual defects, using two different lighting systems at four speeds of inspection. Figure 6 shows the results where the probabilities of the two correct decisions are plotted against time per casting for each lighting condition. There was a clear improvement under the revised lighting system, which consisted of adding a third fluorescent tube above the workplace and a yellow spotlight shining over the inspector’s shoulder to enhance the contrast of the defects. [Note also that the longer the time taken per item, the greater the hit rate, but also the greater the false alarm rate. This point will be discussed later in Section 7.2.4.]

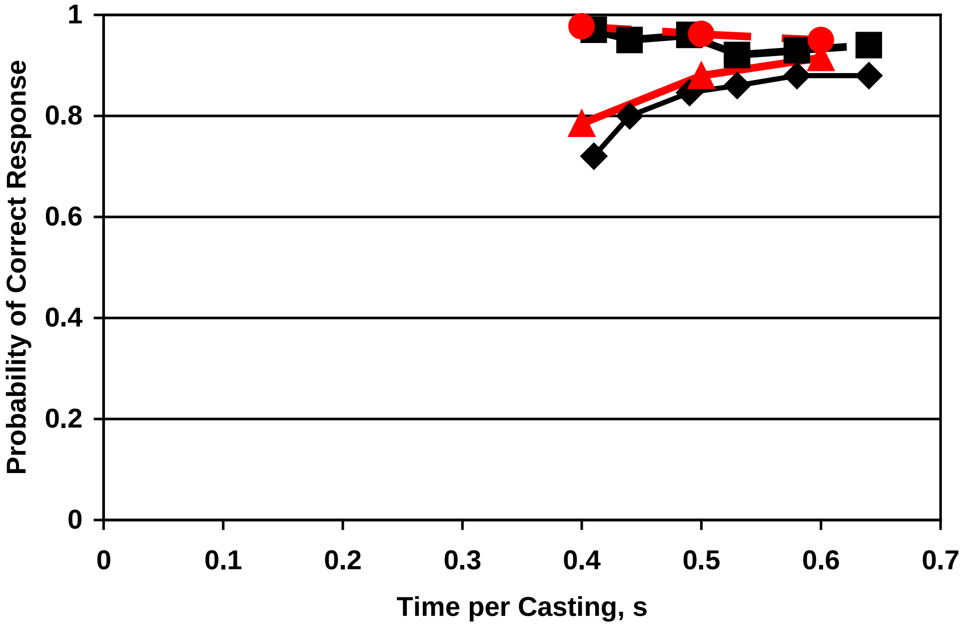


Figure 6. Effects of lighting changes on visual inspection of castings at different speeds. The upper curves represent [1-prob (False Alarm)] while the lower curves represent prob(hit). For each measure the "changed" condition is the higher line.

The second study was of inspection of rivet joints for small cracks, Drury and Spencer (1997).[52] A total of 12 mechanics and 12 inspectors inspected 4 panels of 20 rivets each under two different conditions of general lighting using two different flashlights. The lighting conditions were bright (900 Lux) and dim (90 Lux) measured at the inspection working point. The two flashlights were 3-cell Maglights® one with a normal glass lens and the other with a special diffusing lens known to give a smoother illumination across the flashlight beam (Shagram, 1995).[53] While the experiment was sufficiently sensitive to find a significant difference in performance between mechanics and inspectors, there were no significant lighting effects. Thus even an order of magnitude change in general light intensity and two visibly-different flashlight lenses did not change inspection performance. Thus, improving inspection performance through visual environment changes is not a simple task. As one human factors scientist remarked "There is more to visual inspection than meets the eye". Lighting, and the visual environment in general are covered extensively in reports of the FAA Office of Aviation Medicine, including Chapter 5 of the Human Factors Guide for Aviation Maintenance (1998)[35] and Chapter 6 of the Phase III report (Reynolds, Gramopadhye and Drury, 1993).[54] An older reference by Faulkner and Murphy (1975)[55] is also still relevant, but contains no performance validations of the lighting systems specified. Here we will concentrate on the material directly relevant to visual inspection.

A classification of lighting sources employed in aircraft inspection is as follows:

Ambient (requires no action by inspector)

General	Daylight (outside, through windows/doors) Area (from fixed luminaries)
Specialized	Built-in (e.g., cabin lights, cargo area lights)

Task (requires overt action by inspector)

Portable	Set up at an inspection site
Personal	Carried on the inspector's person, e.g., flashlight

The ambient lighting represents the minimum lighting level available in a task and also the minimum glare condition as it cannot be removed by the inspector, except for turning off the built-in lighting if this is convenient for other maintenance/inspection personnel. Task lighting represents the maximum lighting level, both from lighting devices set up to cover an inspection area, and from personally carried lighting. Note that to provide adequate lighting for any task it should be possible to reduce glare from ambient lighting and use the task lighting in a focused manner to illuminate the task without causing unnecessary glare.

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)[56] recommends that surface areas requiring visual inspection be provided with 750-1000 Lux of illumination. Vision can be improved by increasing the lighting level, but only up to a point, as the law of diminishing returns operates. 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 a maintenance organization.

According to IES (1987),[56] direct, focused lighting is the recommended general lighting for aircraft hangars. Inspection of aircraft takes place in an environment where specular reflection from airplane structures can cause glare so that low brightness luminaries should be installed. Often, additional task lighting will be necessary when internal work, or shadowed locations around the aircraft, result in low illumination levels.

From IES (1987)[56] pre/post maintenance and inspection operations (e.g., docking) require between 300 and 750 Lux. The choice of value within the range is task dependent. Generally, most maintenance tasks require between 750 and 1000 Lux, although more detailed maintenance tasks may require additional illumination. Inspection tasks require between 500 Lux and 2000 Lux of illumination, dependent upon the level of detail required in the inspection task. General line inspections (e.g., easily noticeable dents) may only require 500 Lux; however, most inspection tasks demand much higher levels. From the site observations of actual defects, it is apparent that many difficult inspection tasks may require illumination levels up to or exceeding 5000 Lux.

Based upon the current IES standards, it is recommended that the ambient light level in a maintenance hangar should be at least 750 Lux, in order to perform pre-post maintenance/ inspection operations and some general maintenance/inspection tasks without the necessity for additional task lighting. Furthermore, adequate illumination levels may be obtained in a majority of inspection tasks and many maintenance tasks through the utilization of task lighting.

Inspection Task	Recommended Illuminance, Lux
Pre/post maintenance and inspection	300-750
Maintenance	750-1000
Inspection	
Ordinary	500
Detailed	1000
Fine	2000

Table 8. Levels of Illumination Required in Aircraft Inspection/Maintenance (IES, 1987)[56]

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. The closer an inspectors direction of sight is to a glare source, the lower the visual performance. The effect of glare is to reduce the inspector's ability to discriminate detail, e.g., to differentiate between cracks and surface scratches. Thus, open hangar doors, roof lights, or even reflections of a white object such as the workcard can cause glare. Glare can also arise from reflections from the surrounding surfaces, and can be reduced by resorting to indirect lighting. Of particular concern is in inspecting partially-hidden areas (e.g., inside door panels), the lighting used to illuminate the defect may cause glare from the surrounding surfaces. The lighting system should be designed to minimize distracting, or disabling glare, using carefully designed combinations of area lighting and task lighting.

Reflectance: Every source reflects some portion of the light it receives as measured by the surface reflectance. High reflectance surfaces increase the effectiveness of luminaires and the directionality of the illumination. Thus for an aircraft hangar, it is important that the walls and floors are composed of reflective materials, or, for existing structures, are of high reflectance so that they help in reflecting light and distributing it uniformly. This can be achieved by having the floor and the walls painted a lighter color. This is more critical under the wings and fuselage where there may not be adequate lighting, due to aircraft shadows. Table 10 presents recommended surface reflective values, to assist in obtaining an adequate visual environment.

Surface	Reflectance
Ceiling	80 to 90%
Walls	40 to 60%
Equipment	25 to 45%
Floors	not less than 40%

Table 9. Recommended Reflective Values (Adapted from IES, 1987)[56]

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

Current Lighting Practice: In the Reynolds, Gramopadhye and Drury (1992)[54] report, the visual environment in a typical maintenance hangar was evaluated quantitatively. In addition, a number of inspection tasks were evaluated in a different facility. In the maintenance hangar, a grid of illumination measurements was used to quantify the ambient lighting. The mean illuminance was 560 Lux on open areas of the hangar, 25 Lux under the fuselage and 19 Lux under the wings. These values are lower than recommended by IES (1987).[56] The floor reflectance values varied around 8%, again far lower than the 40% recommended by IES (1987).[56] A series of inspection tasks were measured for both ambient and total (including personal) lighting. The results were that ambient illumination varied from less than 1 Lux for the air conditioning unit inspection to 440 Lux for the exterior of the fuselage nose area, with the other tasks ranging from 25 Lux to 120 Lux. With task lighting, usually from a 3 D-Cell flashlight, the illuminance was between 977 and 1800 Lux for 4 of the five tasks, but only 425 Lux for the nosewell task. Around each of these means there was a very high variability on different parts of the inspection task. Thus the visual environment for inspection makes interior tasks almost impossible without personal lighting sources, but with these sources it is adequate on average. . In general, we found an over reliance on the flashlight where portable task lighting would have been helpful.

The same study measured inspector perceptions of lighting needs to complement the environmental measurements. Fifty-one inspectors took part, rating lighting adequacy and the desired characteristics on lighting systems. From the latter set of responses, the main design requirements for lighting sources were:

Light output / brightness
Glare / brightness control
Distribution / focus / aiming
Weight / size of equipment
Battery Life

While light output had the greatest number of responses, the other factors to be considered when specifying lighting for each task. In fact, in the same study, a set of personal and portable lighting units were measured for their distribution of light output across the beam, and their characteristics on the factors listed above. A logical system of specifying lighting systems (ambient, portable, personal) based on task demands for illuminance, access etc. As noted at the beginning of this section, good glare-free lighting is important to inspection performance, but is certainly not the major factor, at least above some minimum level of adequacy.

7.2.3 Posture and Inspection Performance

A major characteristic of aviation inspection tasks is that they must be performed wherever the structure dictates, in contrast to industrial inspection tasks where an ergonomically-designed workplace is feasible. This characteristic results in many inspection tasks being performed in awkward postures. Posture can have adverse effects on the inspector, effects that may carry through to detection performance. For this reason we need to consider the posture of the inspector, particularly in light of the restricted spaces within which inspectors have to work. Note that we will use the generic term “Restricted Space” instead of the term “Confined Space” as that has already been used for a specific set of regulations by OSHA.

A study by Reynolds, Drury and Eberhardt (1993)[57] examined the effects of restricted spaces and postures in a set of inspection tasks on a DC-9 C-check. They found that over 20% of reported injuries were related to restricted spaces. Typical injury causation patterns were:

1. Repositioning in cramped or dirty spaces (e.g. fuel tank, tail interior) causes strains and sprains.
2. Head lacerations from walking in cabin or around fuselage exterior
3. Kneeling causes bruises of strains in the knees
4. Lifting equipment in restricted spaces results in back strain
5. Fall on stairs and access stands

Standard rating scales for discomfort and workload were used on four inspection tasks, with observations of the body posture’s difficulty using the OWAS posture analysis system. Results showed that there were many instances of extreme posture in all tasks. OWAS posture ratings at the three most urgent action categories accounted for between 42% and 60% of all postures on the four tasks. About 20% of the postures were in the two highest action categories, implying that a better posture should be sought “as soon as possible” or “immediately”. In terms of discomfort and fatigue, large differences were seen between the beginning and end of each task, particularly for the interior inspection of the tail area. This finding was corroborated by the workload scores on NASA Task Load Index (TLX). Overall TLX scores were very high, about 60 to 80 out of 100, with the main contributor being task physical demands. Clearly, postural problems with visual inspection are quantitatively large.

In terms of performance, inspection has been shown to be sensitive to postural issues. A recent summary of the evidence for this (Drury and Paquet, 2002)[58] found that not all tasks showed performance declines from poor postures, but of those that did, inspection tasks were well represented. For example, in a task of searching a visual field for defects in a laboratory test Bhatnager, Dury and Schiro (1985)[59] compared 3 postural conditions which were manipulated by 3 different display heights, and inspected circuit boards for 3 hours with 2 5-minute breaks per hour. As time-on-task increased during the tasks, participants tended to lean forward, change postures more frequently, report more discomfort, take more time inspecting circuit boards and make more errors when inspecting circuit boards. In a study closer to aviation visual inspection, Spencer and Schurman (1995)[15] measured the performance of 45 inspectors on eddy-current inspection of simulated lap-joints on frames that held the test specimens at two different heights and orientations. The top row was 1.63 m from the floor (approximately at shoulder height) and was roughly vertical. The bottom row was 0.61 m) from the floor (approximately knee level) and was at an angle of 35^o from vertical. The two larger specimens formed 4.3 meters of lap joint for inspection. Each of these larger specimens contained a single lap joint that was 1.20 m from the floor and near vertical. (See section 2.4 for a more detailed description of the specimens.) In this study, analysis on the detection of cracks sizes over 0.1inch showed no significant differences between the three task heights employed. However, for the full range of crack sizes contained in the study analysis of the PoD curves showed a significant influence of inspector position. This was a small effect, equivalent to reducing the size of the crack detected by 0.005inch at a POD of 50%, but could be operationally important in the early detection of marginal cracks.

A series of controlled studies by Mozrall, Drury, Sharit and Cerny (2000)[60] used an inspection task with body space highly restricted in vertical and two horizontal directions. As spaces became more severely restricted measures of discomfort, number of postural changes and stress (measured by variability of breathing rate) all became worse in a consistent pattern. However, there were minimal changes in performance accompanying this increased bodily stress. This was thought to be due to a limited task duration, of about 15 min in each condition, but a subsequent study by Wenner (1997)[61] found the same effect at longer durations. Wenner's study also made the inspection duration task limited (i.e. task must be completed) instead of time limited (i.e. inspect for 15 min). This also showed no effect, although we had expected participants to hurry through the task in the more severe conditions to minimize time spent in difficult postures.

To quote the conclusion from Drury and Paquet (2002)[58]:

> The field and laboratory studies do provide some evidence of a relationship between working postures and performance on tasks that are primarily cognitive in nature. The lack of consistent or dramatic postural effects on

> cognitive performance may be, in part, due to limitations in previous studies such as the evaluation of task performance over short time periods, which might not necessarily be predictive task performance maintained over a workday. In spite of the limitations, the literature demonstrates that, in some cases, body postures do have measurable effects on cognitive task performance.

To apply this to Good Practices in visual inspection, we need to first recognize that there may be postural limitations to performance. Three effects of posture on performance can be distinguished. One mechanism is direct mechanical interference: an inspection task will be affected if it is impossible to make the coordinated head, hand and body movements without interference from the structure of the aircraft or engine. A less obvious mechanism is that the demands on the postural system draw resources from the pool available for task completion. For example, the postural discomfort may be so distracting that tasks requiring complex cognitive processing (e.g. inspection) degrade from the interruptions caused by this distraction. More subtle mechanisms may come into play, such as reduced task motivation in protest against workplace discomfort, or even increased motivation to complete the task early and hence relieve the discomfort.

Having recognized postural effects, inspectors, managers and engineers should be trained to recognize bad postures rather than treating them as "business as usual". One way to do this is to imagine a robot performing the inspection task, and ask whether the awkward posture might upset the stability or the operating ability of the robot. Human inspectors should be given at least the same consideration as robots! An alternative way is to collect quantitative data, as was done in the Reynolds, Drury and Eberhardt (1993)[57] study using the OWAS coding (Kivi and Mattilla, 1991).[62] Having located bad postures, effective interventions are needed. Many of these involve better body support. Easily moved and adjusted seats or adjustable workstands are useful ways to support the body relative to the inspection work point. Note however that they will not be used if they are difficult to move around the whole inspection task. Support for standing tasks involves safe and secure platforms. Too many times we see inspectors on ladders or wobbly supports, having to spend at least some of their mental resources in worrying about balance and safety rather than inspection. Equipment that does not move easily to cover the inspection area encourages inspectors to reach sideways to inspect a larger area before the equipment has to be moved. This gives extreme postures. The good practice is to have rigid, easily controllable support equipment available to the inspector.

Other interventions are possible. Good lighting (see Section 7.2.2) allows the inspector to maintain a proper eye-to-task distance so that extreme postures are not required. Careful use of portable and personal lighting can alleviate postural stress while improving defect detection performance. In the same way, well-designed mirrors and loupes help the inspector maintain a more reasonable posture. Finally, one reason for awkward postures is the desire by the inspector not to get dirt, grease, etc. onto the body or clothes. We have observed inspectors unwilling to lean against a dirty element of structure or a grease-covered control pivot to avoid both spreading the dirt and having to change clothing too frequently. Good cleaning is not just a direct contributor to better defect visibility but also an indirect contributor by allowing a wider range of postures appropriate to the inspection task.

7.2.4 How Closely Should the Inspector Inspect: Effects of Speed on Performance

We have already seen in Figure 6 that the time spent inspecting an item affects the inspection accuracy. That figure, originally presented to illustrate the effects of lighting design, shows that more time per item leads to improved defect detection, but at a decreasing rate. There is a corresponding change in false alarm rates, but they *increase* as more time is spent on each item. In this section we take up the whole issue of the effect of inspection speed on inspection accuracy, providing a rationale for understanding the effects and a set of good practices based on this understanding. This section is based on work carried out for FAA/OAM (e.g. Drury and Gramopadhye, 1990)[58] and more detailed treatments of speed and accuracy beyond just inspection tasks (Drury, 1994).[59]

First we should note that the issue of inspection speed in one of great sensitivity to inspectors. They maintain, often vociferously, that they are there to inspect, not to inspect quickly. However, the issue is raised in other more subtle ways. When an area of structure is inspected, the inspector must make a decision on how closely to inspect. While not explicit, a decision to inspect more closely necessarily implies taking longer over the inspection. Thus, the confusion documented by Wenner (2000)[5] about what exactly constitutes “general” as opposed to “detailed” inspection is implicitly a question of speed as well as of accuracy. We shall return to this example at the end of the section.

As a framework for understanding the Speed/Accuracy Trade-Off (known in human factors engineering as SATO), first we need to note that not all tasks are sensitive to speed. Tasks where changing the speed has no affect on accuracy are known as “Data Limited” tasks. In contrast, there are other tasks where the accuracy goes down as speed goes up, known as “Resource Limited” tasks (Norman and Bobrow, 1985).[65] Data Limited tasks are those such as deciding whether a visible area of corrosion meets the standards for reporting or not. After a very brief interval, merely looking and thinking longer will not produce a better decision. A change of task, e.g. measuring with a ruler and comparing to written standards, can produce more accuracy for more time spent, but the initial visual decision making is essentially Data Limited. Many of the tasks we deal with throughout maintenance and inspection are Resource Limited and do show a SATO function, known as the Speed Accuracy Operating Characteristic (SAOC). Visual search is a prime example. As our earlier treatment showed, Section 4.2.1, the probability of target detection up to time (t) is:

$$p_t = 1 - \exp\left(-\frac{t}{\bar{t}}\right)$$

This is a definite SAOC, as p_t goes from 0.0 (target never detected) for $t = 0$ to 1.0 (perfect target detection) as t increases without limit. This can be illustrated in Figure 7, from Drury (1994)[59] which shows data collected from magnetic particle inspection in the steel industry. The data points have the following curve fitted to them, using the visual search SAOC defined above:

$$p_t = 1 - \exp\{(t - 1.33) / 1.45)$$

Here the mean time, $\bar{t}$, equals 1.45 min. Also note that the SAOC does not start at t = 0.0 but later at t = 1.33 min. What this means is that some time (1.33 min here) is not spent on visual search but on other activities, e.g. handling the steel bar, making decisions or marking defects. This is typical of search activities embedded in an inspection task, although the actual non-search time constant (1.33 min) will differ appreciably between tasks. Note that for the SAOC overall, there is a steep increase in p_t at short search times, but we get diminishing returns in terms of detection improvement for very long search times.

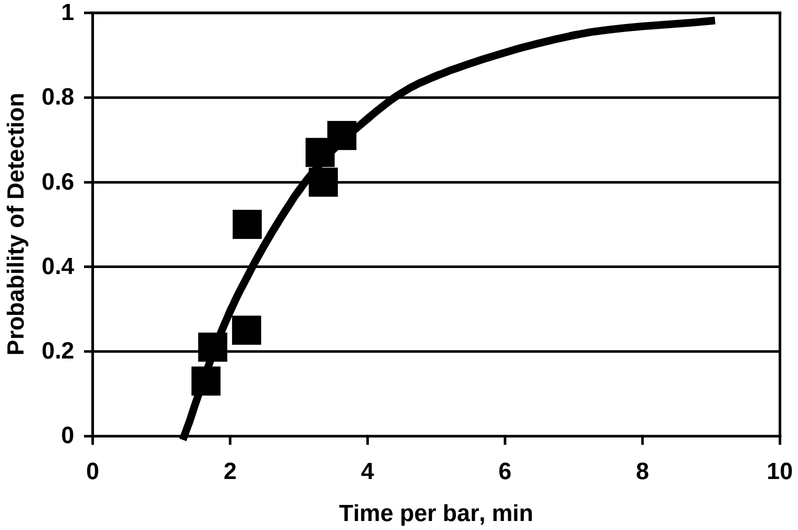


Figure 7. Speed-accuracy Trade-off (SATO) for magnetic particle inspection of steel bars.

From this example, we can demonstrate how improvements in the SAOC can occur, and use these as the basis for good practices. Improvement here means getting a higher p_t for the same search time t, i.e. moving to an SAOC higher on Figure 7. The four strategies for improvement are all shown in Figure 8 that uses the same initial SAOC curve as our magnetic particle task in Figure 7. The four strategies can be summarized as:

1. Reduce the fixed time, here 1.33 min
2. Reduce the search time constant, here 1.45 min
3. Choose an appropriate operating point on the SAOC
4. Choose a consistent SAOC operating point

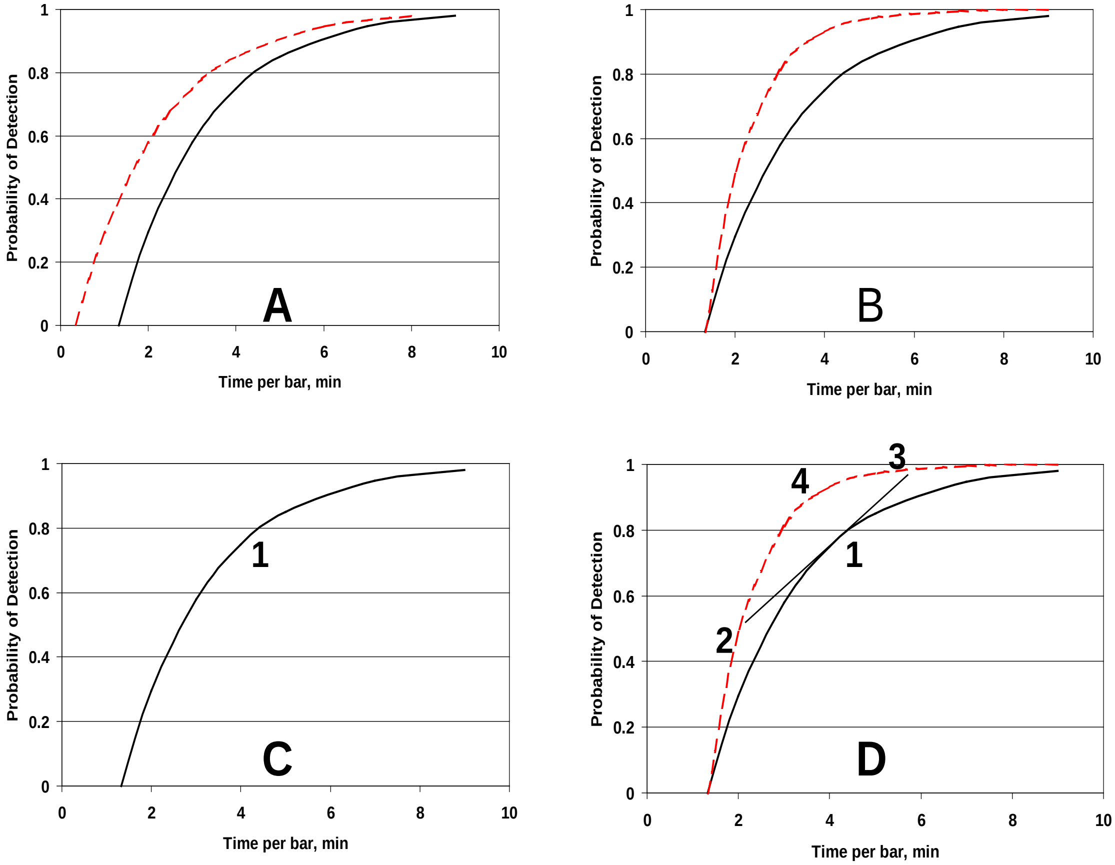


Figure 8. Four strategies for improving SATO in inspection.
A is reducing the fixed time; B is reducing the search time constant;
C is choosing the operating point and D is choosing a consistent operating point.

In Figure 8a, the fixed time has been reduced from 1.33 min to 0.33 min. Note that at short search times, e.g. 2 min, this has a large effect on p_t, but the effect on p_t becomes progressively less as search time increases, because of the diminishing returns in searching for very long times, e.g. 8 min. We can reduce this fixed time by concentrating on reducing the “overheads” of inspection such as setting up tools and equipment, marking of defects found, reference to documentation. Whatever we do to streamline such processes, e.g. by easily searchable electronic workcards, will help us to achieve greater detection accuracy in the same elapsed time.

For improving the search time constant, Figure 8b shows the effect of a reduction in $\bar{t}$ from 1.45 min to 1.00 min. Here the actual efficiency of search has improved, allowing more of the task area to be searched effectively in the same time. Examples of search improvements would be lighting, training or perhaps selection of inspector. Note that in Figure 8b, the main improvement is at medium to long search times, rather than short times as in overhead reduction.

Our third strategy, choosing an appropriate operating point on the SAOC is easy to do but may not be easy to implement effectively and may not be particularly effective. As shown in Figure 8c, there is a single best point to balance speed and accuracy to meet inspection objectives. This was defined in Section 4.2.1 as t* by considering the costs of time and error as well as the original SAOC. Such a strategy means that nothing is changed materially (e.g. the overhead or the search rate) but we are essentially doing the best we can with the current conditions. The way such a strategy is implemented is by specifying the “closeness” of inspection: should the inspector’s eyes be 0.5m from the area being searched (close) or should the inspection be from a standing position perhaps 1m or 2m away? This is the same argument in Wenner’s (2000)[5] work about how closely to inspect. The closer the inspection, the longer the search task will take, but the greater the chance of detecting any defects. Wenner’s study showed that standards for closeness of inspection are interpreted quite differently between different organizations, and even within a single organization. Better, i.e. more quantitative, definitions of such terms as “general” or “detailed” inspection are needed to ensure that all inspectors interpret their instructions in the same way.

This brings us to the final improvement strategy, consistent SAOC operating point. If the choice of operating point is allowed to vary, the overall result will always be worse than sticking to a single, consistent strategy. This can be seen on Figure 8d, where points 1 and 2 represent fast, inaccurate search and slow, accurate search respectively. If these two are mixed, as shown by point 3, the result is identical to working on a lower (= worse) SAOC curve. Thus a consistent strategy could have led to point 4, but an inconsistent one leave us effectively at point 3. As noted earlier, anything that helps a consistent choice of search speed will help by removing this SAOC operating point variability. A prime example would be more quantitative definitions to aid the inspector.

So far we have only used the example of visual search to illustrate the speed/accuracy trade off. It is an important task within inspection (see Table 1 in section 4.2) but it is by no means the only task. We need to consider each task briefly to see how SATO operates. The Initiate task has many Resource Limited components. As more time is devoted to assembling / calibrating equipment or reading workcards, the chance of missing a vital step is decreased. However, the SAOC will saturate at some point: there is only so long an inspector can spend reading a workcard and still obtain anything useful to improve accuracy. Improvements to the calibration procedures or the design of documents can help raise the accuracy of the Initiate task irrespective of the time taken to complete it.

Access tasks have a definite SATO, as was emphasized in the Borescope report (Drury, 2001).[8] The appropriate model here is for moving an object along a path without damage, e.g. a borescope head through an engine interior, or a cherry picker platform around an airframe structure. As the speed of a movement is increased, the probability of hitting a lateral obstruction increases, with obvious implications for damage. Improvements include increasing the controllability of the access equipment, e.g. by having controls move according to good human factors population stereotypes.

Decision, as was noted earlier, is subject to SATO, but the SAOC rapidly saturates. For any given decision task, this would be true after the first second or two of decision time, e.g. considering a potential crack visually. But decision is not limited to a singe sense, or to a single task strategy. Observing inspectors, it is clear that when the search locates a potential target (an indication), the whole task changes as the decision mode is entered. The inspector tries to collect alternative information to aid the decision, e.g. by moving a Maglight to illuminate the indication from several different angles, or by changing to a tactile mode and rubbing a finger across the surface including the indication. Other examples include the use of rulers, measuring loupes or templates to help judge the size of an indication. The inspector may even use knowledge of the aircraft structure and function to perform other tests to confirm the indication, for example inspecting the opposite side of a structural element to check whether the crack is visible from both sides. One of the best knowledge-based strategies the author has observed was by an inspector faced with a dent in the horizontal stabilizer leading edge. He looked at the position of the dent relative to the elevator trim tab, and used his knowledge of the airflow characteristics to predict that the airflow over the critical trim tab region would be disturbed. On the basis of this reasoning, he reported the dent as a defect for repair.

Because decisionmaking can have two distinct errors, there is the additional SATO parameter of False Alarms. We know from Section 4.2.1 that there is a trade off between the two decision errors (Misses and False Alarms), but the additional trade offs involving speed are likely to be small. One trade off that *is* important is a result of inspection having both search and decision components. As shown in Figure 6, used to illustrate lighting effects, not only does the detection rate increase with time spent inspecting, but the probability of declaring a defect-free item good decreases. This means there is an *increase* of false alarms with inspection time. Such a finding arises from the fact that the longer the time spent inspecting an area, the greater the chance that the search task will find an indication. Some of these indications will be true defects, while others will not. The more indications found that are not true defects, the more decisions that have to be made and thus, even if each decision has a constant False Alarm rate, there will be more false alarms in total. Thus there is an effective SATO effect on false alarms just because of the greater opportunity for making them as more time is taken on a particular area. In the extreme, if a very short amount of time is taken, the search will not reveal any indications, so that false alarms are impossible. But so too are detections!

Finally, the Respond task has some SATO effects. If insufficient time is taken to record defects, errors can arise in subsequent repairs. Inspectors are typically not verbose people, and tend to report defects with an economical use of words. This work in everybody’s favor provided that enough information is provided for unambiguous repair and buy-back. The characteristics of the defect (type, size, orientation) and its exact location must be provided if subsequent steps are to be error free. This takes time. Job aids, such as computer menus of defect types in computer-based NRRs, or digital photos of the defect, can save time while still improving accuracy.

We can now generalize the derivation of strategies beyond our visual search example to the other tasks of visual inspection. To summarize the human factors good practices for handling the speed accuracy tradeoff:

1. Reduce fixed or overhead times
2. Improve the task performance parameters
3. Choose the most advantageous operating point for SATO
4. Keep the operating point consistent between inspectors

The first two of these represent physical changes to the system (e.g. more usable job aids, better training) while the latter pair are essentially policy concerns. Only by using all as appropriate can we reach the point where enough time is spent for effective inspection to ensure that defects are detected at the performance level desired.

7.2.5 Training and Selection

It is always an objective of management to find “the best person for the job”, typically through selection of people who should be naturally suited to the job and by training of people in job-specific skills and knowledge. This is an example of the other side of human factors engineering: fitting the worker to the job instead of fitting the job to the worker. It can be effective, but needs to be carefully evaluated as a strategy. Selection and training are recurring costs whereas equipment and environment redesign are one-time costs. The history of most industries contains a long legacy of replacing recurring costs by making up-front investments, and human factors good practices must be treated in the same way. It is possible to use training to compensate for poor equipment and environment design, but this is (a) costly and (b) prone to errors when the inspector does what is natural rather than what is correct. Thus we must examine the evidence for training and selection carefully to show validity before advocating the use of these strategies as good practices.

Clearly, some training is needed for all inspection tasks: they are not natural occurrences in peoples’ every day lives and so must be trained for the specific defects, procedures and terminology before inspectors can even begin. But legitimate questions are: how should training take place, does further training improve performance, can training transfer between tasks and aircraft, and how does experience determine performance.

Training is conducted by all inspection functions in aviation, both initial training and periodic updating or retraining. A mixture of classroom, simulation, computer-based training (CBT) and on-the-job training (OJT) forms the typical training delivery system. Whole magazines are devoted to aviation training e.g. CAC. However, the issue of how to deliver the training is often seen as the issue, for example showing that CBT can be superior to ordinary classroom training. While this is certainly an issue (the author as even observed a classroom training where inspectors took turns in reading paragraphs from the manual!), a more immediate issue is content. What should be trained? Is it better to train both knowledge and skill for defect A, then repeat for defect B, or should the knowledge for both be trained together followed by the skill for both? Many such questions have been answered over the years in the human factors and experimental psychology literature, so we will concentrate here on finding best practices to use in visual inspection.

A number of quantitative studies of training and inspection performance in aviation have been undertaken to complement the typical measures of acceptability of the training to the participants. Building on what experimental psychology has learned about learning behavior, Czaja and Drury (1981a, 1981b)[66,67] showed how such training systems work for inspection. They were mainly interested in training older adults rather than the typical school and college ages, and so their results are directly applicable to aviation inspectors, who tend to be older than AMTs because of the experience requirement for inspection. They found that using performance feedback

These studies were used as the basis for the design of a training scheme for inspectors of jet engine roller bearings at the production factory by Kleiner and Drury (1993).[68] The scheme started with a detailed analysis of how expert inspectors perform the task, and why they behaved as they did. They showed that the whole job could be decomposed into logical components using the generic task description presented earlier. They then used a Progressive Part strategy to teach the knowledge and skill, with many job aids and simulations for practice throughout. Two novices were trained in this way and after two days of training were able to out perform inspectors of 15 years experience, even on this complex and demanding task. The Kleiner and Drury[68] study led directly to the aviation inspection training systems developed tested by Gramopadhye over several years (Gramopadhye, Drury and Prabhu, 1997[69]; Drury and Gramopadhye, 1992[70]).

In the Gramodadhye, Drury and Prabhu[69] study, the departure point was that OJT provided limited opportunities for training in inspection, as there is no control over which defects will be present on any particular aircraft or engine. Between knowledge instruction (from classroom or CBT) and OJT should come carefully developed simulations that allow control over which defects are present, where and when they are presented, and the form and frequency of feedback.

Selection: Firstly, the relationship between measured inspection performance and experience is not always apparent. In the study of rivet crack detection with 24 participants presented in Section 7.2.1, there was no relationship between experience and defect detection performance. Similarly, in the eddy current inspection study of Spencer and Schurman (1994)[15] there was no correlation between inspection performance and inspector experience. However, for the Visual Inspection research Program (VIRP) benchmark study found for the 12 inspectors studied, a significant correlation of 0.607 (p = 0.037) was found between detection of cracks and years in aviation. However, the same study found no significant relationship between the performance of the 12 inspectors on the eight different inspections tasks. Thus good performance on one task does not imply good performance on other tasks, suggesting that inspection performance may be task specific rather than a general ability. Outside the aviation field, there is mixed evidence for experience being related to defect detection performance. In a study of medical doctors reading chest X-rays, Kundel and La Follette (1972)[71] found that as training and experience progressed from medical students to staff radiologists, the time taken to detect an anomaly decreased significantly. However, Kleiner and Drury (1987)[72] showed that a short training program allowed novices to match the defect detection performance of aircraft bearing inspectors with many years of experience. Finally, Rubenstein, Nelson, Cherry and Duke (2002)[73] using a sample of 70 security screeners showed that experience (up to 24 months) did correlate with inspection performance, but mainly by reducing the false alarm rate rather than increasing PoD.

If experience of the inspector cannot always be relied upon to differentiate between good and less good inspectors, then what variables should we consider when using individual differences to select or place inspection personnel? Various approaches have been tried over the years in many industries and laboratories. Some have been tested with aviation inspectors. Individual differences between inspectors are typically large: Gallwey (1982)[74] reviews many studies where the differences between inspectors accounted for the largest amount of variability in the data. This was certainly true for the rivet crack detection task presented in Section 7.2.1. Despite the fact that detection scores on different tasks may not correlate well with each other, or with likely individual difference characteristics, large and consistent individual differences in performance have been reported. For example, where the task is the same from day to day and year to year, Moraal (1975)[75] and Salvendy and Seymour (1973)[76] report correlations between the performance of different inspectors across days and across years. Thus inspectors do have different performance levels and performance on the same task is likely to be consistent over time. These are the necessary prerequisites for finding good correlates of performance differences so that selection tests, or even extra training, may be an appropriate strategy.

Wiener (1975)[77] reviewed many of he early studies and concluded that personnel selection “offers little promise at the present”. Partly this was due to the early studies using invalid criterion measures such as supervisor ratings. These ratings were usually *not* based on defect detection or even false alarms, as these are notoriously difficult to measure. Little wonder then that supervisory ratings tended to produce a picture of the ideal inspector as a diligent and careful worker, prepared to stick to a job until it is complete, and willing to go the extra mile. However, these are exactly the qualities we are looking for in our bus drivers, police officers, medical doctors etc. When we move from supervisory ratings to actual defect detection performance, the results are not so clear cut.

Perhaps the first validated test to help select inspectors was the Harris Inspection Test (HIT), developed by Harris and Chaney (1969).[19] This was used for electronics workers and proved successful, correlating +0.55 with performance for 31 experienced inspectors. Later tests of the HIT on 84 participants in a laboratory inspection task found no correlation, however, suggesting that HIT may be specific to the electronics industry.

Gallwey (1982)[74] looked at several tests as correlates of inspection performance in a laboratory task using both students and experienced inspectors as participants. Incidentally, both groups showed similar performance. Gallwey found that different tests correlated with different aspects of performance, e.g. defect detection, speed, false alarms. He found that of the general tests, the Embedded Figures Test (EFT) had reasonable correlations with performance (+0.3 to +0.5), but that a simplified version of the real task gave the highest correlations. This simplified task approach has also been useful in selection aviation security screeners using the XISST test (Rubenstein, Nelson, Cherry and Duke, 2002).[73] The XISST was validated against on-line measures of screener performance, and found to correlate well with PoD and the Signal Detection Theory measure d′ (r = +0.5). That study also quoted earlier work showing that the cognitive test Hidden Patterns (similar to EFT) correlated at r = + 0.35 with on-line performance measures.

A more rational approach is to use the generic functions of inspection from Table 1 and see which aspects of individual differences correlate with performance on each function. We would not expect the same tests to correlate with Decision as with Search as both are different tasks requiring different abilities. For example, our model of search in Section 4.2 shows that performance depends on the size of the participant's visual lobe (*a* in our formulation). This prediction had been found correct by Eriksen (1990).[78] Wang and Drury (1989)[79] tested such predictions by measuring performance on several of the functions of inspection and devising appropriate tests of individual differences. Generally, search tasks correlated with different tests, but decision tasks tended to correlate together. The tests of individual differences did appear to be appropriate, although none of the correlations were very large, +0.3 to +0.5 being typical. Note that the usual measure of "inspector's eyesight", foveal visual acuity, has never been found to be significantly related to performance for experienced inspectors. This is perhaps because such acuity only comes into play when the indication has been found by the search process, and at this point other visual or non-visual aids can be used to assist the decision function.

This task analytic approach of specific tests being related to specific inspection functions was confirmed to some extent by the VIRP studies, where for the 24 inspectors on rivet cracks there was a significant correlation between search performance and visual lobe size of 0.61. Overall, it appears that there is some promise in selection tests for inspectors provided the task is very similar from day to day, and relatively low correlations are acceptable. The correlations approach +0.5, but this only explains 25% of the variability between inspectors. Where there is a sudden need for many inspectors, as in the current upgrading of aviation security inspectors, such a correlation can help focus attention on the better prospects, but at the cost of rejecting many good prospects and of accepting many who will never be top inspectors. At this time, there are better approaches to improving visual inspection performance, such as training or error-proofing of tasks as shown in our Good Practices.

7.2.6 Documentation Design

The development of standards and guidelines, e.g. ATA Spec 100[4] and the more recent ATA Spec 2100[80], etc. shows that the design of better documentation has long been a concern in aircraft inspection and maintenance. In a recent study (Wenner and Drury, 2000)[81] it was found that 46% of all errors in maintenance and inspection had "Documentation" as one contributing factor. But the guidelines for documentation design have typically been based on existing good practice and user opinion, rather than on documented changes in performance. This has led managers to dismiss calls for improved documentation as "giving in to Joe's preference rather than Fred's." It has also led to changes in just the documents that have caused recorded errors, rather than changing all documents, on the theory that there are just a "few bad apples" in an otherwise fine set of documents. Another favorite reason for inaction is that "we are getting a computer-based system next year, so we won't have documentation problems." The current section is an effort to replace such myths by quantitative data on the causes of documentation errors, and present a proven technique for measurable improvements in design.

A case study of one inspection document demonstrates this quantitative perspective. Drury (1998)[82] analyzed the effects on error rate of different features of a rather humidly-designed document to perform a fleet-wide inspection of a control activation system demanded by the FAA. The two-page document, with some additional diagrams, was used at many sites during an intensive four-night inspection of the whole of the airline's fleet of that aircraft type. There were seven numbered instructions requiring nine responses, plus some additional data to identify the tail number, etc. at the end. These ten responses were split into those five that violated a set of research-based guidelines (Patel, Drury and Lofgren, 1994)[83] and those five that met all the guidelines. The error rate for instructions violating the guidelines was 1.4%. The corresponding error rate for instructions complying with the guidelines was zero, i.e. no errors were made. *All* of the errors were thus foreseeable and therefore presentable. Urgency of working, poor lighting conditions at night, outdoor working, shift work and even individual complacency could perhaps be cited as contributing factors, but the bald fact remains: the well designed instructions eliminated errors. Good design is <u>not</u> following the whims of particular users, it is designing with the use of quantitative findings. "Management by Fact" is one of Deming's[84] 14 Quality Axioms, and applies as much to aviation inspection tasks as it does to manufacturing industry.

How then to produce better document design? First, one needs to note that Documentation is only one part of the information needed and used by an inspector. Information, in turn, is only a part of the total knowledge available to be used (Yovits et al 1987).[85] Much knowledge does not get used, either because it is not known to the right people, or has never been collected and systematized, e.g. by knowledge engineering. Within the knowledge that is used, i.e. information, Drury and Prabhu (1996)[86] drew a distinction between:

1. **Directive information**. e.g. documents, direct instructions
2. **Feedforward information**, e.g. training material, recent experience specific to that task, hangar knowledge.
3. **Feedback information**, e.g. measured information given to inspectors about their inspection performance or strategy. (This is notoriously difficult to provide in inspection tasks.)

Here, our concern is with the fist of these, the work documentation available to the inspector (and used?) to accomplish the task. But we have to ensure that any documents we design are compatible with inspection training and do support feedback, e.g. even on paperwork error rates (Drury, Wenner and Murthy, 1997).[87]

Good documents are both accurate and usable (Chaparro and Groff, 2002).[88] Accuracy is of great concern in the industry as a wrong number or reference can lead directly to an error in inspection. Particular sources of errors are in effectivity, i.e. whether this inspection procedure applies to the configuration of this particular aircraft or engine. Many workcards handle effectivity by listing the tasks for all configurations with an effectivity list specifying which tail numbers have each configuration. In these days of ubiquitous computer systems such a clumsy approach is unnecessary. The computer system has the data to write a specific workcard for each tail number, eliminating one source of potential errors.

A second source of inaccuracy lies in procedures that cannot be performed as written. It is a sad fact that most procedures specified by manufacturers or operators have never been validated. *No matter how many high-ranking engineers have signed off the document as acceptable, unless the procedure has been verified by an inspector (or other actual user) on aircraft, the document is only an untested theory, not a practical job aid.* During the ETOPS certification of the B-777, about half of the procedures were validated, and about 11% of these were found to be in error (e.g. Hall, 2002).[89] Validation is perhaps the single most effective error prevention strategy in documentation design.

Given that the procedure is accurate and can be performed correctly, there is still the question of physical design and layout of the resulting document. Here, human factors findings can also help to ensure error free performance. Document usability is a matter of both acceptability and performance. Human factors design guidelines are based on both criteria, but mainly on performance. For example, use of UPPER CASE text increases reading time by 13% to 20% (Reynolds, 1984;[90] Tinker, 1963[82]). In 1992, Patel, Prabhu and Drury [92] compiled a set of data-supported good practices for document layout, design and wording. They tested these by comparing acceptability of workcards designed with and without these guidelines and found overwhelming support for the “new” designs.

In a follow-up study, still using acceptability as the criterion, a comparison was made between an existing workcard, the same workcard designed to Patel et al (1994)[83] guidelines and a hypertext-based computer version of the workcard (Drury, Patel and Prabhu, 2000).[93] There were significant improvements in 17 measures between the original workcard and the redesigned paper-based workcard. Between the original workcard and the computer-based workcard there were similarly 17 significantly improved measures. However, between the redesigned paper-based workcard and the computer-based workcard, the only significant acceptability differences were for “degree of interest” and “would like to use again.” On an overall measure of ease of use, 84% of the improvement of the computer-based workcard over the original workcard was also found for the redesigned paper-based workcard. Clearly, most of the benefits of computerization came from the redesign of the workcard itself rather than from the computer specific features. This is the justification for undertaking documentation redesign *before* moving to electronic documentation. Most of the rewards will come from having more usable documents even before the computer systems are installed. Note that computer-based systems can have many features not available in paper-based systems, such as instant access to up-to-date manuals and AD information, or automatic generation of NRRs. However, unless valid guidelines are followed for document design and layout, the result will merely be to preserve the current, often poorly designed, documents in electronic form. This leads to error propagation as well as to a complacent belief that if it is on a computer, it must be correct. Goldsby (2002)[94] refers to this as preserving existing documents “under glass.”

One issue in documentation design is sentence structure and wording. The AECMA Simplified English (SE) restricted language was developed by major airframe manufacturers to standardize and simplify the language used in maintenance and inspection documentation. It uses a standard syntax for all instructions, and an allowed list of words to describe actions and outcomes. In an extensive test of SE, Chervak, Drury and Oullette (1996)[95] measured comprehension of workcards produced with and without SE using a sample of 175 AMTs and inspectors. SE reduced comprehension errors significantly, especially for more complex workcards and for non-native English speakers. There was no “downside” to SE for native English speakers, as has been expected anecdotally among AMTs, inspectors and technical writers.

The Patel et al (1994)[83] guidelines plus SE have been codified into a Documentation Design Aid (DDA) by Drury and Sarac (1997).[96] This computer program operates in a window, illustrating each good practice by an example, and detailing why the practice should be followed. These explanations of “why” are based on published research showing performance improvements, for example the improvement due to used of upper and lower case quoted above. The DDA also includes a SE checker to determine whether a word is allowed in SE, or has a synonym in SE. The DDA was tested using six technical writers, who were able to make significant improvements to test documents within one hour of first trying the program. In a direct test of the effectiveness of documents designed using the DDA, Drury, Sarac and Driscoll (1998)[97] found substantial and statistically significant improvements in comprehension errors when the DDA was used.

A final test of the DDA was performed in conjunction with a study of repair station errors (Drury, Kritkausky and Wenner, 1999).[98] It had been feared that in repair stations, the frequent switching between different companies’ documentation systems would lead to increased errors. This was tested using two different carriers’ versions of two workcards based on two AD notices, plus a DDA version of both of the same notices. Thirty-six AMTs at a repair station and 18 at a large carrier’s maintenance base performed comprehension tests on the two tasks, using either the same version for each task or different versions. No comprehension error differences were found between same or different formats, or between repair station and air carrier personnel. However, the DDA designed versions of both workcards had error rates significantly lower than those of the two carriers. For one document, the error rates on comprehension test answers for the two carriers was 51% and 35%, but only 4% for the DDA (Figure 9).

By now it should be obvious that documents designed in accordance with proven guidelines reduce errors, and are more usable to inspectors. The question now is how to implement the needed changes without huge costs. Fortunately, rewriting in SE is not difficult or time consuming, and many of the guidelines can be incorporated directly into company templates for workcards and other documents. The most difficult part of the redesign exercise is to ensure that the task steps themselves represent the most appropriate may to accomplish the task. This can most conveniently be done by having inspectors complete a task using the existing workcard, and then immediately use a different colored non-operational copy of the workcard to suggest changes. In this way engineering approval for task changes can proceed in parallel with other redesign work on each workcard, adding very little time to the task accomplishment time and reducing memory errors for difficulties, errors and improvements. This technique also gives the users an obvious stake in their own job design, making broad acceptance more likely.

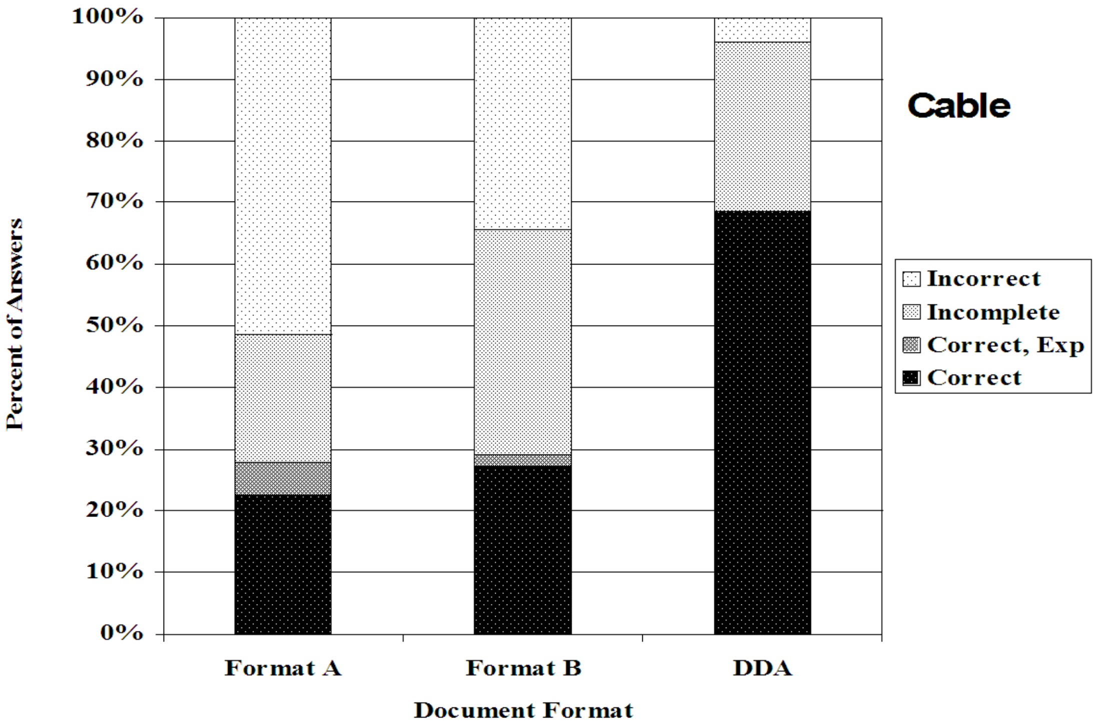


Figure 9. Accuracy of comprehension for three document designs (From Drury, Kritkausky and Wenner, 1999)[98]

8. CONCLUSIONS

This study has examined visual inspection as defined in various FAA and ATA documents, and as practiced in the hangar. A task analytic approach, used in previous studies of FPI[7] and borescope[8] inspection, was utilized to find the points in the seven functions comprising visual inspection where human capabilities were not well matched to task demands. These points became the key potential sources or error, and the task analytic approach allowed the derivation of human factors good practices appropriate to each potential source of error. Overall, 58 specific good practices were derived from both current best practices in the industry and human factors knowledge of error causation mechanisms. Based on these, six major areas were selected for in-depth discussion to allow readers to progress beyond the prescriptions of the specific good practices and understand the issues likely to confront visual inspection in the future.

9.0 REFERENCES

1. Federal Aviation Administration (1997). *Visual Inspection for Aircraft*. Advisory Circular ACNO. 43-204.
2. Goranson, U.F. and Rogers, J.T. (1983). Elements of Damage Tolerance Verification, *12th Symposium of International Commercial Aeronautical Fatigue*, Toulouse.
3. McIntine and Moore (1993). *Handbook of Nondestructive Investigation*, American Society of Nondestructive Testing, Volume 8 on Visual Inspection.
4. Air Transport Administration (199x). ATA Specification 100
5. Wenner, C. (2000). *The Impact of Instructions on Performance and Search Strategies*, Unpublished PhD Dissertation, State University of New York at Buffalo, Dept of Industrial Engineering, Buffalo, NY.
6. *NDE Capabilities Data Book* (1997). Nondestructive Testing Information Analysis Center.
7. Drury, C. G. (1999). Human Factors Good Practices in Fluorescent Penetrant Inspection. ***Human Factors in Aviation Maintenance - Phase Nine,*** Progress Report, FAA/Human Factors in Aviation Maintenance, via http://hfskyway.faa.gov**.**
8. Drury, C. G. and Watson, J. (2000). Human Factors Good Practices in Borescope Inspection. FAA/Office of Aviation Medicine, Washington, D.C.
9. Rummel, W. D. (1998). Probability of Detection as a quantitative measure of nondestructive testing end-to-end process capabilities. *Materials Evaluation*, **56**, 29-35.
10. Hovey, P. W. and Berens, A. P. (1988). Statistical evaluation of NDE reliability in the aerospace industry. In D.O. Thompson and D. E. Chimenti (eds), *Review of progress in Quantitative Nondestructive Evaluation*, Plenum Press, *Vol 7B*, 1761-1768.
11. *Materials Evaluation* (2001). Volume 9.7, July 2001.
12. Spencer, F. W. (2001). Estimating probability of detection curves from regression data. *Materials Evaluation*, Volume 9.7, July 2001, 866-870.
13. Panhuise, V. E. (1989). Quantitative Nondestructive Evaluation - Introduction. *Metals Handbook*, **17,** 663-665.
14. Spencer, F., Schurman, D. and Drury, C. G. (1996). Visual inspection research project report on benchmark inspections, prepared for the FAA Aging Aircraft NDI Validation Center.
15. Spencer, F. and Schurman, D. (1995). Reliability Assessment at Airline Inspection Facilities. Volume III: Results of an Eddy Current Inspection Reliability Experiment, DOT/FAA/CT-92/12. Atlantic City, FAA Technical Center.
16. Drury, C. G. (1992). Inspection Performance. In G. Salvendy *(ed.)*, *Handbook of Industrial Engineering*. New York, John Wiley & Sons**,** 2282-2314.
17. Rummel, W. D., Hardy, G. L. and Cooper, T. D. (1989). Applications of NDE Reliability to Systems. *Metals Handbook*. **17,** 674-688.
18. Yang, J. N. and. Donath, R. C. (1983). *Improving NDE capability through multiple inspection with application to gas turbine engine disks*. Wright-

Patterson Air Force Base, OH 45433, Materials Laboratory, Air Force Wright Aeronautical Laboratories.

19. Harris, D. H. and Chaney, F. B. (1969). *Human Factors in Quality Assurance*. New York, John Wiley and Sons.

20. Drury, C. G. and Fox, J. G. (1975). *Human Reliability in Quality Control*. London, Taylor & Francis, Ltd.

21. Rasmussen, J. (1983). Skills, rules, knowledge: signals, signs and symbols and other distinctions in human performance models. *IEEE Transactions: Systems, Man and Cybernetics*, **SMC-13** (**3**), 257-267.

22. Morawski, T., Drury, C. G., and Karwan, M. H. (l980). Predicting Search Performance for Multiple Targets.*Human Factors***, 22.**6, 707-7l8.

23. Drury, C. G. and Hong, S.-K. (2001). Generalizing from single targt search to multiple target search. *Theoretical Issues in Ergonomic Science*. **1(4),** 303-314.

24. Hong, S.-K. and Drury, C. G. (2002 in press). Sensitivity and validity of visual search models for multiple targets. *Theoretical Issues in Ergonomic Science*, 1-26.

25. Morawski, T.B., Drury, C.G., and Karwan, M.H. (1992). The Optimum Speed of Visual Inspection Using a Random Search Strategy. *IIE Transactions*, **24**, 122-133.

26. Karwan, M., Morawski, T. B. and Drury, C. G. (1995). Optimum speed of visual inspection using a systematic search strategy. *IIE Transactions* (1995), **27,** 291-299.

27. Chi, C.-F. and Drury, C. G. (1998). Do people choose an optimal response criterion in an inspection task? IIE Transactions (1998), **30**, 257-266.

28. Bavejo, A., Drury, C. G., Karwan, M. and Malone, D. M. (1996). Derivation and test of an optimum overlapping-lobe model of visual search, *IEEE Transactions on systems, man, and cybernetics*, **28**, 161-168.

29. Chi, C.-F. and Drury, C. (2001). Limits to human optimization in inspection performance. *International Journal of Systems Science*, 32 (6), 689 – 701

30. Wang, M.-J. J., Lin, S.-C. and Drury, C. G. (1997). Training for strategy in visual search. *International Journal of Industrial Engineering*, 101-108.

31. Gramopadhye, A. K., Drury, C. G. and Sharit, J. (1997). Feedback strategies for visual search in airframe structural inspection. *Int. Journal of Industrial Ergonomics*, **19**(5), 333-344.

32. McNichol, D. (1972). A Primer of Signal Detection Theory, Sydney, Australia, Allen and Unwin.

33. Drury, C.G., and Addison, J.L. (l973). An Industrial Study of the Effects of Feedback and Fault Density on Inspection Performance. *Ergonomics***, l6**, l59-l69.

34. 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*, Human Factors and Ergonomics Society, 1181-1185.

35. Maddox, M. M. (ed) (1995). *The Human Factors Guide for Aviation Maintenance 3.0*, via www.air-transport.org.

36. Air Transport Association of America, Inc. (199x), Spec 113 Human Factors Program, Washington, D.C.
37. Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, **1,** 6-21.
38. Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors*, **16(4),** 339-353.
39. Huey, B. M. and Wickens, C. D. (1993). Workload Transition: Implications for Individual and Team Performance, The National Academy of Sciences.
40. Craig, A. (1985). Field studies of human inspection: The application of vigilance research. *Hours of Work: Temporal Factors in Work-Scheduling*. Chichester, John Wiley and Sons, 133-145.
41. Wickens, C. D. and J. G. Hollands (2000). *Engineering Psychology and H*uman *Performance*. NJ, Prentice Hall.
42. Parasuraman, R., Warm, J. S. and Dember, W. N. (1987). Vigilance: Taxonomy and utility. In L. S. Mark, J. S. Warm and R. L. Huston (Eds), *Ergonomics and human factors*, New York, Springer-Verlag, 11-31.
43. Warm, J. S. and W. N. Dember (1998). Tests of a vigilance taxonomy. In R. R. Hoffman, M. F. Sherrick and J. S. Warm (eds), *Viewing psychology as a whole: The integrative science of William N. Dember*. Wash, D.C., American Psychological Association.
44. Molloy, R. and Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors*, **38**, 311-322.
45. Murgatroyd, R. A., Worrall, G. M. and Waites, C. (1994). A Study of the Human Factors Influencing the Reliability of Aircraft Inspection, AEA/TSD/0173. Risley, AEA Technology.
46. Thackray, R. I. (1992). Correlates of Individual Differences in Nondestructive Inspection Performance, Phase I Report: Background, Statement of Needed Research, and Research Approach. Washington, D.C., FAA Office of Aviation Medicine.
47. Gramopadhye, A. (1992). Training for Visual Inspection, Unpublished M.S. Thesis, State University of New York at Buffalo, Buffalo, NY.
48. Chapman, D. E. and Sinclair, M. A. (1975). Ergonomics in inspection tasks in the food industry. In C. G. Drury and J. G. Fox (Eds), *Human Reliability in Quality Control*, London, Taylor & Francis, 231-252.
49. Fox, J. G. (1977). Quality control of coins, In H. G. Maule and J. S.Weiner (eds), *Case Studies in Ergonomics Practice*, London, Taylor & Francis, Vol. 1.
50. Hartley, L. R., Arnold, P. K., Hobryn, H. and MacLeod, C. (1989). Vigilance, Visual Search and Attention in an Agricultural Task. *Applied Ergonomics*, **20(1),** 9-16.
51. Schmidt, W., Schmidt, D. and Jarabek, T. (1976). Case Study #1: Chain Saw Castings. Internal course report, Department of Industrial Engineering, State University of New York at Buffalo, Buffalo, NY.

52. Drury, C. G. and Spencer, F. W. (1997). Human factors and the reliability of airframe visual inspection. *Proceedings of the 1997 SAE Airframe/Engine Maintenance & Repair Conference (AEMR'97),* August 1997.
53. Shagam, R. N. (1995). *Light Shaping Diffusers for Improved Visual Aircraft Inspection*, DOT/FAA/AR95/32, October 95.
54. Reynolds, J.L., Gramopadhye, A. and Drury, C.G. (1992). Design of the aircraft inspection/maintenance visual environment. *Proceedings of the Seventh Federal Aviation Administration Meeting on Human Factors Issues in Aircraft Maintenance and Inspection,* Washington, DC, Office of Aviation Medicine. 151-162.
55. Faulkner, T.W and Murphy, T.J. (1975). Lighting for difficult visual tasks. In C.G. Drury & J.G. Fox (eds), *Human Reliability in Quality Control*, New York, Taylor & Francis.
56. Illuminating Engineering Society (1987). IES Lighting Handbook, Application Volume, 8th Edition.
57. Reynolds, J. L., Drury, C. G. and Eberhardt, S. (1994). Effect of working postures in confined spaces. *Proceedings of the 8th FAA/OAM Meeting on Human Factors Issues in Aircraft Maintenance and Inspection,* Alexandria, Va, November 16-17, 1993, 139-158.
58. Drury, C. G. and Paquet, V. (2002 in press). Posture and performance, In N. Delleman, C. Haslegrave and D. Chaffin (eds), *Working Postures and Movements - Tools for Evaluation and Engineering,* New York, Taylor & Francis.
59. Bhatnager, V., Drury, C.G. and Schiro, S.G. (1985). Posture, postural discomfort, and performance, Human Factors, 27(2), 189-200.
60. Mozrall, J. R., Drury, C. G., Sharit, J. and Cerny, F. (2000). The effects of whole-body restriction on inspection performance. *Ergonomics*, **43(11),** 1805-1823.
61. Wenner, C. (1997). Effects of Restricted Space on Cognitive Task Performance, Unpublished M.S. Thesis, State University of New York at Buffalo, Buffalo, NY.
62. Kivi, P. and Mattilla, M. (1991). Analysis and improvement of work posture in the building industry: application of the computerized OWAS method. *Applied Ergonomics,* **22**, 43-48.
63. 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, 149-164.
64. Drury, C. G. (1994). The speed-accuracy trade-off in industry. *Ergonomics,* **37,** 747-763.
65. Norman, D. A. and Bobrow, D. J. (1979). On data limited and resource limited processes, Cognitive Psychology, 5, 44-64.
66. Czaja, S. J. and Drury, C. G. (1981a). Training Programs for Inspection, *Human Factors*, 23(4), 473-484.
67. Czaja, S. J. and Drury, C. G. (1981b). Aging and Pretraining in Industrial Inspection. *Human Factors,* **23.4**, 485-494.

68. Kleiner, B. M., and Drury, C. G. (1993). Design and Evaluation of an Inspection Training Program. *Applied Ergonomics,* **24(2),** 75-82.
69. Gramopadhye, A. K., Drury, C. G. and Prabhu, P. (1997**).** Training for aircraft visual inspection, *Human Factors and Ergonomics in Manufacturing*, Vol. (3), 171-196.
70. Drury, C. G., and Gramopadhye, A. (1992). Training for visual inspection: Controlled studies and field implications. *Proceedings of the Seventh Federal Aviation Administration Meeting on Human Factors Issues in Aircraft Maintenance and Inspection,* Washington, DC: Office of Aviation Medicine, 135-146.
71. Kundel, H. S. and Lafollette, B. S. (1972). Visual search patterns and experienced with radiological images. *Radiology*, **103**, 523-528.
72. Kleiner, B. M. and Drury, C. G. (1987). Sensitivity of human tactile inspection. *Human Factors,* **29 (1)**.
73. Rubinstein, J., Nelson, T., Cherry, J. and Duke, P. (2002). Development of an X-ray image screener selection test for aviation checkpoint security. *Proc. of Human Factors and Aviation Security Conference,* January 2002.
74. Gallwey, T. J. (1982). Selection tests for visual inspection on a multiple fault type task. *Ergonomics,* **25**(11), 1077-1092.
75. Moraal, J. (1975). The analysis of an inspection task in the steel industry. In C. G. Drury and J. G. Fox (eds), *Human Reliability in Quality Control,* 217-230.
76. Salvendy, G. and Seymour, J. W. D. (1973). *Prediction and Development of Industrial Work Performance*, John Wiley and Sons, NY.
77. Wiener, E. L. (1975). Individual and group differences in inspection. In C. G. Drury and J. G. Fox (eds), *Human Reliability in Quality Control*, London, Taylor & Francis**:** 19-30
78. Eriksen, C. W. (1990). Attentional search of the visual field. In D. Brogan (ed), *Visual Search,* London, Taylor & Francis**,** 3-19.
79. Wang, M. J. and Drury, C. G. (1989). A Method of Evaluating Inspector's Performance differences and Job Requirements. *Applied Ergonomics,* **20.3**, 181-190.
80. Air Transport Administration (2000?). ATA Spec 2110
81. Wenner, C. and Drury, C. G. (2000). Analyzing human error in aircraft ground damage incidents. *International Journal of Industrial Ergonomics,* **26(2),** 177-199.
82. Drury, C. G. (1998). Case study: error rates and paperwork design, *Applied Ergonomics,* **29(3),** 213-216.
83. Patel, S., Drury, C.G. and Lofgren, J. (1994). Design of workcards for aircraft inspection. *Applied Ergonomics 1994,* **25(5),** 283-293.
84. Deming, W. E. (1986). *Out of the Crisis*. Cambridge, Massachusetts Institute of Technology.
85. Yovits, M. C., R. Kleyle, De Korvin, A., Medsker, L. and Mascarenhas, M. (1987). The relationship between information and decision making and the effect on the reliability and failure of information systems. In J. A. Wise and

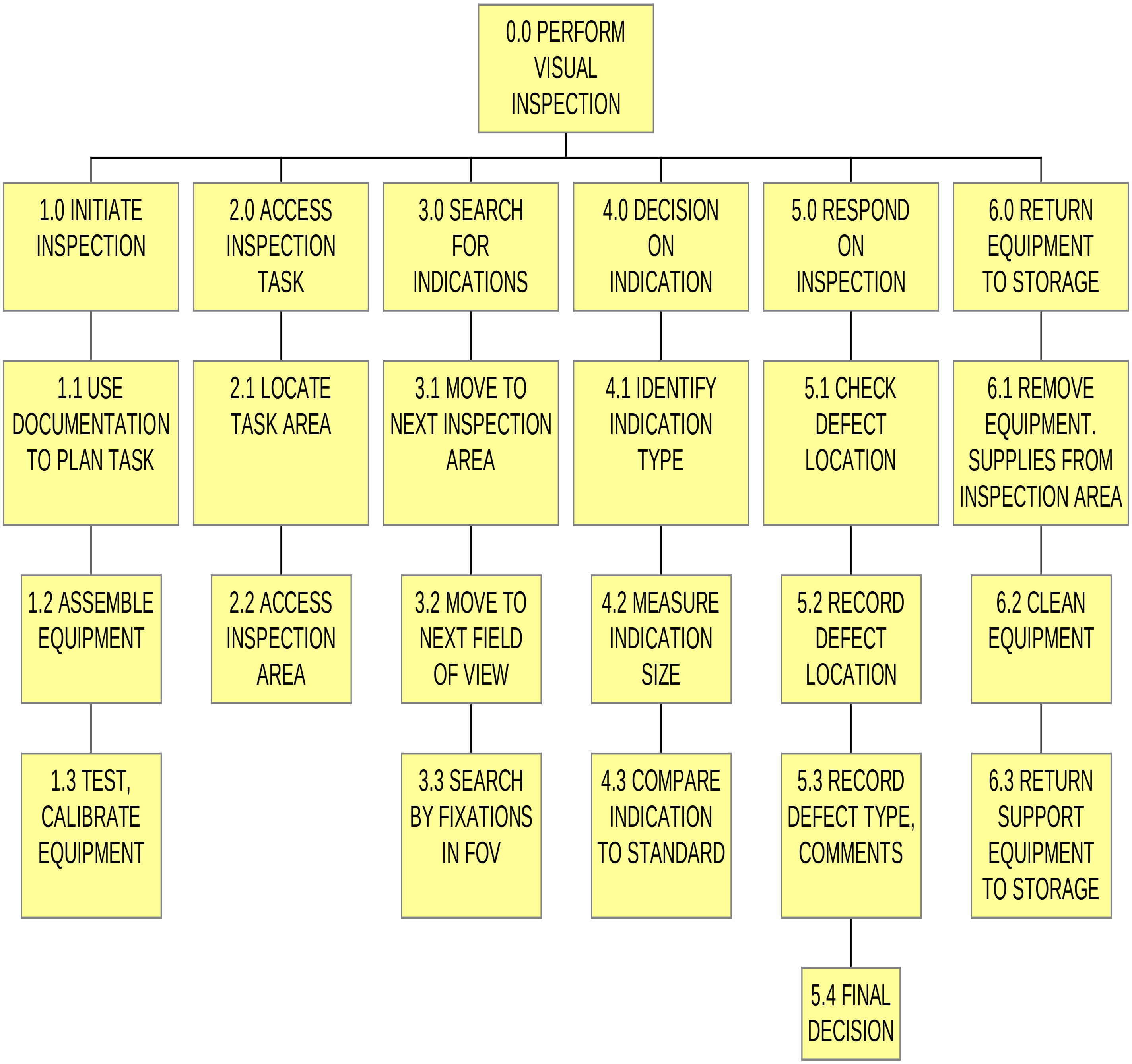
A. Debons (eds), *Information Systems: Failure Analysis,* Heidelberg, Springer-Verlag, 37-46.
86. Drury, C. G. and Prabhu, P. (1996). Information requirements of aircraft inspection: framework and analysis, *International Journal of Human-Computer Studies,* **45**, 679-695.
87. Drury, C. G., Wenner, C. L. and Murthy, M. (1997). A proactive error reporting system. *FAA/AAM 11th Meeting on Human Factors Issues in Aviation Maintenance and Inspection, SanDiego, CA,* 91-104.
88. Chaparro, A. and Groff, L. S. (2002). Human factors in aviation maintenance technical documents. *Proceedings of the 16th Human Factors in Aviation Maintenance Symposium,* April 2-4, 2002, San Francisco.
89. Hall, D. (2002). Maintenance documentation – a UK perspective. *Proceedings of the 16th Human Factors in Aviation Maintenance Symposium*, April 2-4, 2002, San Francisco.
90. Reynolds, L. (1984). The legibility of printed scientific and technical documents. *Information Design: The Design and Evaluation of Signs and Printed Material*, New York, J. Wiley & Sons, **2,**187-207.
91. Tinker, M. A. (1963). *Legibility of Print*. Ames, The Iowa State University Press.
92. Patel, S., Prabhu, P., and Drury, C. G. (1992). Design of work control cards. *Proceedings of the Seventh Federal Aviation Administration Meeting on Human Factors Issues in Aircraft Maintenance and Inspection,* Washington, D.C., Office of Aviation Medicine, 163-172.
93. Drury, C. G., Patel, S. C. and Prabhu, P. V. (2000). Relative advantage of portable computer-based workcards for aircraft inspection. *International Journal of Industrial Ergonomics*. The Netherlands: Elsevier, 26(2), 163-176.
94. Goldsby, R. P. (2002). Maintenance Documentation - Still on the improvement trail. *CAT Magazine*, Issue 2.
95. Chervak, S., Drury, C. G. and Ouellette, J. L. (1996). Simplified English for Aircraft Workcards. In *Proceedings of the Human Factors and Ergonomic Society 39th Annual Meeting 1996,* 303-307.
96. Drury, C. G. and Sarac, A. (1997). A design aid for improved documentation in aircraft maintenance: a precursor to training. *Proceedings of the 41st Annual Human Factors and Ergonomics Society Meeting,* Albuquerque, NM, 1158-1162.
97. Drury, C. G., Sarac, A. and Driscoll D. M. (1998). Evaluation of Documentation Format. ***Human Factors*** in Aviation Maintenance - Phase Eight, Progress Report, DOT/FAA/AM-97/xx**,** National Technical Information Service, Springfield, VA.
98. Drury, C. G., Kritkausky, K. and Wenner, C. (1999). Outsourcing Aviation Maintenance: Human Factors Implications, ***Proceedings of the Human Factors and Ergonomics Society 43th Annual Meeting,*** Houston, TX, 762-766.

10.0 ACRONYMS

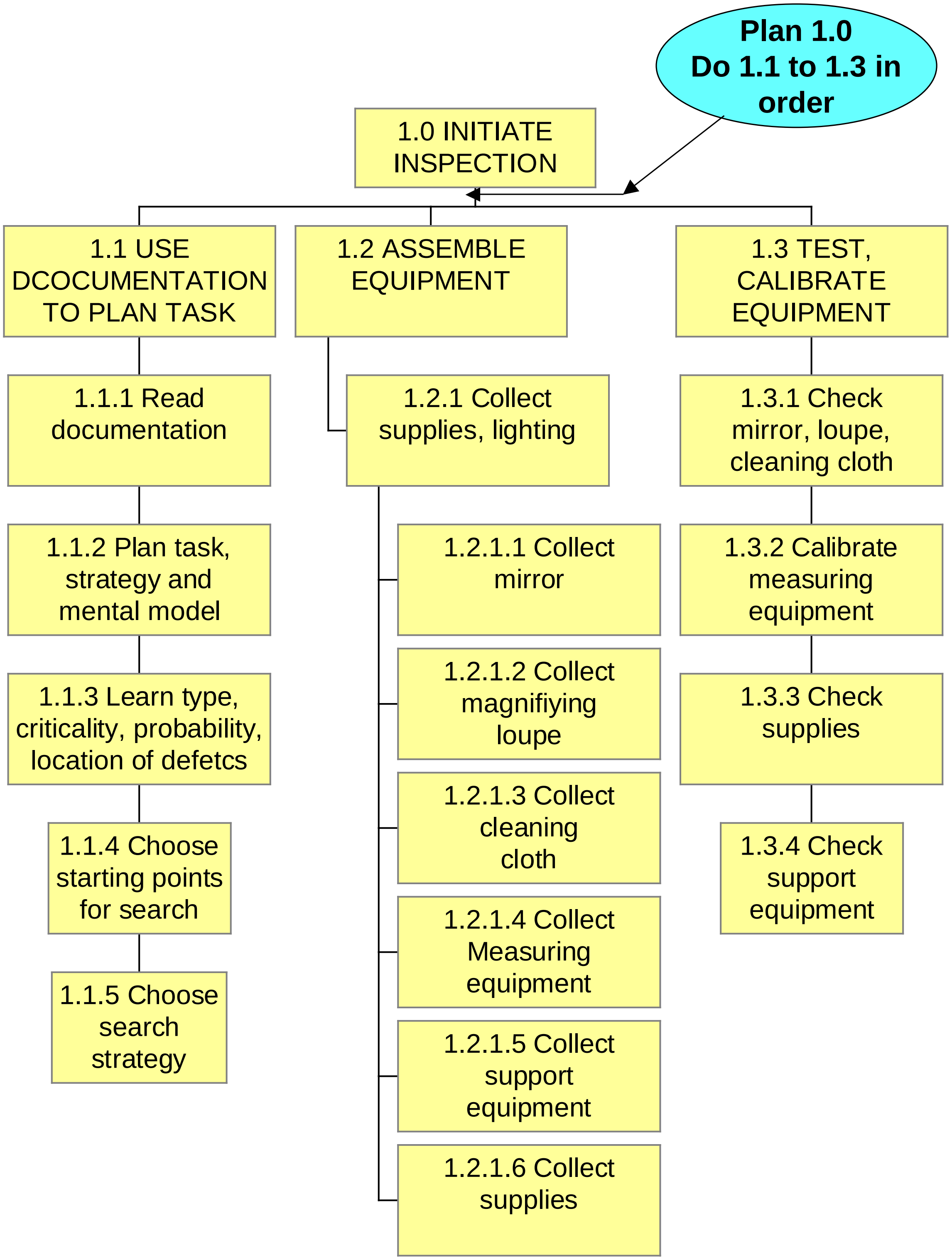
AAM	FAA's Office of Aviation Medicine
AC	Advisory circular
AD	
ASNT	American Society of Non-Destructive Testing
ATA	American Transport Administration
CASR	Center for Aviation Systems Reliability
CTSB	Canadian Transportation Safety Board
FAA	Federal Aviation Administration
FOV	Field of View
FPI	Fluorescent Penetrant Inspection
HCI	Human / Computer Interaction
HTA	Hierarchical Task Analysis
MSG-3	
NAD	Non-Aqueous Wet Developer
NTSB	National Transportation Safety Board
NDI	Nondestructive Inspection
NDE	Nondestructive Evaluation
PoD	Probability of Detection
PoFA	Probability of False Alarm
ROC	Relative Operating Characteristics
SATO	Speed/accuracy tradeoff
SDT	
SNL/AANC	Sandia National Laboratories

APPENDIX 1 -TASK DESCRIPTION AND TASK ANALYSIS OF EACH PROCESS IN VISUAL INSEPCTION

The overall process is presented first as a top-level key (same as Figure 1). Next, each of the six processes is presented in detail as an HTA diagram. Finally, each process is presented in the most detailed level as a Task Analysis table.



1.0 Initiate Inspection

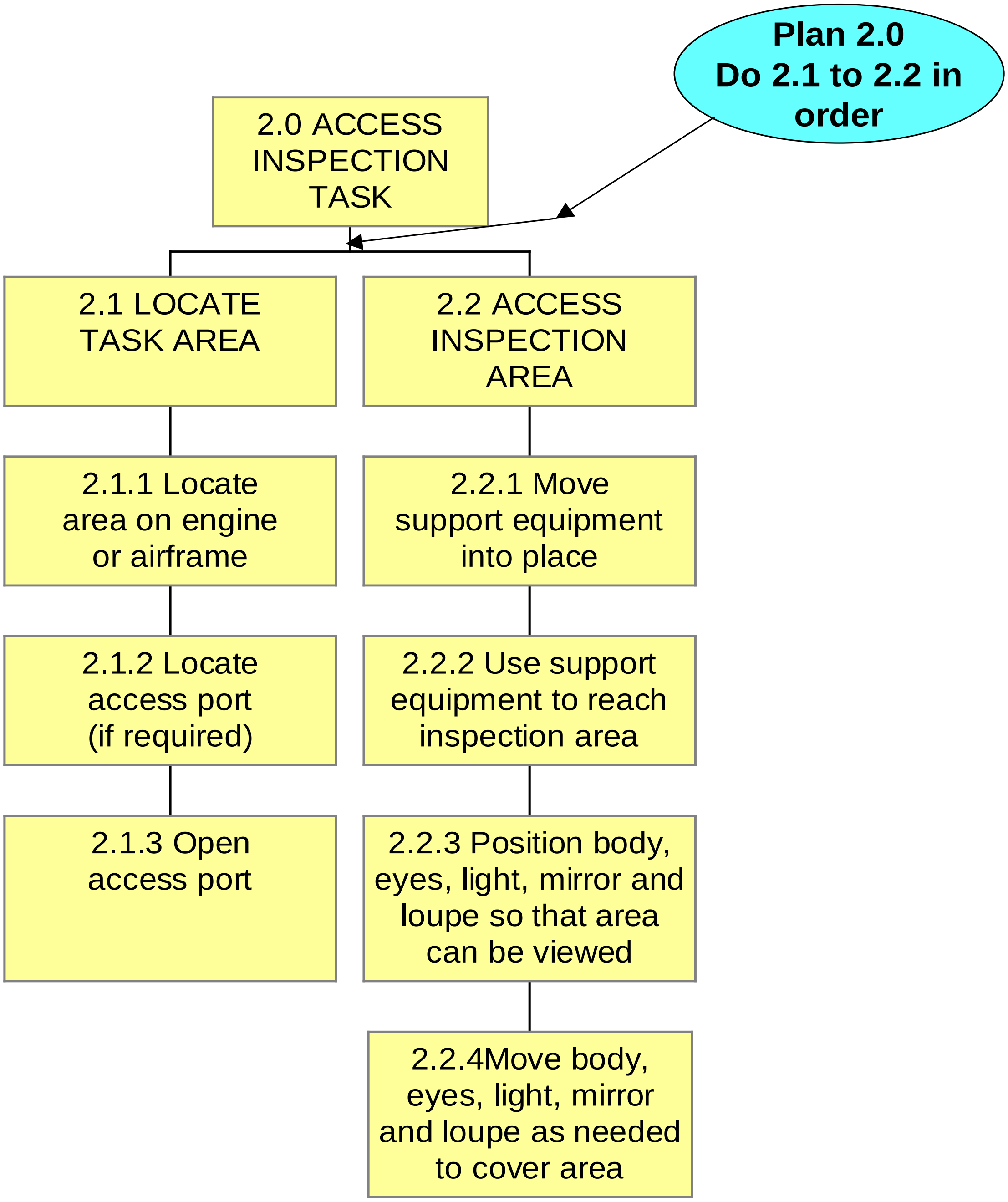


1.0 Initiate Inspection

	Task Description	Task Analysis
1.1 Use documentation to plan task	1.1.1 Read documentation on task, e.g. workcard	Is workcard available and current? Are there variances or EA's that modify the task? Is workcard self-contained or does it require access to manuals? Is terminology for areas, defects consistent between MM, workcard and hangar practice? Is workcard well human-engineered for layout, content, figures, ease of handling?
	1.1.2 Plan task for equipment setup and mental model of area to be inspected	Is there clear overview of whole task on workcard? Are the diagrams of the area to be inspected designed to allow for an accurate mental model of the structure? Does inspector have an accurate mental model of the structure where the task will be performed? Does workcard indicate mechanisms for defect occurrence that can help plan the task?
	1.1.3 Learn defects: types, criticality, probability, location, standards	Are all possible defects listed? For each defect type are criticality, probability and location listed? Are standards available in a form directly usable during visual inspection? How much does inspector trust information on workcard? Does workcard include all possible defects with information on how probable each type is in each location?
	1.1.4 - 5 Choose search strategy and starting point	Is inspection starting point specified in workcard? Is strategy (eg. Defect-by-defect vs. Area-by-area) specified in workcard? Does strategy specified fit the task from the inspectors viewpoint?
1.2 Assemble equipment	1.2.1 Collect supplies, lighting	Is there a realistic listing of tools, supplies? Can all equipment be used together,

1.2 Assemble equipment (continued)		e.g. mirror, light, measuring equipment?
	1.2.1.1- 4 Collect lighting, mirror, magnifying loupe, cleaning cloth, measuring equipment	Is kit available and complete for the task to be performed? Is mirror correct size? Is loupe of correct magnification for *this* task? Is power supply available for area lighting? Is cleaning cloth of approved type? Is measuring equipment, e.g. ruler, graticule, in same units as on workcard?
	1.2. 1 5 Collect support equipment	Is correct support equipment specified in workcard? Is correct support equipment readily available? Is non-approved support equipment more easily available?
	1.2.1.6 Collect supplies	Are only approved supplies (e.g. cleaning fluids) listed in workcard? Are approved supplies readily available? Are non-approved supplies more easily available?
1.3 Test, calibrate equipment	1.3.1 Check loupe, mirror, lighting, measuring equipment, cleaning cloth	Are all pieces of equipment functioning correctly? Are batteries in personal light adequate? Is area lighting clean and well-maintained? Is mirror clean, unbroken?
	1.3.2 Calibrate measuring equipment	Does test procedure include feedback for each step in a form appropriate to the inspector? Do inspectors have short-cuts, heuristics or informal recovery procedures to allow task to continue despite failure?
	1.3.3 Check supplies	Does cleaning fluid smell right for label?
	1.3.4 Check support equipment	Is support equipment safe and well-maintained?

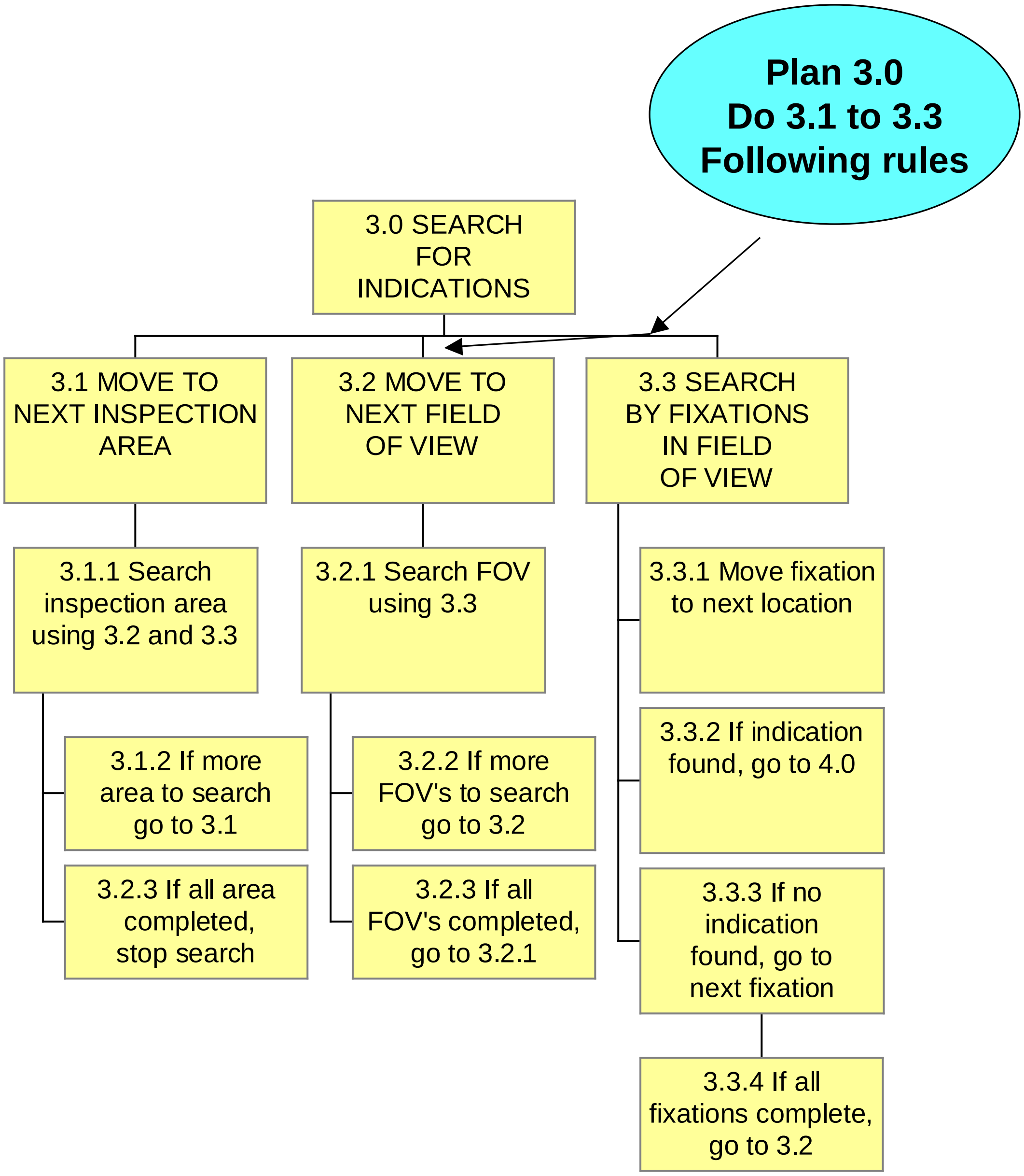
Errors/Variances: 1.0 Initiate Inspection
Documentation not available
Documentation not self-contained
Documentation not well-human-engineered
Documentation does not specify inspection strategy
Wrong, broken lighting used
Wrong, broken equipment used
Non-approved support equipment used
Non-approved supplies used
Measuring equipment not calibrated, or mis-calibrated


Plan 2.0
Do 2.1 to 2.2 in order
2.0 ACCESS INSPECTION TASK
2.1 LOCATE TASK AREA
2.2 ACCESS INSPECTION AREA
2.1.1 Locate area on engine or airframe
2.1.2 Locate access port (if required)
2.1.3 Open access port
2.2.1 Move support equipment into place
2.2.2 Use support equipment to reach inspection area
2.2.3 Position body, eyes, light, mirror and loupe so that area can be viewed
2.2.4Move body, eyes, light, mirror and loupe as needed to cover area

2.0 Access Inspection Task

	Task Description	Task Analysis
2.1 Locate task area	2.1.1 Locate correct area on airframe or engine	Does Inspector know aircraft numerical locations? Does documentation give clear landmarks on to help define boundaries of inspection task?
	2.1.2 Locate correct entry port	Does documentation view correspond to inspector's view? Is there visual confirmation that correct port has been selected?
	2.1.3 Locate access equipment	Is required equipment (e.g. ladders, stands, tables) specified in workcard? Is required equipment available for use? Do inspectors select substitute equipment if correct equipment not available?
2.2 Access inspection area	2.2.1 Move support equipment into place	Is access equipment safe for this task? Can support equipment be moved easily?
	2.2.2 Use support equipment to reach inspection area	Is access equipment adequate for task performance, e.g. tables/stands for holding equipment and accessories?
	2.2.3 Position body, eyes, light, mirror and loupe so that area can be viewed	Is area lighting adequate for task in terms of ability to manipulate, amount and direction of illumination? Is personal lighting adequate for task, in terms of ability to manipulate, amount and direction of illumination? Does support equipment allow a comfortable body position while viewing area? It an initial position possible where body, eyes, light, mirror and loupe can be set up to view area?
	2.2.4 Move body, eyes, light, mirror and loupe as needed to cover area	Can mirror, lighting and loupe be handled together easily? Can eyes be moved easily to cover area? Can lighting be moved easily to cover area? Can mirror be moved easily to cover area? Can loupe be moved easily to cover area? Does support equipment move when inspector changes position?

Errors/Variances: 2.0 Access Inspection Task
Wrong choice of area /access port
Missing access equipment
Inadequate access equipment
Inadequate body support during task
Poor posture for simultaneous manipulation and viewing
Difficulty handling light, mirror, loupe together to view area
Wrong inspection area limits chosen


Plan 3.0 Do 3.1 to 3.3 Following rules
3.0 SEARCH FOR INDICATIONS
3.1 MOVE TO NEXT INSPECTION AREA
3.2 MOVE TO NEXT FIELD OF VIEW
3.3 SEARCH BY FIXATIONS IN FIELD OF VIEW
3.1.1 Search inspection area using 3.2 and 3.3
3.1.2 If more area to search go to 3.1
3.2.3 If all area completed, stop search
3.2.1 Search FOV using 3.3
3.2.2 If more FOV's to search go to 3.2
3.2.3 If all FOV's completed, go to 3.2.1
3.3.1 Move fixation to next location
3.3.2 If indication found, go to 4.0
3.3.3 If no indication found, go to next fixation
3.3.4 If all fixations complete, go to 3.2

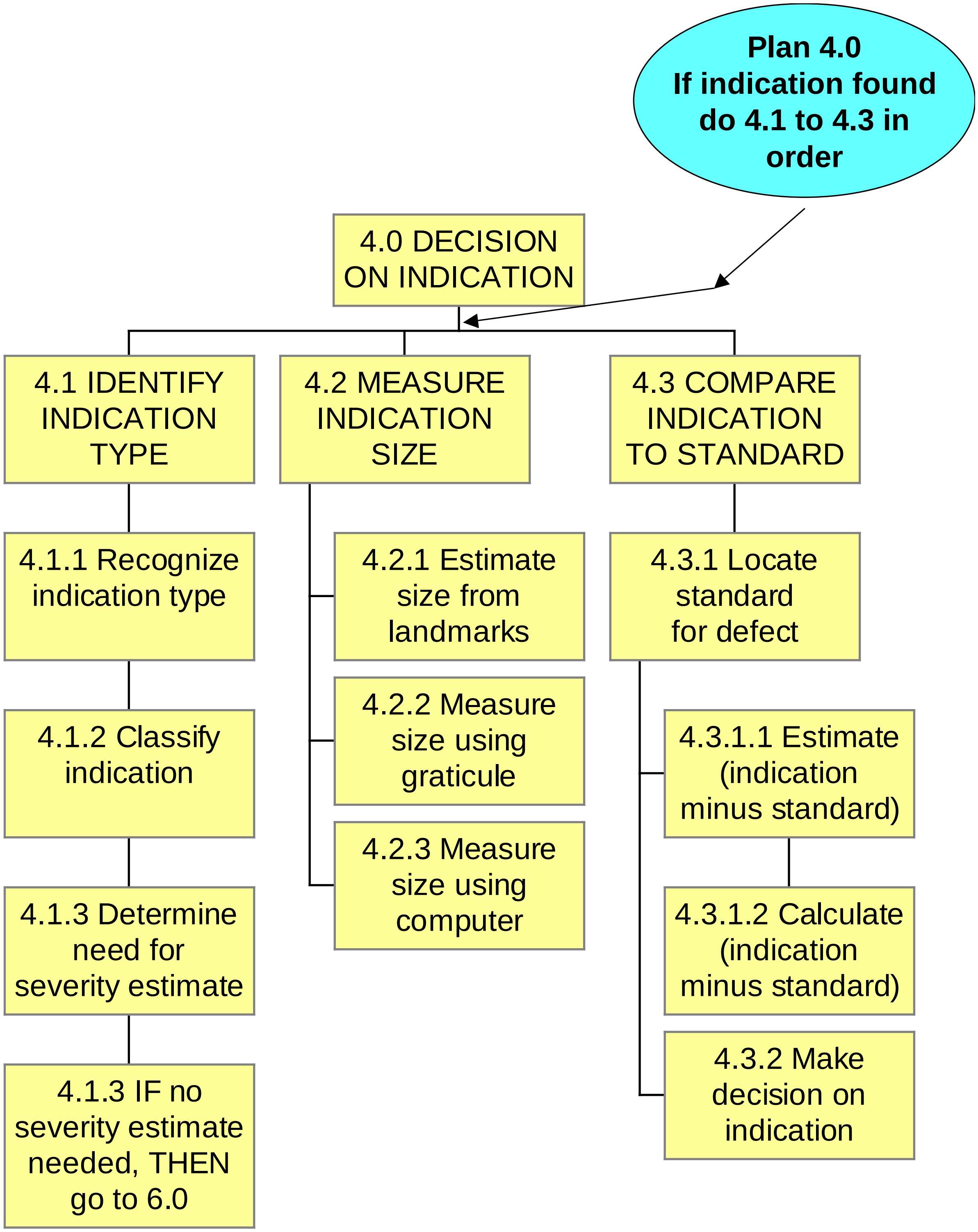
3.0 Search for Indications

	Task Description	Task Analysis
3.1 Move to next inspection area	3.1.1 Search inspection area using 3.2 and 3.3	Is area to be inspected remembered by inspector? What path (strategy) is followed by inspector to move FOV's over inspection area? Is search coverage complete? Is sufficient time allowed for reliable search for whole blade?
	3.1.2 If more areas to search, go to 3.1	
	3.2.3 If all area completed, stop search	
3.2 Move to next field of view (FOV)	3.2.1 Search FOV using 3.3	Is FOV movement needed to cover whole inspection area at adequate magnification? Can FOV be moved to all positions needed, e.g. mirror, lighting in correct positions? Can inspector maintain situational awareness as FOV moves? What is scan path followed by inspector? Does scan path cover complete FOV?
	3.2.2 If more FOVs to search, go to 3.2	
	3.2.3 If all FOVs completed, go to 3.2.1	
3.3 Search by fixations in FOV	3.3.1 Move fixation to next location	Does eye scan path across FOV cover whole FOV? Are fixations close enough together to detect indication if it is in the fixation? Is fixation time sufficient to detect a target? Is inspector expecting all possible indications each time search is performed? Are some indications expected in particular parts of the structure? Do inspector's expectations correspond to reality for this task? Does inspector return to area where possible indication perceived? Does inspector have high peripheral visual acuity? Is contrast between indication and background high? Is indication visible to inspector if an direct line of sight (Fovea)?
	3.3.2 If indication found, go to 5.0	Is there a clear protocol for what is an indication?

3.3 Search by fixations in FOV (continued)		Is there a clear protocol for remembering how much of search was completed before going to decision?
	3.3.3 If all fixations complete, go to 3.2	Does inspector remember whether fixations are complete? Is the policy to scan whole FOV once before stopping? Does inspector try to continue fixations for search while moving FOV?
	3.3.4 If no indication go to next fixation 3.3.1	

Errors/Variances: 3.0 Search for Indications
Incomplete search coverage by area, FOV or fixation
Incomplete coverage due to time limitations
Fixation movement too far to ensure reliable inspection
Loss of situational awareness by area or FOV or fixation
Loss of SA and coverage when finding indication stops search process

4.0 Decision on Indication

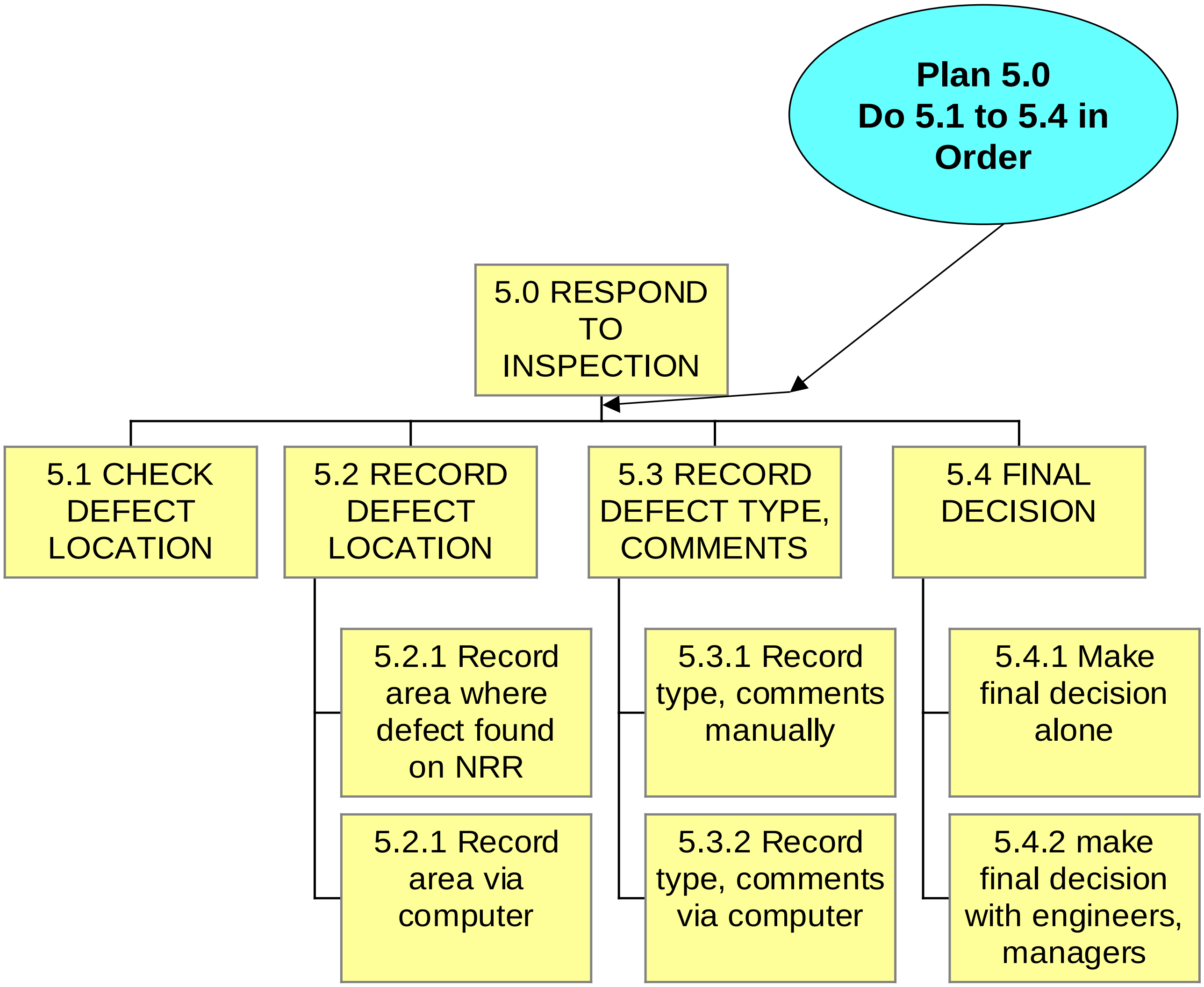


4.0 Decision on Indication

	Task Description	Task Analysis
4.1 Identify Indication Type	4.1.1 Recognize indication type	Does inspector have comprehensive list of possible indication types? Are some indication types under special scrutiny on this inspection? Does inspector have wide enough experience to be familiar with all indication types? Does inspector's view of indication correspond to prototypical indications in workcard? Is lighting of correct quality and quantity to ensure adequate recognition of indication?
	4.1.2 Classify indication	Are the correct terms for each indication type listed prominently in workcard? Are there local terms used by inspectors in place of official indication terms?
	4.1.3 Determine need for severity estimate	Does this class of indication need an estimate of size or severity or is any severity level rejectable?
	4.1.4 If no severity estimate needed, go to 6.0	
4.2 Measure indication size	4.2.1 Estimate indication size from landmarks	Are correct landmarks identified in workcard? Can inspector locate and recognize correct landmarks (e.g. structure, fasteners)? Are landmarks visible in same FOV as indication? Is there distance parallax between indication and landmark? Is there angular difference between indication and landmark? Does landmark correspond closely in size to indication? If not, can inspector make accurate judgments of relative magnitude between indication and landmarks? Does inspector have to remember size / severity or can it be entered immediately onto workcard?
	4.2.2 Measure size using graticule	Is measuring graticule available, e.g. as part of loupe? Can graticule be aligned with critical dimension(s) of indication? Is there distance parallax between indication and graticule?

4.2 Measure indication size (continued)		Is there angular difference between indication and graticule? Is numbering on graticule in a left-to-right direction? Are units on graticule the same as units specified in workcard for this indication? Does inspector have to remember graticule reading or can it be entered immediately onto workcard?
4.3 Compare indication to standard	4.3.1 Locate standard for defect	Is a standard specified on workcard? Are physical comparison standards available at inspection area?
	4.3.1.1 Estimate difference between indication and standard	Can standard be placed for direct comparison with indication? Can inspector make reliable estimate of difference between indication and standard?
	4.3.1.2 Calculate difference between indication and standard	Is measurement from 4.2 written down? Is judgment a simple >,< comparison? Can inspector calculate difference?
	4.3.2 Make decision on indication	Is decision based upon single indication or must multiple indications be evaluated before decision? Does inspector write down decision as soon as it is made?

Errors/Variances: 4.0 Decision on Indication
List of all possible indication types not available.
Inspector does not recognize indication type correctly.
Inspector uses wrong term to classify indication.
Measurement of indication size inaccurate.
Judgment of difference between indication and standard incorrect
Failure to record measurement size accurately.
Failure to record decision immediately

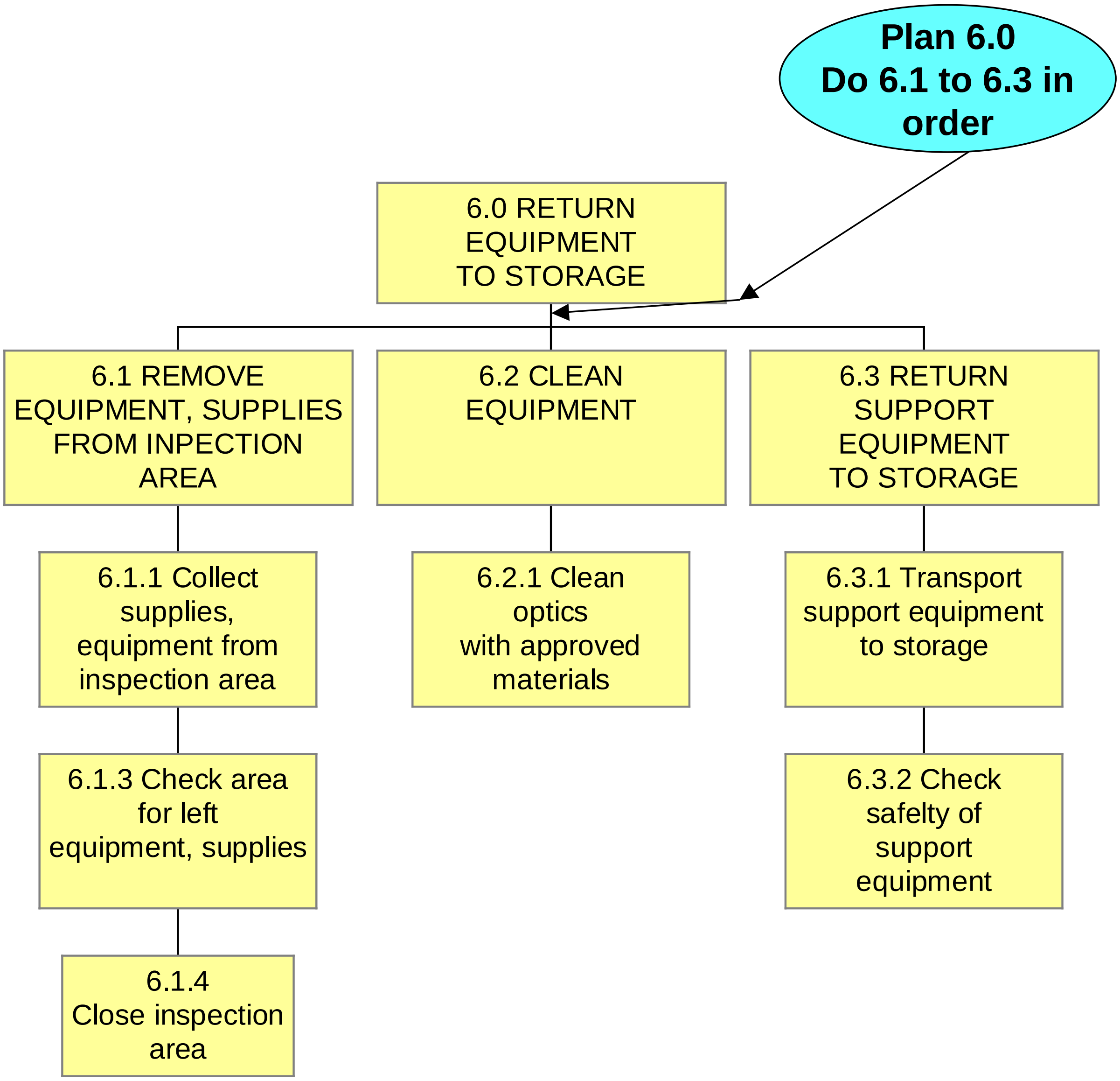

Plan 5.0
Do 5.1 to 5.4 in Order
5.0 RESPOND TO INSPECTION
5.1 CHECK DEFECT LOCATION
5.2 RECORD DEFECT LOCATION
5.3 RECORD DEFECT TYPE, COMMENTS
5.4 FINAL DECISION
5.2.1 Record area where defect found on NRR
5.2.1 Record area via computer
5.3.1 Record type, comments manually
5.3.2 Record type, comments via computer
5.4.1 Make final decision alone
5.4.2 make final decision with engineers, managers

5.0 Respond on Inspection

	Task Description	Task Analysis
5.1 Check Defect Location	5.1.1 Check defect location	Does location visually correspond to numerical location data (e.g. stations) on workcard? Has known reference mark been determined correctly?
5.2 Record Defect location	5.2.1 Record area where defect found on workcard	Should a workcard or an NRR be used for recording? Is workcard/NRR conveniently located with respect to the inspection site? Is there enough room on workcard /NRR to allow writing all defect locations?
	5.2.2 Record via computer	Is computer conveniently located with respect to the inspection site? Is computer program in correct mode for recording? Does computer program allow room for all defects to be recorded?
5.3 Record Defect type and comments	5.3.1 Record Defect Type and comments manually	Are defect types classified unambiguously? Is there a checklist of proper defect types? Is there room for comments on the workcard / NRR? Are inspectors encouraged to write sufficient comments for later use of data?
	5.3.2 Record defect comments via computer	Are defect types classified unambiguously? Is there a checklist of proper defect types? Is there room for comments on the computer program?
5.4 Final Decision	5.4.1 Make final decision alone	Was difference between indication and standard clearly beyond acceptance limit? Is there a clear record of the findings to back up the decision? Does inspector have to weigh consequences of lone decision, e.g. costs, schedule delays? Will managers typically stand by inspector in lone decision?
	5.4.2 Make final decision with engineers, managers	Does the procedure call for others to share the decision? Can engineers / managers be contacted with minimal delay? What happens if correct engineers / managers are not available for contact? Do engineers / managers display resentment at being contacted? Can facts be transmitted rapidly to engineers, managers, e.g. by engine, using documents / fax, sending computer files? Do engineers / managers respect inspector's

5.4 Final Decision continued		skills and decisions in coming to final decision? If inspector is overruled, what are consequences for future inspector performance?

Errors/Variances: 5.0 Respond to Inspection
Defect location not recorded
Defect type not recorded
Defect comments not recorded.
Defect location incorrectly recorded
Defect type incorrectly recorded
Defect comments incorrectly recorded.
Decision differences between inspector and engineers, managers


Plan 6.0
Do 6.1 to 6.3 in order
6.0 RETURN EQUIPMENT TO STORAGE
6.1 REMOVE EQUIPMENT, SUPPLIES FROM INPECTION AREA
6.2 CLEAN EQUIPMENT
6.3 RETURN SUPPORT EQUIPMENT TO STORAGE
6.1.1 Collect supplies, equipment from inspection area
6.1.3 Check area for left equipment, supplies
6.1.4 Close inspection area
6.2.1 Clean optics with approved materials
6.3.1 Transport support equipment to storage
6.3.2 Check safelty of support equipment

6.0 Return Equipment to Storage

	Task Description	Task Analysis
6.1 Remove equipment and supplies from inspection area	6.1.1 Remove supplies, equipment from inspection area	Is there a checklist of equipment and supplies to ensure nothing is left in inspection area?
	6.1.2 Check inspection area for left equipment, supplies	Is area where supplies and equipment could be placed easily visible when leaving area? Is there any other check on forgotten equipment and supplies, e.g. by another person?
	6.1.3 Close inspection area	Is correct closure of access port easily confirmed visually?
6.2 Clean equipment	6.2.1 Clean optics with approved materials	Are correct cleaning materials (cloths, solvents) available at workplace? Does inspector have training in correct cleaning procedures? Do inspectors have local work-arounds (informal and unsanctioned procedures) using easily-available materials? Can cleaning be accomplished without optical damage?
6.3 Return support equipment to storage	6.3.1 Transport support equipment to storage	Is correct location for access equipment known and available? Do personnel use "work arounds" (informal and unsanctioned procedures) if location not available? Is weight of support equipment low enough to be safety transportable? Does equipment have well-designed handling aids, e.g. Handles, wheels? Is there correct storage place for equipment? Is correct storage place available? Do inspectors have "work arounds" (informal and unsanctioned procedures) if storage place not available?
	6.3.2 Check safety of support equipment	Is there a procedure for safety check of equipment prior to storage? Is procedure always followed? What happens if support equipment is needed immediately on another job? Does it get signed in and out correctly?

Errors/Variances: 6.0 Return to Storage
Inspection access port not correctly closed
Supplies or equipment left in aircraft / engine
Support equipment not returned to correct storage
Equipment damage during cleaning
Support equipment not signed back into storage

APPENDIX 2 HUMAN FACTORS BEST PRACTICES FOR EACH PROCESS IN VISUAL INSPECTION

Process	Good Practice	Why?
1. Initiate	Design documentation to be self-contained	1. If multiple sources must be accessed, e.g. workcard, maintenance manual, this increases the probability that the inspector will rely on memory, thus increasing errors.
1. Initiate	Design documentation to follow validated guidelines, e.g. Documentation Design Aid (DDA).	1. Well-designed documentation has been proven to decrease comprehension errors 2. Application of validated guidelines ensures consistency across different inspection tasks, reducing errors. 3. Beware of reliance on all-inclusive terms such as “damage” or “general” inspection as inspectors are less consistent if trying to search for “everything”.
1. Initiate	Provide list of tools / supplies required, but indicate if it is the default list	1. Tools that are required, but are forgotten until at the inspection site, increase the chance the inspector will use a handy but incorrect substitute, increasing likelihood of inspection error or structure damage. 2. There is no need to repeat the standard tools / supplies list on every document. Establish a standard list and only call out changes from this list, to minimize the chance of the inspector skipping “boiler plate” paragraphs at the beginning of the documentation.
1. Initiate	Use documentation and training to help inspector form an appropriate mental model of the inspection task. E.g. provide diagrams showing the area to be inspected, using multiple angles if that increases clarity. Start with a view from the point of view of the inspector. E.g. Link new training and retraining directly to the documentation	1. The inspector should have an appropriate mental model of the structure being inspected. Particularly for complex structures, inspectors need to be able to recognize landmarks and map their documentation onto their view of the structure. This will allow the inspector to plan the task ahead, so that the task proceeds without surprises. 2. The inspector’s mental model includes failure mode knowledge, e.g. where maximum stresses occur and where defects are most likely. Documentation should be in accordance with this model, e.g. emphasizing the reasons behind probable defect locations, so that inspector can use knowledge-based

		reasoning. This helps the inspector to generalize acquired knowledge, making the inspection system more robust when faced with new threats.
1. Initiate	Define defect types, critical sizes and potential locations early in the documentation.	1. With good information on defects, inspectors can better plan their inspection task strategy. 2. If inspectors know the likely position and size of defects, they can better plan how to search, reducing the chance of missing defects.
1. Initiate	Include updated and explicit information on probable locations and likelihood of each defect.	1. The more complete the "feed-forward" information given to the inspector, the greater the probability of detection. This is the information inspectors seek from their peers, and will use if available.
1. Initiate	Ensure single terminology for structures and defects.	1. If inconsistent terminology is used, errors are likely to result at the Decision stage as the wrong standards for defect reporting may be used. 2. Terminology for defects and for standards can vary between organizations, between hangars and between inspectors in the same hangar. Inconsistent terminology can lead to inconsistent NRR wording, and potentially wrong repair or buy-back actions.
1. Initiate	Maintain configuration control by ensuring that the documentation applies exactly to *this* aircraft	1. Any variances or EA's should be incorporated into the documentation with visible indication in the document showing recent changes or modifications. This prevents inspection error arising from using (erroneous) memory of previous workcards. 2. If the documentation is too general, the inspector can become frustrated with trying to understand which parts apply to this aircraft, and revert to remembered versions.
1. Initiate	Recognize that inspectors' prior experience will influence their approach to each inspection task.	1. Inspectors develop their own expectancies of defect types, probabilities, severities and locations based on prior tasks and hangar wisdom" as well as from documentation. If these non-documentation sources are in error, defects can be missed. Use training and regular discussions of tasks to keep inspectors' expectancies in line with latest data.
1. Initiate	Ensure that standard equipment kit is available, in good condition and convenient to carry	1. Poor equipment can make detection and decision more difficult. Equipment can be forgotten or inconvenient to carry to inspection

		sites with difficult access. A standard inspection kit (personal lighting, mirror, cleaning rag, loupe, ruler, etc.) makes checking and transport of equipment easier, helping ensure that correct equipment is used.
1. Initiate	Ensure that measuring equipment is correctly calibrated and uses the same units as the documentation	1. If measuring equipment is not calibrated, or in different units from the documentation, errors will be made at the decision stage and defects not reported correctly.
2. Access	Ensure that inspector can locate correct inspection area on aircraft or engine	1. Make documentation compatible with task (see Initiate above). Include clear and unique landmarks and measurements to ensure that the correct area is inspected. Do not have implied but vague additions, such as "…and associated parts". Failure to inspect the correct area negates the inspection plan and can lead to missed defects.
2. Access	Is marking inspection area possible and/ or needed	1. Inspectors mark the area to be inspected to assists them in access and planning. Is this possible without structural compromise? Is it necessary? If possible and necessary, then provide non-damaging temporary markers for inspector use. Ensure that marking does not encourage leaving items in the inspection area.
2. Access	Specify correct access / support equipment in work documentation	1. If correct equipment is not specified, inspectors will be tempted to find an alternate "work arounds" (informal and unsanctioned procedures) so as not to delay the task. This can lead to poor working conditions and hence increased errors.
2. Access	Provide access equipment that facilitates ease of use E.g. support equipment should allow the inspector to stand or sit comfortably and safely while reaching the inspection area with all associated equipment (loupe, mirror, flashlight etc.)	1. Sub-optimal equipment leads to poor working postures and / or frequent body movements. Both can increase inspection errors. 2. Ensure that support equipment can be moved into place easily and moved precisely as the task progresses. If not easy to use, inspectors will be tempted to uses alternate equipment or not to move support equipment often enough. This can lead to unsafe overreaching, hence to incomplete inspection coverage or injury to inspector.
2. Access	Ensure direction-of-movement of cheery picker controls is directly compatible with bucket movement from the inspector's position.	1. Wrong direction-of-motion stereotypes can lead to damage to aircraft structures when an inspector moves a control in the wrong direction. Any uncertainly in proper control operation encourages less bucket movements, resulting in improper inspection coverage or injury to inspector.
2. Access	Keep support equipment available and well-maintained.	1. If equipment is not available, or the time taken to locate and procure the equipment is excessive, alternate non-approved equipment may be used,

		resulting in improper inspection coverage or injury to inspector. 2. If equipment is not maintained properly, alternate non-approved equipment may be used, resulting in improper inspection coverage or injury to inspector.
2. Access	Design access ports to reduce possibility of incorrect closure after inspection. E.g. fasteners that remain attached to the closure, tagging or red-flagging system, documentation procedure to show that port was opened and must be closed before return to service.	1. A common error in maintenance is failure to close after work is completed. Any interventions to reduce this possibility will reduce the error of failure to close.
2. Access	Ensure that equipment, such as mirror, lighting, loupe can be used together effectively.	1. If inspector cannot manipulate these tools together, then not all of them will be used. For example, if mirror, flashlight and loupe are needed for a closer examination of a potential defect. But if they cannot all be used together, then the flashlight may be propped in a non-optimum position while the other tools are used. This can result in missed defects or wrong decisions on reporting.
2. Access	Design loupe for direct viewing display to provide eye relief	1. High eye relief reduces the need to a rigidly fixed body posture for direct viewing. This in turn reduces the need for inspector movements required to provide relief from muscular fatigue. Such movements can result in incomplete search and hence missed defects
3. Search	Allow enough time for inspection of whole area	1. As shown in section 4.2.1, the time devoted to a search task determines the probability of detection of an indication. It is important for the inspector to allow enough time to complete FOV movement and eye scan over the whole area. When the inspector finds an indication, additional time will be needed for subsequent decision processes. If the indication turns out to be acceptable under the standards, then the remainder of the area must be searched just as diligently if missed indications are to be avoided.
3. Search	Provide clear instructions to inspector of expected intensity of inspection	1. The documentation should give the inspector enough information to provide a consistent choice of inspection intensity. Terms such as “general”, “area” and “detailed” may mean different things to different inspectors, despite ATA definitions. Well-understood instructions allow the inspector to make the intended balance between time taken and PoD. If the inspector looks too closely or not

		closely enough then PoD may not be that intended by the inspection plan.
3. Search	Inspector should take short breaks from continuous visual inspection every 20-30 minutes	1. Extended time-on-task in repetitive inspection tasks causes loss of vigilance (Section 4.2.1), which leads to reduced responding by the inspector. Indications are missed more frequently as time on task increases. A good practical time limit is 20-30 minutes. Time away from search need not be long, and can be spent on other non-visually-intensive tasks.
3. Search	If search uses a loupe, ensure that magnification of the loupe in inspection position is sufficient to detect limiting indications.	1. The effective magnification of the loupe depends upon the power of the optical elements and the distance between the lens and the surface being inspected. Choose a loupe magnification and lens-to-surface distance that ensures detection. This may mean moving the lens closer to the surface, thus decreasing the FOV and increasing the time spent on searching. The cost of time is trivial compared to the cost of missing a critical defect.
3. Search	Use combination of area, portable and personal lighting to make defects more detectable.	1. Area lighting from overhead luminaries and portable lighting, e.g. from floodlights, ensures that the inspection area is generally well-illuminated, but can cause glare from illuminated metal and glass structure. Glare reduces visual effectiveness dramatically, and can lead to missed defects. 2. Where hangar doors are open to sunlight, or even snow cover, glare can occur where this light source is within the inspector's visual field. Glare reduces visual effectiveness dramatically, and can lead to missed defects.
3. Search	Provide lighting that maximizes contrast between indication(s) and background.	1. The better the target / background contrast, the higher the probability of detection. Contrast is a function of the inherent brightness and color difference between target and background as well as the modeling effect produced by the lighting system. Lighting inside a structure mainly comes from the illumination provided by the personal lighting (flashlight), which is often directed along the line of sight. This reduces any modeling effect, potentially reducing target background contrast, so that lighting must be carefully designed to enhance contrast in other ways.

3. Search	Provide lighting that does not give hot spot in field of view	1. Hot spots occur where the lighting is not even across the FOV. This may be inevitable as light source to surface distance changes, but should be minimized by good lighting design. If a hot spot occurs, it can cause the eye to reduce pupil diameter, which in turn limits the eye's ability to see shadow detail. This effect can cause missed indications.
3. Search	Provide the inspector with approved tools to prevent tools being improvised.	1. Inspectors will improvise tools if the correct one is not available. For example, inspectors use a knife to check elasticity of elastomer seals, or use a rag that catches on frayed control wires to inspect for fraying. While these may be adequate, they have not been tested quantitatively. Wrong indications may result.
3. Search	Use a consistent and systematic FOV scan path	1. A good search strategy ensures complete coverage, preventing missed areas of inspection. 2. A consistent strategy will be better remembered from task to task, reducing memory errors. 3. Searching for all defects in one area then moving to the next (Area-by-Area search) is quicker than the alternative of searching for all areas for each type of defect in turn (Defect-by-Defect search), but the probability of detection is reduced. It may be difficult to help inspectors to work Defect-by-Defect.
3. Search	Use a consistent and systematic eye scan around each FOV	1. A good search strategy ensures complete coverage, preventing missed areas of inspection. 2. A consistent strategy will be better remembered from task to task, reducing memory errors.
3. Search	Do not overlap eye scanning and FOV or blade movement.	1. It is tempting to save inspection time by continuing eye scans while the FOV is being moved. There is no adverse effect if this time is used for re-checking areas already searched. But search performance decreases rapidly when the eyes or FOV are in motion, leading to decreased probability of detection if the area is being searched for the first time, rather than being re-checked.
3. Search	Provide memory aids for the set of defects being searched for.	1. Search performance deteriorates as the number of different indication types searched for is increased. Inspectors need a simple visual reminder of the possible defect types. A single-page laminated sheet can provide a one-page visual summary of defect types, readily available to inspectors whenever they take a break from the borescope task.

3. Search	Provide training on the range of defects possible, their expected locations and expected probabilities to guide search.	1. If inspectors know what defects to look for, how often to expect each defect, and where defects are likely to be located, they will have increased probability of detection. 2. If inspectors rely on these feed-forward data, they will miss defects of unexpected types, in unexpected locations, or unusual defects. Training and documentation should emphasize both the expected outcome of inspection and the potential existence of unusual conditions.
3. Search	When an indication is found, or the inspector is interrupted, ensure that inspector can return to exact point where search stopped.	1. Loss of situation awareness during blade rotation and after interruptions can lead to missed areas. With visual inspection it is possible to mark the current point in the search, e.g. with a pen or attached marker. Marking the search point reached when an interruption occurs will lead the inspector back to at least the current FOV.
3. Search	Remember that visual inspection often includes functional checks that are non-visual. Provide adequate job aids.	1. Not all visual inspection is visual. The odor of leaking solvents can alert the inspector to functional leaks. Looseness of fasteners may be checked by feel (haptic perception), particularly if the fastener is not readily accessible visually. In these cases, provide the inspector with approved procedures and training to ensure consistent inspection performance. 2. If functional checks require equipment, provide calibration and inspection procedures.
3. Search	Provide adequate structural cleaning	1. If the area is not cleaned well, defects may be hidden. In particular, some defects such as radius cracks, occur in structural positions that are difficult to clean fully. If the defect is hidden, its probability of detection decreases. 2. Over cleaning can remove indications of defects, such as leaks, leading to search errors.
3. Search	Some functional tests may have visual indications	1. Impaired movement of control runs may be visually indicated by paint rubbing at the point where the movement is impaired. Provide visual standards for later decision on reporting.
4. Decision	Ensure that inspector's experience with all defect types is broad enough to recognize them when they do not exactly match the prototypes illustrated	1. In recognition of a defect, inspectors use their experience and any guidance from the documentation. Illustrations show typical versions of a defect that may be different in appearance from the indication seen on the structure. Inspectors' experience should allow them to generalize reliably to any valid example of that defect type. In this way, defects will be correctly recognized and classified so that the correct standards are used for a decision. 2. Training programs need to assist the inspector

		in gaining such wide-ranging examples of each defect type. They should use multiple, realistic indications of each defect type to ensure reliable recognition.
4. Decision	Design lighting system to assist in defect recognition E.g. provide alternate lighting systems for search and decision.	1. The ideal lighting for recognition and classification may not be the ideal for visual search. Search requires contrast between indication and background, while recognition requires emphasizing the unique visual features of each defect type.
4. Decision	Help inspectors use other senses besides vision for accurate decision	1. Some decisions may need non-visual senses, e.g. touch, feel, thermal, as well as vision for a good decision. For example, a rivet may be found loose through touch, or a control run binding felt by roughness of movement. Ensure that these decision tests are supported by the documentation and training of the inspector. Failure to provide adequate non-visual standards will reduce decision accuracy.
4. Decision	Use consistent names for all defect types	1. Unless indications are correctly classified, the wrong standards can be applied. This can cause true defects not to be reported, and false alarms to disrupt operations unnecessarily.
4. Decision	Provide clear protocol for identifying landmarks used to judge defect size	1. If indication size is to be judged by reference to landmarks (not the most reliable system), then ensure that they are applied correctly. Providing a protocol in the documentation can assist the inspector in size estimation, reducing decision errors.
4. Decision	Provide direct means of measuring defect severity.	1. Often sizes are judged visually (“That’s about a quarter of an inch”) rather than measured. Measurement may be difficult due to structural interference, e.g. a 6”ruler may be impossible to place next to the defect. Alternatively, the measuring device may not be present at the inspection point, e.g. when an inspector had had to clean out his pockets in order to get his body into a fuel tank. The difficulty of exit to get the ruler and reentry may cause the inspector to conclude that a visual estimation is accurate enough not to affect the outcome. Errors can occur in visual estimation. 2. Providing readily usable tools for measuring lengths, angles and forces will help ensure they are used. But they need to be able to be easily transported together to the inspection site. Ensure that they are compatible with each other, and with the other tools the inspector must carry, to avoid inspectors using direct estimation of defect severity.

4. Decision	If ruler or graticule used to measure indication size, ensure that it can be used with minimal error E.g. Ruler / Graticule and indication are not separated causing parallax E.g. Indication and ruler / graticule have no angular foreshortening	1. Parallax and angular foreshortening can change apparent size relationships between indication and ruler / graticule scale. There are formulae for dealing with both, but if the indication and the landmark are in the same plane such formulae, and any associated errors, are eliminated.
4. Decision	Make it clear whether inspection is for first fault or all defects need to be located. If one defect causes a component to be replaced, then for good record keeping, all other defects on that component need to be found.	1. In many tasks the aim is to check whether the aircraft is fit for return to service or not. If a single critical defect is found, the aircraft cannot be returned to service, so that finding *all* the defects at that point becomes a secondary job. Inspectors need to be clear about whether they are supposed to find all of the defects, or stop after the first is found and delegate the finding of other defects to the subsequent stage of inspection. If both the inspector and those performing subsequent inspections think that it is not their task to find all the defects, then defects may be missed. The record keeping may also suffer if not all defects are reported, even if the component containing the defects is replaced. 2. Where the standard involves counting the number of defects, e.g. “not more than 3 areas of corrosion exceeding 5mm diameter”, provide the inspector a reliable means for keeping count of the number of defects found so far. Miscounting can cause missed reports, invalidating the inspection program.
4. Decision	Encourage knowledge based reasoning where appropriate	1. Not all of the inspection task can be captured by rigid rules. Unexpected defects may not be covered in the written instructions, e.g. a dead bird in a structure. Other defects may have different consequences depending on where they are located, e.g. a dent in a wing leading edge would cause harmful airflow breakaway if in front of a trim tab, but the consequences would be much less severe away from all control surfaces. Use training in why events occur and what are their consequences to help the inspector reason out the implications under unusual conditions. Encourage inspectors to share such experiences with management and peers.

4. Decision	Have clear guidelines on when use of adjacent structures are correct comparison standards.	1. Inspectors often use an adjacent "identical" structure as a comparison standard to help judge free play, warping, discoloration etc. At times this can be appropriate, but not always. Management needs to recognize that this happens and discuss guidelines in training and retraining to avoid wrong decisions.
5. Respond	Have a clear policy on what action to take when an indication does not meet defect reporting criteria,	1. Although the general wisdom among inspectors is to avoid writing down anything that does not have to be recorded, this can reduce overall inspection effectiveness by requiring subsequent searches to be successful. If ways can be found to record indications that do not yet meet defect criteria, then these can be tracked in subsequent inspections without having to search for them. Search unreliability is one of the major causes of missed defects in inspection.
5. Respond	Design a reporting system for defects that minimizes interruption of search process E.g. Use of stick-on markers in search, so that inspector can return to the correct point after interruption	1. Interruptions of the search process give the possibility of memory failure, hence re-starting the search in the wrong place, resulting in incomplete coverage and missed defects. Recording of findings is an interruption of search, so that keeping recording as rapid and easy as possible minimizes the chance of poor coverage.
5. Respond	Consider having the inspector return to all marked locations after search is complete.	1. For some tasks any break in search may lead to missing areas, e.g. in confined or awkward postures. In these cases, consider performing all of the search first, using stickers to mark indications until the end of search. When all the area is searched, the inspector can change modes to perform all of the decisions together, writing NRRs or not as appropriate for each sticker. This will both prevent search interruptions (reducing search failure) and make decisions more consistent for all being done together. 2. Use numbered stickers so that they are not left in the structure and can be counted out after inspection is complete. Stickers left in the structure may get loose and cause fouling of controls.
5. Respond	Reporting system should have sufficient space to describe defect type, location, severity and comments.	1. Inspectors have a tendency to be terse in their reporting, yet subsequent checking and repair depend on clear indications of defect type, location and severity. Consider the use of audio recording to amplify the information recorded on the workcard or NRR.

5. Respond	Automate paperwork where possible, but ensure flexibility	1. Writing NRRs by hand means that all common heading information must be entered repeatedly. This is an error-prone activity in itself and should be avoided by sensible automation. However, do not let computer code limitations force inspectors to act unnaturally, e.g. limited character lengths for word descriptions, lack of easy "undo' feature.
5. Respond	Provide a standard list of defect names and ensure that these names are used in defect reports.	2. Unless defect names are consistent, errors of severity judgment and even repair can arise. One technique is to use barcodes in the recording system for all defect types.
5. Respond	Have clear and enforced policy on when inspectors can make decisions alone and when others are needed to help the decision making.	1. Inspectors either make decisions on return to service / repair alone or with colleagues (engineers, managers). The requirements for choosing which decision mechanism is appropriate should be clearly communicated to the inspector and others. If not, there will be recriminations and loss of mutual trust when the decision made turned out to be incorrect.
5. Respond	If inspector makes decisions alone, consider the consequences if their decisions are later countermanded.	1. Inspectors, like all other people, need timely and correct feedback in their jobs if they are to make regular decisions effectively. They take feedback seriously, and will respond with changes in their own decision criteria. If a decision to change an component is countermanded, inspectors will tend (despite instructions and management assurances) to be more certain before calling for changes in future. Conversely, a decision to sign-off an component or area, if countermanded, may lead to tightened standards. If inspectors make the wrong decision, they need to be informed, but the effects of this feedback need to be considered.
5. Respond	Provide a means for rapid and effective sharing of information with other decision makers. E.g. Provide raw images of defects using a digital camera E.g. Provide two-way real time communications.	1. For the best possible shared decision making, there needs to be sharing of information. Modern digital cameras and computer based systems allow remote decision makers access to both the raw data, and two-way communications about the data and its implications. Two-way communications mean that remote decision makers can ask for new views or different lighting and receive the results rapidly. All of these enhancements can lead to more reliable decisions.
6. Return to storage	Design access ports to reduce possibility of incorrect closure after inspection.	1. A common error in maintenance is failure to close after work is completed. Any interventions to reduce this possibility will reduce the error of failure to close.

	E.g. fasteners that remain attached to the closure, tagging or red-flagging system, documentation procedure to show that port was opened and must be closed before return to service.	2. Ensure that procedures for close-up are adhered to, despite interruptions and time pressures, to prevent loss of closure errors.
6. Return to storage	Provide well-marked cleaning materials for cleaning optics and other tools.	1. Different materials, e.g. cloths or solvents, may be needed to cleaning optical surfaces and working surfaces. Materials need to be easily available and clearly marked if unauthorized substitutions are to be avoided. Relying on manufacturers labels is not enough. Labels specific to inspection can easily be printed and added, ensuring that tools are both cleaned and not damaged.
6. Return to storage	Provide reliable sign-in / sign out procedure for tools.	1. The signing in and out of tools should be as painless as possible or it will be violated sooner or later. The inspector may be under time pressure to start the inspection, or another inspector may be waiting for the equipment. Under such challenges, the simplicity of the procedures will determine their reliability.