SE9900184

SKI Report 99:8

Human Factors Aspects of Non-Destructive Testing in the Nuclear Power Context

A review of Research in the Field

Johan Enkvist
Anne Edland
Ola Svenson

February 1999

30 - 23

ISSN 1104-1374
ISRN SKI-R--99/8--SE

SKi

SKI Report 99:8

Human Factors Aspects of Non-Destructive Testing in the Nuclear Power Context

A review of Research in the Field

Johan Enkvist[1]
Anne Edland[1, 2]
Ola Svenson[1, 3]

[1] Stockholm University, Department of Psychology, SE-106 91 Stockholm, Sweden
[2] The Swedish Nuclear Power Inspectorate (SKI), SE-106 58 Stockholm, Sweden
[3] Netherlands Institute for Advanced Study in the Humanities and Social Sciences, NL-2242 PR Wassemaar, The Netherlands

February 1999

SKI Project Number 98127

Table of Contents

Summary

The present report reviews literature relevant to human factors and non-destructive testing. The purpose is to cover research that has been done, and to find out what still needs to be done to improve inspection performance. Methods of non-destructive testing (e.g., ultrasonics, eddy current) are complex diagnostic tools used by operators to inspect materials, e.g., components of a nuclear power plant. In order to maintain the integrity of a plant, recurrent inspections are made while the components are still in service. To control the quality of inspections, operators have to follow a procedure that determines what equipment to use and how to use it. The procedure also guides the operator in assessment of indications. There are a number of factors that can affect the inspection quality (e.g., heat, time pressure, and fear of radiation). In earlier studies, experience, organizational practices, and work conditions have been shown to affect on the quality of inspections. The quality of inspection performance is considered to benefit from adapting equipment and procedure to man's abilities and limitations. Furthermore, work conditions and feedback are considered determinants of performance quality. However, exactly how performance is affected by these factors, and the combined effect of them, need to be studied further. Further research is needed in decision criteria, procedure, and work conditions, and their affect on the quality of inspection performance.

Sammanfattning [Summary in Swedish]

Föreliggande rapport sammanfattar litteratur som är relevant för människa - teknik och oförstörande provning. Syftet med rapporten är att samla in tidigare forskningsresultat och dra slutsatser om ytterligare forskning som behövs för att öka kvaliteten på oförstörande provning. Metoder för oförstörande provning (t.ex., ultraljud, virvelström) är komplexa diagnostiska verktyg som operatörer inom process industri använder för att kontrollera material t.ex., komponenter i ett kärnkraftverk. För att säkerställa kärnkraftverkens integritet utförs återkommande provning av komponenter medan de är i drift. För att säkerställa dessa provningars kvalitet måste provarna följa en procedur som bestämmer vilken utrustning som skall användas och hur den skall användas. Dessutom hjälper proceduren provaren i bedömningen av indikationer. Det finns flera faktorer som kan påverka kvaliteten hos en provning (t.ex., värme, tidspress, oro för strålning, vakenhet). Tidigare forskning har visat att erfarenhet, organisatoriska faktorer och arbetsförhållanden påverkar provningskvaliteten. Att utrustning och procedur är anpassade till människan antas vara viktigt för provningskvaliteten. Vidare antas återkoppling och arbetsförhållanden vara kopplade till provningarnas kvalitet. Exakt hur oförstörande provning påverkas av ovanstående faktorer och faktorernas sammanlagda påverkan måste studeras vidare. Fortsatt forskning behövs inom områden som: procedurutformning; beslutskriterier; samt arbetsmiljöfaktorers påverkan och den effekt de har på provningskvaliteten.

1. Introduction

The purpose of the present report is to review existing theories and empirical results that are relevant to human factors research in general, and to non-destructive testing (NDT) of solid materials in particular. Relevant literature has been collected in order to see what has been done, and what needs to be studied further. The present review will present critique of earlier research as well as make suggestions for further research.

The structure of the present report is as follows. In the first part, there is a brief introduction to NDT, and some of the problems associated with NDT. In the second part there is a short presentation of some of the more interesting research findings in the field of human factors in NDT. The third part presents conclusions of the present review. Finally, there are suggestions for further research.

1.1. Non-Destructive Testing (NDT)

Non-destructive testing is a number of often complementary methods for evaluating the integrity of a material. These techniques include visual inspection, penetrant inspection, magnetic particle inspection, ultrasonic inspection, radiographic inspection, eddy current inspection, and acoustic emission testing (Bray & Stanley, 1989).

The goal of NDT is to determine the presence and location of discontinuities in materials, and in some cases also to characterize and determine the size of discontinuities. NDT is applied both as a quality control measure in the manufacturing process, and as an in-service damage assessment tool. The successful NDT program first of all requires reliable inspection equipment. Also, well trained engineers to plan the inspection and inspection procedure. Last, and certainly not least, the successful NDT program requires highly skilled operators performing the inspection. In addition, the success of the inspection process (inspection reliability) depends on the level of attention (vigilance) and judgment of the individual performing the inspection.

1.1.1. The effectiveness of an inspection

According to Firth (1992), a completely effective inspection would be one that gives clear and unambiguous answers to the following questions (Firth, 1992 p.165):

-Is this item fit for the purpose which it is required to fulfill?

-Are the flaws known to be present still benign or have any of them grown since the previous inspection?

Such a completely effective inspection would render conclusions like:

-The item is not fit for its purpose because of this type of defect of this size in that position

Or:

-The item is guaranteed to fit its purpose

Before an actual inspection takes place, a thorough investigation of what to be inspected is made. This is done to make sure that all flaws that may exist, and can jeopardize the integrity of each component, can be found with an acceptable level of certainty. For most NDT inspections, peak performance is not necessary. Flaws that are very large and very small are relatively easy to assess highly relevant or irrelevant respectively. However, as the line between an acceptable and an unacceptable discontinuity becomes harder to pinpoint, sharper skill is required. Since it is not possible to know, beforehand, if the flaws are going to be easy or difficult to assess, flaws must always be considered to be difficult to assess. Therefore, it is very important that human factors aspects are thoroughly considered before an inspection is initialized in order to reach, and maintain, a high level of quality in the field of NDT, especially in the nuclear industry.

1.1.2. The process of planning an NDT inspection

Since the consequences of a component not meeting its service requirements in a nuclear power plant can be disastrous, it is very important to be thorough, and to control and double-check the process. The different components have different purposes, and are therefore sometimes made of different materials.

Mordfin's axioms on NDT (Mordfin, 1985):

- All materials contain flaws
- Flaws in a material do not necessarily render it unfit for service
- The detectability of a flaw (also the risk of failure of the component) generally increases with the size of the flaw

As a consequence of the above axioms, components need to be inspected. Before a nuclear power plant is taken into service, its components are carefully controlled to make sure that no defects are left behind. When the plant is in service the components are put under mechanical and chemical conditions that could generate defects. Therefore, in order to secure the long-term integrity of the plant, in-service inspections, especially of components critical to safety, need to be made.

Bray and Stanley (1997) claim the following four points to be important to answer before engaging in an inspection.

- What is the prior probability of the existence of a defect?
- Given the existence of a defect, what is the probability that its size is in a specified (critical) range?
- Given that the size of a defect is in a specified range, what is the probability of detection for a particular NDT system?
- What is the probability of a defect of a certain size, propagating to failure prior to the next inspection?

In the above, one can say that the probability for human error is hard to determine (for each step, as well as for the whole procedure). Still the probability of human error has to be addressed (see elsewhere this report).

Dickens & Bray (1994) suggest that preceding an NDT inspection program the following points should be addressed.

- Type, location, and size of expected discontinuities
- The frequency of inspections
- The method to use for best detecting the expected discontinuities

- The equipment to use
- Preparation of the component.

For NDT in nuclear power plants in Sweden, the responsibility for the above points lies on the plant owner.

1.1.2.1. Type, location, and size of expected discontinuities

When planning an inspection, one very important aspect is to consider the types of discontinuities that can be anticipated in the material, and more important, to what degree such discontinuities interfere with the service requirements of the component. Also, the extent to which the expected propagation of a discontinuity interferes with the safety margin of a component, should be considered. One of the primary considerations in selecting an inspection process is the maximum allowable size of discontinuity and its location in each component. This establishes the threshold rejection criterion, the decision point of whether the component is fit for service or not.

When setting the decision point, it is important to take into account the consequences of Type I (misses) and Type II errors (false calls). If the rejection criterion is set so low as to ensure that no discontinuities (cracks) are missed, some satisfactory components are rejected as well (false calls). Thus, by reducing the probability of misses, the probability of committing false calls is simultaneously increased. This phenomenon becomes more of a problem when components with very high quality requirements are inspected (Dickens & Bray, 1994).

Ultimately there will be a tradeoff between misses and false calls. The following questions arise: "-What are the consequences of accepting a part with a critical flaw, versus those of rejecting a satisfactory one?" If there is a much higher potential cost (in effort, security, time, money) associated with accepting a part with a discontinuity, then the rejection criterion will be set low enough to ensure that no discontinuities are missed.

1.1.2.2. Frequency of inspections

When deciding on what components to inspect at a particular outage, the result of earlier inspections should be considered. The more critical a component is, the more frequent it needs to be inspected, and vice versa. The inspections must be made frequently enough to prevent failures, but not so frequent that the costs of inspection outweigh the costs of failure. Seen with the eyes of the plant owners, the goal of an inspection is to determine if the components inspected will remain satisfactory (meet their service requirements) until the next outage, or if they need to be repaired or replaced during the current outage. Due to the great costs, the plant owner does not want to make costly repairs or replacements without a good reason. Also, repairing/replacing a failing component is much more costly while the plant is in service.

1.1.2.3. The method to use for detection of expected discontinuities

The NDT methods used in Swedish nuclear power plants for in-service inspection of materials and welded joints in are mainly ultrasonics and eddy current, but to some extent, also radiography, magnetic particle, and penetrant testing are used. Ultrasonics and radiography are used for inspecting volumes (depth). Eddy current, magnetic particle, and penetrants are used for inspection of surfaces. All NDT methods are not physically capable of detecting *all* kinds of discontinuities.

The most common technique for inspection of thick power plant components is ultrasonic testing (UT) (e.g., Ammirato et al., 1989; Pond, Donohoo, & Harris, 1998). The technique has a long history of effective use in many industries, including both fossil and nuclear power plants. The performance can, however, vary from operator to operator and occasion to occasion depending on the situation and type of component to inspect (low reliability). The variation could lie in, for example, what the operators consider most important when they make their assessments. The assessments of e.g., through-wall size (height) of a crack are cursed with poor accuracy, and large variability. The biggest cause of this is probably the difficulty in identifying the diffracted signal from the heads of the cracks since these signals typically are only marginally stronger than those of the background noise.

Eddy current testing (ET) is a very sensitive method under good conditions. When the material to be examined is very thin, the entire volume can be examined e.g., inspection of thin tubes. However, ET is mainly used as a surface method. Eddy current relies on the electromagnetic abilities of the material. By exposing the material to periodically changing magnetic fields through a transducer, a secondary electromagnetic field is created. Cracks and other discontinuities affect the properties of this secondary field and can thereby be observed as differences in conductance.

Penetrant testing is exclusively used for inspection of surfaces. The method is very sensitive and normally easy to use. Since the technique is so easy to use, it is not always treated with the care and respect required to reach a valid result (Skånberg, 1991). In short, penetrant testing is performed by applying a fluid that has special abilities (e.g., low viscosity), if there are cracks in the surface were the fluid is applied, the fluid will go into the crack. When the excess liquid is wiped off, surface breaking cracks become visible, often with the aid of fluorescent lighting.

In manual inspection, the operator applies the technique on the surface of the material. In manual UT and ET the operator scan the material by moving the transducer across the area of interest. By scanning in a certain pattern, the operator makes sure that the entire area of interest is scanned. The operator has to rely on memory so that the area is neither over-, nor under-scanned. Not scanning the entire area could have serious consequences (i.e., missing a critical flaw). Scanning the same area more than once is a waste of precious time, and thereby money. In manual inspection data are not collected per se. Data are typically presented on a screen, and the operator assesses the data at the same time as he is scanning. If an indication of a flaw is detected the operator either makes notes of where the indication is and returns to it after the scanning is done; or, the operator interrupts the scanning, assess the indication, and later resume the scanning. The main forte of manual inspection is that it does not require much installation. Also, manual inspection is versatile and adaptable to sudden changes in the task.

Even though manual inspection is versatile, it is still performed by man, and man makes mistakes. Some human limitations of manual inspection can be overcome by mechanizing the inspection. The potential reliability of data-acquisition is improved, and the scanning is much more exact. Also, the fact that data are recorded alleviates the operator from potential time stress in analyzing the data. Mechanized inspection also makes it possible to augment scans form different directions with each other. However, mechanized inspection (sometimes) requires extensive installation, and one main problem during annual outages is shortage of time. Cost - benefit analyses of mechanized inspection does not speak in its' favor according to Karimi (1988).

One might be tempted to ask: "-why not just remove the human from the inspection routine?" In many cases a complete automation of the whole NDT process is, for several reasons, not possible (Dickens & Bray, 1994; Harris, 1990; Taylor et al., 1989).

In semi-mechanized inspection, the operator is aided in some way. For example when using MAPP-Scan or P-Scan, an operator scans the material manually, but is aided in making sure that the entire area is scanned. Furthermore, the signals are not only observed, but also saved for later evaluation. By applying a set of sensors on the component to be examined it is possible to record not only the ultrasonics but also the exact position of the transducer at any given time. This results in two major improvements of the manual inspection. Firstly, the operator can see, in real-time, what parts of the area he has covered (scanned) and not. Secondly, when all data are gathered and saved the data can be evaluated over and over again without the evaluator having to perform a second inspection (and being exposed to more radiation).

Other advantages with equipment such as the MAPP-scan are, firstly, that it makes it possible for qualified (level two) operators to do more work since they are not exposed to as much radiation. Secondly it enables the evaluator to assess at the data material whenever he pleases, and the possibility to confer with others.

1.1.2.4. Preparation of the component

Before an inspection starts, it is important that the components to be inspected are adequately prepared. The shape of the component, its surface roughness, and accessibility to the inspection surfaces, may make the inspection difficult and lead to errors. The components should also be cleaned, have paint removed etc., as necessary, so the operator can devote his time to the inspection. Perhaps the most important preparation is reducing the level of radiation as much as possible (within reason). This should be done so that the operators do not get suspended too early for receiving too much radiation. Furthermore, if the operator in some way feels uneasy this can seriously affect his performance (see 2.7.). The plant should also remove insulation etc., and build scaffolding so that the components are fully accessible. The operator can require additional preparations; he might make these himself or order them from the plant.

1.1.3. The process of performing NDT

In order to control performance quality, inspections follow a (predetermined) written procedure. The procedure that the operators follow determines what equipment to use in what situations, and how to calibrate and use the equipment. Also, the procedure determines how to evaluate signals received. The principal aim of the procedure is to secure the quality of the inspection. Since there is much evidence of high variability in the results of inspections (e.g., PISC III, Harris, 1990), a procedure has to be very strict. This in order to, as far as possible, minimize the risk of operators performing inspections in individual ways. So, if the operator follows a good procedure, the inspection meets the demands on security, safety, and accuracy. The aim of the inspection is thereby met, and the inspection is reproducible (i.e., subsequent inspections will render the same result, even if performed by another operator).

With increasing experience, the operators can develop what might be called "workmanship". Their NDT experience gives them the ability (intuition) to "feel" when the equipment shows an incorrect picture of the "reality". This workmanship may cause the operators to disagree with the guidelines of the procedure. The feeling of workmanship may be treacherous in that it may induce an exaggerated sense of competence in the operator.

Of course, it is one thing to follow the procedure blindly and to the letter. It is another thing to follow the procedure, i.e. doing what the procedure prescribes, and in the correct order. It is entirely possible to follow the procedure and still give expression for individual initiative. For example, when having followed the procedure and still having ambiguous information, the operator seeks additional information. In interviews made by the authors of the present review, operators expressed a conviction that: 1) all flaws cannot be found with the predetermined procedure; 2) all indications that the procedure "calls" geometry are not geometry (the same for cracks); and 3) the height as indicated by following the procedure could also be another (higher/lower) than the actual one (Enkvist, Edland, & Svenson, 1998). Depending on how confident the operator feels in his own ability, he might adjust his result in the direction his workmanship indicates is more correct. This is not at all desirable.

To make sure that each operator performing inspections in nuclear power plants can follow the procedure, and perform an inspection with a satisfying result, the operators have to pass a qualification test. The procedure and the equipment to be used at the inspection also have to pass a qualification test. The qualification test can be seen as a time limited "drivers license" for the operator.

Harris (1992) compared the performance of operators who where aided in their inspections with operators who where not aided. The operators who where aided had a 25% better performance (hit rate of 37.6 vs. 47.1%) than the non-aided operators. This implies that by following the chain of procedure, here in the form of clear and unambiguous instructions, performance is improved.

1.1.4. The assessment of defect size measurements

In NDT there is also considerable uncertainty in both the assessments of length and depth of discovered cracks. Studies have shown only highly skilled operators to reach a good performance in assessing the size of cracks. This is the case for through wall sizing in wrought austenitic stainless steel with manual ultrasonic techniques. Only the best operators reach a confidence interval of +/- 4 mm (Skånberg, 1991).

Although the procedure is strict and to a high degree guides the operator through the inspection, and the rigorous control preceding the inspection, there is still room for human error. The critical point of NDT is the assessment made by a human. Furthermore, the components used in nuclear power plants are individuals as well, and the stress the components are put under is also highly unique for each component. This means that the result of an inspection much lies on the most advanced part of the NDT system - the human.

1.2. Human performance

Human performance depends on three components (Bailey, 1989); (1) the skill and motivation of the person performing the task; (2) the complexity and duration of the activity; (3) the nature of the environment in which the task is performed. We can expect near-perfect or ideal performance to occur when the *person* is highly skilled and motivated, the *activity* is familiar and satisfying, and the *environmental* conditions are favorable. These three requirements are, however, not likely to be fully satisfied simultaneously.

Personal, social, and organizational factors can affect human performance in several ways. Previous training and experience will typically affect performance in a positive manner. The social influence of contemporaries/peers may contribute to the inspection process by the operator having his assessment of a signal conformed or disputed. Organizational factors such as the extent and quality of supervision; the support or frustration attributed to the organization; and the faith and trust placed in individuals or work-teams by the employer or contracting organizations; all influence the human reliability aspects of NDT.

The ability/status of man is influenced on the physical and mental environment, and of course, how he interprets them. Both physical and mental stress as well as anxiety and boredom can take their toll on performance and deplete the available resources. Furthermore, in situations where informational and supportive feedback is not readily available, man can develop negative beliefs about his capabilities. Under such circumstances, performance can quickly deteriorate due to the inhibiting influence of such beliefs. This could be remedied by the social support (from peers) at the workplace, and/or managerial/organizational practices.

1.2.1. Performance shaping factors (PFS)

Performance can not be said to be "good" or "bad" per se, it is not a static entity. Rather it is a highly dynamic interactive process, between the individual and the context in which he operates (e.g., Karimi, 1988). In order for the operators to perform well it is important that they are competent (skilled and experienced). Furthermore, the reliability of the NDT process is improved when boredom, vigilance decrement, and negative environmental influences are reduced (Dickens 1992). However a competent individual does not always perform well. There are factors that influence the performance; they are called *Performance Shaping Factors* (PSF). Factors that have been shown to influence (shape) man's performance adversely are for example: noise, heat, and fatigue. In a study of NDT, Harris (1990) organized PSF according to the following six areas: environmental conditions, protective clothing, time stress, organizational structure, knowledge and skills, personal work habits and attitudes. Below, some important human factors, that all are involved in the above areas, are presented.

1.2.2. Memory

Since the performance of an (NDT) inspection requires skillful performance of perceptual-motor skills and execution of cognitive skills indicating memory, memory systems are interesting and have implications for the education and training of NDT operators. For example, crack and signal prototypes have to be remembered and later compared to the results in an NDT session (inspection).

The division of memory into short-term and long-term memory has, more or less, become part of the "folk-psychology". Not everybody knows that long term memory also is divided into different types of memory (Tulving, 1993). One of the memory types that are of great importance to NDT is procedural memory, "how to do things". Tulving (1993) claim procedural memory to be, "expressed in the form of skilled behavioral and cognitive procedures, independently of any cognition". "Skillful performance of perceptual-motor tasks..." is an example of a task "...that depend heavily on the procedural memory system" (Tulving, 1993 p. 29). Manual inspection is a typical example of such a task.

Another type of memory important to NDT is PRS, or perceptual priming. PRS "is a specific form of learning that is expressed in enhanced identification of objects..." (Tulving, 1993 p. 29) (cf., pattern recognition). After an object (e.g., characteristic of an indication) has been encountered once, the perception of the same, or a similar, object on a subsequent occasion is facilitated (priming). Or, as Tulving puts it: "...in the sense that the identification of the object requires less stimulus information or occurs more quickly than it does in the absence of priming", (Tulving, 1993 p. 29). A typical perceptual priming task is where the operator has to identify a target (e.g., a crack) based on incomplete information.

1.2.3. Vigilance

In order for an operator to perform optimally his vigilance has to be sufficient during all critical moments of the inspection. However, research has shown that performance decreases with time, especially if the task requires judgment or integration of information (e.g., Mahan, 1990). When performing inspections, operators are subject to a stream of information from which they must extract possible signs of discontinuities. NDT is inherently repetitive, boring, and often tedious. Inspections are performed over extended periods of time, often hours at a time. At the same time, the operator must remain alert to possible signs of discontinuity, while realizing that not every indication is a discontinuity. For these reasons, vigilance is an important factor in the goodness of NDT (Webster 1988).

According to Dickens & Bray (1994) it is essential to recognize that NDT is essentially a vigilance task and to understand that maximum (optimal) human performance cannot be expected to persist over time. As a consequence, inspection reliability depends significantly on the vigilance of the operator. When having worked a long time, the operator may fail to notice significant stimuli (make misses), and may become more likely to produce false positives (false calls). The manual operator's attention is likely to be divided between the inspection search unit (e.g., transducer) and some sort of display.

During inspection, the operator must watch closely for changes in the stimulus; correctly interpret any detected changes; accurately record observed information; and correctly analyze the information (Webster, 1989). After a period of inspecting in which no discontinuities have been found it can become natural for the operator to expect that no discontinuity will be found with further inspection. This expectation increases the likelihood that the operator will miss the cracks that do exist. Since the probability of finding a discontinuity in nuclear power plants generally is low, the operator is continually being conditioned to believe that discontinuities are rare, and thereby "not likely".

1.2.4. Environment

The place of work for NDT inspectors is sometimes the hostile environment of a nuclear power plant containment. The operator will be required to handle the inspection equipment effectively and record data even in situations that may be dangerous, awkward, or frightening (e.g., hanging over the edge of a bridge or working in an area exposed to radiation).

1.2.4.1. Motivation

Motivation is also important for human performance and motivation can not be trained or taught independently. The conditions for high operator motivation depend on good organization factors such as management.

When performing a real life inspection in a plant, reporting a crack can lead to extensive repair and a prolonged outage, regardless if it is correctly assessed or not. A crack that is missed, on the other hand, is not discovered until it leads to leakage or breakage, or until found at later inspection. The component might function without problems for a long time. Thus, the operators only receive negative feedback; sometimes though, the negative feedback is delayed, perhaps infinitely. The fact that the operators only receive negative feedback can have a negative affect on their motivation.

A detected possible flaw will most certainly be double-checked before taking further action. Reported absence of flaws is, however, rarely double-checked.

1.2.4.2. Organization

Human performance exists in an organizational context, this must always be considered when trying to enhance the quality of performance. Even the performance of the best worker can be negatively affected if the organization is ineffective. For in-service inspections in nuclear power plants, on a macro-level of organization the different entities can be plants and contractors. On a micro-level the entities could be seen as different responsibilities and tasks. Some organizations can be temporary e.g., inspections; others might live longer (e.g., nuclear power plants).

1.2.5. Stress

Stress can be said to occur when the discrepancy between perceived demands and perceived resources is too high (e.g., Frankenhaeuser, 1986). Stress can affect human performance in different ways. In the long-term perspective it is associated with adverse health effects (e.g., Frankenhaeuser, 1986). In the present review, only the direct effects of stress on human performance will be considered.

Factors that adversely affect the perceived demands and/or the perceived resources, making the task more difficult and/or depleting the resources, are called stressors. Direct stressors such as lighting and noise can affect the input of data. For instance, if the lighting is bad, or if there is loud background noise, human perception might be impaired, resulting in our decision not becoming as elaborate as it could have been. Direct stressors can also affect performance when decisions are executed; for example, vibrations can make operations relying on delicate motor-skill, almost impossible to perform.

Stress has also been shown to impair working memory, thus reducing the possibility to learn from experience when under stress (Keinan & Friedland, 1984). The retrieval of information from long term memory is not disrupted by stress, if the information is well rehearsed and memorized (Stokes, Belger, & Zhang, 1990; Wickens, Stokes, Barnett, & Hyman, 1991). However some studies suggest that retrieval from long term memory under stress is prejudiced towards the well learned or the over-learned (Eysenck, 1976; Allnut, 1987). This implies that to replace an old information/behavior with a new one, the new information/behavior needs to be *over*-learned. The more well learned a fact or process is, the more effort has to be put into replacing it, if this is desired.

Earlier, researchers have been intrigued by the fact that individuals react differently when exposed to the same physical stimulus such as for example heat (Bailey, 1996), and noise (Grandjean, 1988). That is, one individual can feel stress in a situation where another individual feels no stress. Not only do different individuals react differently to the same stressor, but depending on the current status of the individual, the same individual can react differently to the same stressor at different points in time (Trumbull & Appley, 1986). Also, as the Yerkes - Dodson law from 1908 states, performance is best when arousal (stress) is not too high or too low (Yerkes & Dodson, 1908). With this in mind, it is easy to see that it is very difficult to supply an environment (physical and social) that result in optimal level of arousal for all individuals all the time. Hence, by providing the operators with the best environment possible, and giving them support, there is a better chance to reach an acceptable performance of NDT.

Operators performing an inspection in a nuclear power plant are exposed to numerous stressors. The most prominent stressors are fear of radiation, time pressure, and the alarm from the dosimeter. Roughly speaking, stress in the workplace can be avoided in two ways, design solutions, and personal solutions (Wickens, 1992). Examples of design solutions can be designing displays so that information is focused and organized (Zhang & Wickens, 1990), or data represented in graphs rather than digits (Schwartz & Howell, 1985). By employing such solutions the effects of stress and time pressure are reduced. Training has its positive effects in reducing stress, information retrieval from long term memory is relatively unaffected by stress. Other ways to cope with stress is to plan, anticipate, and rehearse actions that might be needed, thus gaining an advantage on a stressful situation.

Furthermore, social support has been shown to reduce the adverse effects of stress. Man can have social support in a number of ways; however, it is best if the support is available in the workplace. This way, the build-up of stress under the workday is limited.

1.2.6. Human error

Reason (1990) puts forward a taxonomy for human errors. He divides human errors into three categories: a) skill-based slips and lapses, where action deviate from intention; b) rule-based mistakes, where action follows the plan but the plan is insufficient (or applied wrong) to reach the desired goals; and c) knowledge-based mistakes, where actions are based on incomplete or inaccurate information.

In short, human *errors* are slips, mistakes etc. made involuntary or by lack of knowledge and/or understanding. *Violations*, on the other hand, are purposely made deviations from procedure/protocol. Violations can very well be made with the best of intentions and can sometimes be the only thing to do to stop a catastrophe. Thus, it is difficult, and perhaps not desirable, to protect processes from all violations.

1.3. Human information processing

In his book "Engineering Psychology and Human Performance" Wickens (1992) presents a model of human information processing in order to provides a useful method for examining potential limitations in human performance. Wickens provide an illustrative model that displays the important role of attention resources. This is presented in Figure 1.

In this model (Fig. 1), Wickens divides human information processing into five critical stages, each in some way transforming the data (information). The critical stages are "Sensory Processing", "Short-Term Sensory Store (STSS)", "Perception", "Memory", "Decision and Response Selection", and "Response Execution" (Wickens, 1992 p. 17). The last stage (response execution) can be one of the stimuli that start the next process, thus forming a closed feedback loop. In each of these stages the process can be disturbed by external factors (e.g., environmental stress) and/or lack of attention recourses.

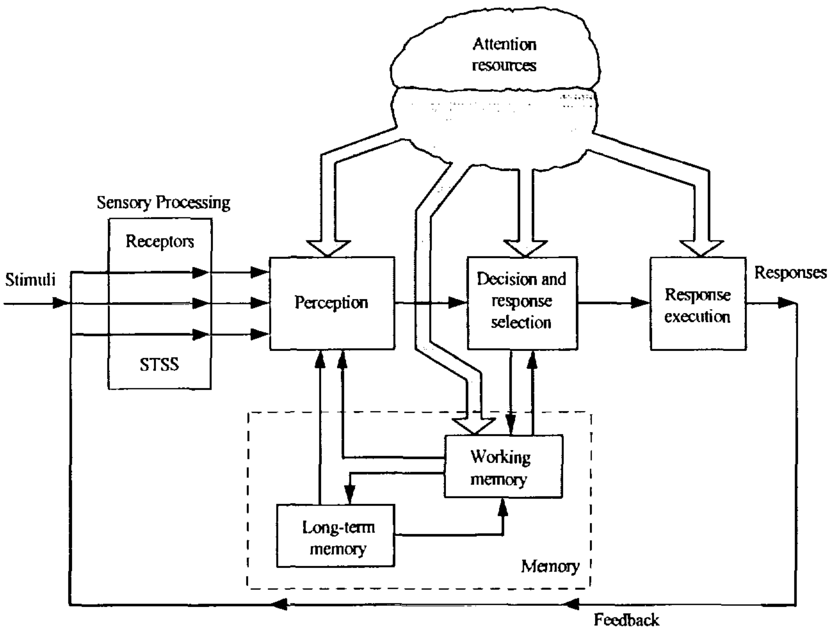

Figure 1. *Wickens model of Human Information processing. Redrawn from Wickens, 1992 p. 25. For further details, see text.*

If one process requires more attention resources, it follows that fewer resources are left for later processes whose performance quality then will decline. Experience leads to less need for attention resources. This is one of the main reasons for why training improves performance. Also, experience sometimes provides the operator with less common situations. Such experiences can make the operator knowledgeable, more flexible and thereby a better at inspection.

1.3.1. Signal detection

We typically look at signals as just signals that are received. However, signals seldom occur in a vacuum, typically there is background noise present as well. When the signal is much stronger than the noise it is easy to distinguish the signal (e.g., light in darkness). Sometimes, however, the signal can be drenched in noise (e.g., normal conversation in a loud nigh-club).

According to "Signal Detection Theory" there are two kinds of correct decisions, and two types of errors the operator can make, in determining if a crack is present or not (e.g., Wickens, 1992, and Dickens & Bray, 1994). A model of the correct and incorrect answers according to the signal detection theory is presented in Figure 2.

As stated earlier, either the operator assesses an indication to be a crack, or he does not. Nevertheless, in the latter case (no crack) the operator might detect some kind of stimulus (signal) of the same dimension as that of a crack, but not sufficient for a different assessment. These other kinds of stimuli/signals can be viewed as "noise". This noise is determined by a number of factors, e.g., discontinuities in the material, the surface of the component, or interference from equipment.

The prior probabilities of the two states of the world (crack - no crack) need not be equal. An operator can have certain experience (or common knowledge) that make him convinced that an occurrence of a crack at one particular place is not very probable (Swets, 1996). This may lead to him to discriminate some stimuli; that is, he might raise the decision criteria (i.e., making it "harder" for an indication to be assessed as a crack). With the decision criteria set high, statistically speaking, the number of false calls is decreased, at the same time so is the number of hits. The operator's experience (prior probabilities) tells him that the indications, he in the above case discriminates, are in fact not cracks.

True state of the world

Operators assessment		Crack	No crack
	Crack	Hit	False call
	No crack	Miss	Correct rejection

Figure 2. *Depending on the true state of the world and the operator s assessment of it, there are four possible outcomes according to signal detection theory. For further details, see text.*

Conversely, if the operator believes one particular place to be especially defect-prone he might set the decision criterion low. Thus, statistically decreasing the number of misses, but then again, increasing the number of false calls. The risk of severe consequences due to a miss can be seen as a good reason to prefer to set the criteria high. On the other hand, a high set criteria results in a large number of false alarms, and can thereby result in unnecessary, and extremely expensive, repairs (Swets, 1996), and thereby lack of trust in NDT. In other words, the "pay-off" (cost - benefit) of the decision has to be considered as well as prior probabilities.

1.3.2. Pattern recognition

The major crux of the NDT process is the interpretation and analysis of received stimuli (is the stimulus an indication of a crack or not). In the case of mechanized, and semi mechanized inspection, the measurement output can be recorded and stored in a meaningful way. This means that assessments can be made, using the recorded data, in a less disturbing context, than when performing actual inspections. Furthermore, it is possible to have more than one person assessing exactly the same (objective) information.

In the case of manual inspection however, the output is not stored; data are collected and analyzed simultaneously by the operator. For all types of inspections, the interpretation of stimuli requires the operator to recognize the stimuli as indicating certain characteristics of the inspected material. This means that the received stimulus pattern must be compared to patterns previously received (e.g., from calibration, or training/education) which indicate typical characteristics (e.g., weld-root or crack) of the material. These patterns are stored in long-term memory and brought into working memory in comparisons with the current measurement results. If a pattern stored in long-term memory matches the perceived one, this match will result in the experience of recognition. However not all *recognition* is conscious. Man responds more quickly to a stimulus that is known to him, than he is to a new stimulus; that is, a sort of a reflex can evolve. Training provides the operator with experience that is useful in inspection. This means that less attention resources are required, and the operator would not be fatigued because he has automatic and immediate access to memory "templates" to match with the current patters.

1.3.2.1. Influences of pattern recognition on information processing

As shown in Wickens (1992) model (Fig. 1.) human performance is affected, among other things, by the attention resources available. The attention resources can be depleted by numerous factors (e.g., PSF), some of which will be discussed below.

During an annual outage in a nuclear power plant there is only a small window of opportunity in which to perform NDT. Furthermore, there are increasing demands from the utilities to make the, typically very costly, outages shorter. By taking these factors (time and cost) into account it becomes clear that the system must be well tuned (optimized). It is important to study how the information processes affected by e.g., different amounts of time pressure; "-Is time pressure worse when operators work long periods?" "-Can an acceptable level of performance in NDT be achieved when the NDT operator is working under time pressure?" It can be possible to optimize pattern recognition by appropriate training, and by organizing the task in a suitable manner.

Firstly, performance in pattern comparison tasks improves if the comparisons are allowed to take more time. Therefore, NDT operators should have enough time to be able to make correct measurement-assessments. Regrettably, ideal conditions of any kind, let alone time, seldom exist for an NDT operator during an outage in a nuclear power plant.

Secondly, to ensure a fast and accurate pattern recognition process, the operator must undergo adequate training. In NDT, the target pattern to which the actual stimulus pattern type/characteristics is compared has to be stored very thorough, again indicating the importance of proper training. The goal of this training should be to provide the operator with patterns/characteristics that are typical and atypical of e.g., cracks and lack of fusion. By retrieving these patterns from memory the operator can make comparisons rather than logical inferences. It is easier to pick the correct alternative, based on recognition, that it is to make logical inferences.

When a stimulus is presented in a suitable context, recognition is made faster than if presented in the wrong context or in no context at all. Thus, parts of the background information make parts of the information redundant. Redundancy is beneficial when level of noise (i.e., non-relevant information) is high. That is, each "evidence" by itself is circumstantial; put together, they close "the case". On the other hand, it is possible to make wrong inferences; cues that are ambiguous might be interpreted erroneously if the operator is not focused (low vigilance). The operator has to consider the information in the correct context. Cues that are not important in one context can be vital in another. In the area of NDT, an operator that he has a mental set more fitted to another context than the context, from which the stimulus comes, might erroneously interpret vital information as noise, and thereby ignore it, or vice versa. Thus, by transferring behavior from a situation in which it is appropriate to a situation, in which it is not, can have grave consequences. By assuring that the operator is considering the information in the correct context, the probability of a correct assessment increases.

Looking at the findings above, some general inferences can be made. Some critical aspects of pattern recognition that can affect the quality of NDT can be pointed out. For example, recognition of a pattern can be facilitated by: 1) ensuring enough time for each task; 2) reducing the information complexity of the stimuli; 3) assuring that each task is assessed in an appropriate contextual approach or mind set.

1.4. Decision making

Non-destructive testing can be seen as a man - machine system with the task to detect and diagnose indications of defects (Skånberg, 1994). The NDT-system consists of equipment, procedure, and operator. The quality of the inspection depends on how well equipment and procedure fits the task, and on how the operator acts as observer and decision maker under the conditions at hand (Skånberg, 1994). The task can vary in difficulty on several levels e.g., homogeneity, texture and surface of the material; the properties of the flaws; the situation at hand; and perhaps foremost, the operator as a decision-maker. By looking at NDT as a decision making task, inferences might be made from research to NDT. Below some findings in, and conclusions of, research in decision theory are presented.

There are numerous models describing how we *should* make decisions (i.e., normative decision making). However, these models are rarely followed when decisions are made in real-life (for a review of empirical findings see Svenson, 1979). In stead of being rational we are *fairly* rational and instead of basing our decisions on all information available we use only a *part of* all information. One reason for decisions not being made "the right way" (normatively) is that it takes too much time and/or effort. We tend to invest more effort in decisions that are more important. The decisions made under NDT inspections are typically very important, an incorrect decision could have grave consequences.

In inspection situations where there are no clear evidence in either direction ("crack" or "no crack"), the tiniest thing can be enough to tip the scale. The proverb "First impressions last" has a scientific correspondence in the anchoring effect (Tversky & Kahneman, 1974). This means that people tend to "anchor" their decision according to the first given information; later received information only affects the decision slightly. The operators must thus keep from basing their decision/assessment on less than all information. Also, cognitive dissonance reduction (Festinger, 1964) may take place on a level not conscious to the decision-maker; the dissonance reduction can become an obstacle when a person needs to *react* to incongruent information. If, for example, an operator performing an inspection finds much information indicating that the component at hand is flawless; the operator might believe the component to be flawless. Information received after the opinion is formed can thus be (non-consciously) ignored or down graded in order to reduce the cognitive dissonance. This can seriously affect the operator's detection performance.

Awareness of this phenomenon is of particular importance in manual inspections, since there are no objective records. In short, the operator can have the ability to perform an impeccable inspection, but due to circumstances (e.g., PSF, or dissonance reduction) the performance is below par (cf., Dahlgren & Skånberg, 1993).

To cope with limits of attention and cognitive resources man uses heuristics ("mental shortcuts"). With the emanation of experience, man reduces pressure on short time memory by dividing the information into chunks that are easier to process (chunking). Furthermore, as a result of the experience man gains a repertoire of hypotheses and actions stored in long-term memory (heuristics). When only little effort can be invested in a task, a heuristic strategy can lead to a better, or more economical, performance than the optimal strategy (Wickens, 1992). When much effort can be invested however, the optimal strategy leads to improved performance. The heuristics produce results that often are good enough, but usually not as precise as they could be (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974). Training is important for the building up of heuristics, and to free attention resources so that heuristics can be avoided. These findings have implications for the training and education of NDT personnel.

When an operator is under extreme time pressure, his performance is likely to be affected. For example, in order to cope with the time pressure the operator might not take all information into account when assessing an indication. In cases where *any* decision is better than *no* decision, coping-strategies can be recommended. For all other situations it is better to organize the task so that there is enough time to make a decision with a quality worthy of the task.

1.5. MTO, a systems approach

Superiority in systems can never be achieved without
excellence of human components, of hardware components, and
of optimization of the interactions among them
(Christensen, 1987, p. 8)

Christensen (1987) claims that a systems approach is something that the development of even the simplest products can benefit from, and that this is paramount to the successful development of complex systems.

The present review looks at NDT from a "Man - Technology - Organization" (MTO) perspective. The concept of MTO is a system perspective that focuses on the interface between human, technological, and organizational subsystems. MTO puts emphasis on the performance of a system as a whole, rather than the performance of each subsystem (Kecklund, 1998). The Man-Technology-Organization (MTO) approach has previously mainly been taken to analyze the security of complex systems (e.g., in the Swedish nuclear industry).

By applying an MTO-perspective, one looks at human behavior as an interaction between the human and technology, in an organizational context (e.g., allocation of resources, general working-conditions, and educational level). The MTO approach considers the human as a resource to meet the purpose of the system. When applying an MTO-perspective one can utilize existing knowledge in a new way. With an MTO-perspective it is possible to find out how (simplified laboratory) research findings can be applied to complex real-life tasks. This may be done by analyzing real-life tasks with results from different research areas in mind.

The types of errors that can be made in the NDT setting are false negative calls (misses), the operator reports no discontinuities although such are present (Type I error); and false positives (false calls), the operator reports discontinuities although there are none (Type II error). However both these types of errors presupposes that the inspection technique being used is appropriate and able to detect any discontinuities present in the material. That is, all variations in the quality of the inspection are seen as a result of variations in the performance of the operator.

1.6. The applied problem

The quality of an inspection is influenced by many factors; most factors originate from either the physical environment or the human factors area. Examples of environmental factors are accessibility and lighting.

Research has shown that, even after controlling for equipment, procedure and formal NDT-qualification of operator, the quality of an inspection to a high degree depends on the operator performing the inspection, (e.g., PISC III Report, 1994; Wheeler, Rankin, Spanner, Badalamente, & Taylor, 1986; Taylor et al., 1989; Skånberg -91; Swets, 1996). These variations can result in costly and time-consuming repairs being made unnecessarily (i.e., as a result of a miss, as well as a false call). Furthermore, these variations can have a negative effect on the credibility of in-service inspections (Skånberg, 1994), which are important to control the integrity of the power plants.

To perform a high quality inspection is no easy matter; the information received can be ambiguous. It is the operator's task to interpret and assess the information and to make a decision concerning whether there is a crack present or not.

2. Findings

Taylor, Spanner, Heasler, Doctor, & Deffenbaugh (1989) tried to quantify the detection of inter-granular stress corrosion cracking (IGSCC) and found a wide variation between the operators' performance. Ranging from one operator with acceptable performance to the majority of operators with unacceptable performance. Swets (1996) examined a study on the inspection of fatigue cracking in aircraft parts, using either ultrasonic or eddy current techniques. Swets found that "The variation across technicians [operators] is about as large as it could be" (Swets, 1996, p. 137). In a study made for the Federal Aviation Administration (FAA), Spencer & Schurman (1986) found no single individual factor to explain the varying performance. Environmental factors were considered to be of importance for the quality of the inspection. However, no single environmental factor was found.

Stephen Doctor suggests that (UT) inspection performance can be enhanced if the operator is assisted in verifying the scan path, and storing "*the flaw signal for later* *review*" (Doctor, 1995). According to Doctor, this would be particularly beneficial for small flaws. The findings in Doctor (1995) study suggest that an application of a semi-mechanized technique such as the P-Scan or the MAPP-Scan (see semi-mechanized testing, this review) can supply this support. The aim of the present review is to look at what have been done to resolve the problem of varying results, and to come with suggestions on how to improve inspection performance.

Research results show a high correlation between human factors and the number of errors (Harris, 1990). That is, the less that has been done to prevent human error to occur, the more errors occur. The relative attention and resources spent on equipment, organization, and operator respectively has not been in proportion to the relative importance equipment, organization, and operators respectively have on the quality of inspection process as indicated by field experiences and round robin tests (Hoogstraate, 1998). This shows that human factor studies are an important approach that has a potential for improving the reliability of NDT. Earlier research in the area of human factors in NDT have tried to analyze and quantify the effects of tasks, tools, skill, and cognitive information processing elements on man - technology systems. The operator him/herself and how the performance of the operator can be improved must be studied more closely in order to reach an acceptable level of quality in NDT.

Below, earlier findings that have implications for inspection performance will be presented. First, findings that have bearing on the operator will be presented; second, follow findings that are important to consider on an organizational level. Third, findings will be presented that concern the working conditions of the operators.

2.1. The operator

"...the reliability of an inspection process hinges on
the performance and judgment of the operator
carrying out that inspection."

(Dickens & Bray, 1994. p. 1033)

In order to carry out an inspection successfully, there are some requirements for performing an inspection correctly (e.g., certain motor skills and abilities). The level of skill and knowledge that an operator possesses is determined by the amount of training, education, and experience he or she has. Psychological abilities that have been coupled with high performance in NDT are e.g., spatial ability and locus of control (Harris, 1992).

The performance of an operator basically depends on his "Knowledge, Skills, and Abilities" (Pond, Donohoo, & Harris, 1998). Training and education can improve knowledge, skill, and performance abilities. However, it might be advantageous to ensure that the people hired, to do such complex tasks as NDT, have the basic abilities required, especially in such a quality demanding area as the nuclear power industry.

Bray & McBride (1992) stress that to improve the performance of NDT, both training and education must be considered. The training of operators should be *specific*, and *task oriented.* Bray & McBride argue that the goal of training could be to have operators perform inspection, without necessarily having a depth of understanding the underlying science. Education should according to Bray & McBride be broad and more long-term than the training.

Karimi (1988) explains the organized pattern of behavior (performance) across individuals and across occasions of performance within individuals. The results of the Karimi study imply that operator performance could best be understood as being a function of *contextual* and *motivational* factors. Although skill is a crucial component of effective performance, Karimi suggests that skill can readily be acquired through (technical) training.

The training and education that the operators undergo does not always reflect the difficulties that may occur under real life inspections. The procedure is not always vast enough, and is seldom stringent enough to provide the guide and aid necessary for evaluating signals and assessing the size of a defect. Therefore, experience is paramount for successful inspections.

Karimi (1988) suggests that inspection performance could be improved by improving the *selection of personnel.* Through active recruitment practices and careful application of selection criteria the industry can ensure that the most consistently competent workers enter the work force. The recruitment of persons with the best abilities for the inspection task is mentioned by others as well. The Electric Power Research Institute (EPRI) has a software test called "Dynamic Inspection Aptitude Test" (DIAT). This test is developed by Doug Harris, and it tests abilities similar to those required in manual ultrasonic NDT (Harris, 1996). The DIAT has been shown to have a high validity, but it is still under evaluation. A test like this, together with a test of extrinsic/intrinsic motivation could be very valuable in the selection of personnel. By hiring the inspectors best suited for the task and giving them proper training, a lot could be won as far as getting more homogenous results of NDT from different inspectors. One of the conclusions in the PISC III study was that to improve predictability of the performance of inspectors it is important that psychological test techniques are continuously developed.

Another way of tackling the problem of "the right man at the right place" is suggested by Halmshaw (1991). Halmshaw states that many NDT methods are so well developed, that even semi-skilled operators (cf., MAPP-scan this report), can apply them. Halmshaw's opinions agree with Bray & McBride (1992) in that with the advent of microcomputers, the operators do not necessarily need to have an understanding of all the physics and the theory behind the technique used. However, it is desirable that supervisors, designers, and development engineers "have a thorough scientific understanding of the fundamental physics involved" (Halmshaw, 1991 p. 24).

Norros (1998) discovered that operators working with detection preferred the use of standard methods whereas operators with higher theoretical training working with both detection and characterizing emphasized the use of the interpretive skills of the operator. Norros concludes that, as in earlier studies, the attitudes of the operator influence the performance of the inspection. Norros claims that the low frequency of defects and the strong division of tasks among operators promote a standardized way of acting. Finally, Norros states that changes in work organization, and further theoretical training is needed for the operators to get the experience needed for attaining adaptive and reflective expertise.

In a study for EPRI, Harris & McCloskey (1990) found cognitive processes to be important to the success of ultrasonic inspections. The following four cognitive elements were found to be associated with successful inspections:

- Development of explicit hypotheses as an integral part of the inspection process
- Avoidance of reaching a conclusion too early in the inspection process
- Application of knowledge during the inspection by putting it in the form of if-then logic
- Avoidance of the arbitrary elimination of information during the process of reaching an inspection conclusion

Utilizing the findings of the above study, Harris (1992), compared the inspection performance of operators who used a decision aid, with operators who did not. The aided operators had a 25% better performance than the non-aided operators. The operators who used a decision aid, had had only one hour special training in using it. Still, the aided operators had 47.1% successful outcomes, to be compared with 37.6% for the non-aided operators. The small experimental difference between the two groups of operators (a decision aid, and one hour training), and the big difference in performance, points to the importance of following the procedure. The fact that the aided operators reached only 47.1% successful outcomes, led Harris (1992) to conclude that there were still substantial room for improvement.

After a failure to detect cracks in a Swedish power plant, Dahlgren & Skånberg (1993) performed an analysis of the records relevant to the failing inspection, and interviewed all personnel involved. Dahlgren & Skånberg also reviewed the procedures used at the time. This was done to get an understanding of how the cracks could have been missed, and how such a scenario could be avoided in the future. Did the miss occur solely as a result of a miss made by the operator responsible, or were there ambiguities in the procedures that contributed? Dahlgren & Skånberg discovered that the operator did not scan (inspect) areas of the components where, according to his experience, cracks do not appear. That is he left the procedure because he "knew" it was unnecessary to scan the whole inspection area. In interviews with the operator, he claims that he, at the time, felt "rusty" and "very stressed", he also felt alone and anxious due to the radiation. He remembered the inspection site as "dirty" with "lots of radioactive grinding dust all over" Dahlgren & Skånberg, 1993).

In a Finnish study (Norros, 1998), the operators themselves did not agree with the international experience that time pressure, fear of radiation, and motivation affect the performance of NDT. On the other hand, also in interviews, (Wheeler et al., 1986) the operators implied that their performance was affected by the attitude of the plant, as reflected by plant cleanliness. In a study reported by Wheeler et al. (1986) the fear of radiation decreases with experience. In nuclear power plants there are different levels of radiation. It is hard to believe that any amount of experience can keep an operator calm in areas of extreme radiation.

Berglund & Lindberg tried to improve the accuracy and reliability of inspections by combining the results of operators in pairs, favoring "calls" (see Berglund & Lindberg, 1990, for details). The performance of the operators was categorized as good, average, and poor, a total of 55 pairs were combined. Results showed that the performance of operators, whose individual performance was average or poor, improved by being combined with another operator's. The same effect was found even though two operators with poor performance were compared.

As mentioned above, motivation also plays an important role for the quality of any qualified task. A significant component of an inspector's motivation is the understanding of the importance of the inspection task. It requires proper training, supervision, and adequate provisions for the inspection including suitable inspection equipment, guidelines, working space, and time, to perform a thorough inspection of a high quality.

2.2. The organization

Pond et al. suggest a shift in the emphasis of research "*away from such variables such as effects of heat and noise and towards those such as managerial policies, job design, and defining UT/ISI-relevant KSAs* (Pond et al., 1998, p. 12). That is away from performance shaping factors, towards knowledge, skills, and abilities relevant to (ultrasonic) in-service inspection.

Findings in the (Karimi, 1988) study suggest that contextual factors (e.g., supervisory and managerial practices) are major determinants of the outcome of performance - assuming that workers (operators) have the basic skills. Karimi (1988) suggests that inspection performance could be improved by improving the motivational state of workers, for example by giving feedback geared toward immediate task performance. That is, providing the workers with a supportive and learning environment.

In a Finnish study (Kettunen, 1997) people in the NDT-business rated the reliability of NDT methods to be very high and depending on various human and organizational factors with the operators' attitudes perhaps being the most important factor. In another Finnish study, Norros (1998) felt it necessary to interpret NDT not as an abstracted signal detection problem but as an intentional activity in meaningful physical and social contexts. Whereas some studies on NDT concentrate on what went wrong, Norros (1998) wanted to use an approach that would better reflect the intentional and contextual nature of the activity as a whole. By analyzing interviews with the operators Norros first examined what the operators thought affected the reliability of inspections. Second, Norros examined the operators' conceptions of the decision-making demands of their work. In the study Norros found that the operators did not find the traditional (performance shaping) factors a problem for inspection reliability, instead the operators emphasized the role of the foremen. The foremen were considered to be a significant resource in planning, preparing and coordinating the work and monitoring the radiation exposure (Norros, 1998). Norros interpreted this as good communication being vital for high reliability of an inspection. This points to the organizational factors being very important for the quality of the inspection.

One of the responsibilities of management is to encourage the work force. If management feels that the inspection task is important, the operators will be well provided for. On the other hand, if the task has low priority with management, the operators will be perceptive enough to realize that. When this happens, work performance will decline accordingly. Typically the inspector will become more lenient and will pass questionable items since the consequences of accepting a faulty item are less immediate than the consequences of rejecting it (Warm, 1984).

Motivation can be divided into intrinsic and extrinsic motivation. People that are *intrinsically* motivated believe that doing a good job, or mastering the challenge of work, is its own reward. People that are *extrinsically* motivated, on the other hand, work to achieve other goals (e.g., make money, fear of losing ones job). Karimi et al. found that workers who are *intrinsically* motivated to perform well might be the most consistently effective workers. In contrast with this are workers with *extrinsic*", they are more likely to behave in a somewhat more erratic and less competent manner (Karimi et al., 1989). Again pointing to the need for proper training, and perhaps selection of personnel. But also pointing to the importance of social support and management providing enough security for the operators to feel intrinsically motivated.

One of the findings in the Taylor et al. (1989) study was interpreted to imply that both NDT-supervisors and utilities are more concerned with false alarms than misses (Taylor et al., 1989). The reason for this can probably be found in risk assessment or decision theory. Supervisors and utilities probably viewed false calls as unnecessary and resulting in high costs and time lost without "real" cause. This can lead to the operators receiving negative feedback. Positive feedback is an important factor in learning, and in gaining experience. Also, in tasks where perception or motor skills are involved, man perform better when he gets feedback (e.g., Drury & Fox, 1975). Negative feedback, on the other hand can lead to loss in self-confidence (c.f., "learned helplessness", Seligman, 1975). In the Wheeler et al. (1986) study only 3 of 12 operators received feedback at least 50% of the time, and 5 of 12 received feedback less than 10% of the time. This have implications for securing enhanced experience of the operators, and not least important not receiving feedback can affect the attitude and motivation of the operators.

2.3. The work conditions

From a human factors study at Batelle PNNL, Taylor and collaborators concluded that the reliability of UT is affected by performance shaping factors (PSF) as well as environmental factors (Taylor et al. 1989). The operators in the Taylor et al. study expressed only mild concern regarding unusual work schedules due to plant priorities and work involving other plant personnel. Most operators felt that their performance would remain efficient for up to 12 hours long workdays. However, research (e.g., PISC III) show that working long shifts (in a hostile environment) can cause tiredness and de-motivation. These effects can have strong effects on human reliability in the inspection process.

Attempts at keeping a constantly high vigilance are mentally and physically fatiguing, and performance will deteriorate over time (Wheeler et al., 1986). Still, in spite of this, in manual UT, constant monitoring of the instrument is required in order to spot critical signals (i.e., indications of flaws) in a constantly shifting noise - signal pattern. In the Wheeler et al. (1986) study, the operators reported two types of fatigue. The first type of fatigue occurs after having worked long hours especially when the operator wears protective clothing. The second type of fatigue occurs after working long periods without a day off. While the first type of fatigue leads to single errors and mistakes, the second type of fatigue affects the general disposition, and attitude of the operator, which is likely to affect performance in general (Wheeler et al., 1986).

Human factors specialists in the Taylor et al., (1989) study concluded that, because manual UT examination involves tactile feedback of probe location, and contact with the surface, small transducers pose problems, especially when the operator wears bulky gloves. Workplace ergonomics is an important human factors research and application area. To illustrate, in NDT, the interface between the testing equipment and the component inspected is the transducer. In manual ultrasonics, the transducer is typically fairly small and square (cf., Spanner, Badalamente, Rankin, & Triggs, 1986). Findings in the Wheeler et el. (1986) study imply that UT equipment have design deficiencies which make them sub-optimal. With many high demands on the operator's performance (e.g., moving the transducer in a fine pattern, scanning at certain angles, at all times keeping contact between the transducer and the material, at no time tilting the transducer); it is quite astonishing how little effort has been put into adapting the equipment design to the operator.

Doctor (1995) suggests that further studies, like the PISC III, is needed, however, a suitable experimental design is necessary to better understand "*factors that lead to errors in inservice inspection*". Also, Doctor concludes, future studies "*must employ sufficient numbers of subjects so that the results can be interpreted in a manner meaningful to the nuclear power plant inservice inspection environment* (Doctor, 1995). Next in the present review, further conclusions from earlier research are drawn.

3. Conclusions

"A good and coherent picture of human and organisational factors influencing the reliability of NDT operations is still missing" Kettunen & Norros, 1996

If the man - technology system is to work optimally, experts of technologies and experts of the human have to cooperate, and encompass knowledge from both areas. It is clear that research on the most complex part of complex systems - the human - is under-invested (Hoogstraate, 1998). If one looks at the money spent optimizing equipment and technology as compared to human factors, and compares that with the probability of failure for those respective areas, a question arises: "-Would it not be logical to place more resources on the area where the most failures occur?"

When designing equipment, the operator is usually considered late in the development process. First, either a need is discovered or created, then the equipment is developed around its function. It is only at the end of the developmental process that man is considered; man is taught how to use the equipment. In dealing with such uncontrollable entities as natural forces it is perhaps best for man to adapt, but when dealing with man-made equipment it must be better to adapt it to man (cf., "The human use of the human" Wiener, 1950; "Fitting the Task to the Human" Kroemer & Grandjean, 1997). In doing so, man does not have to waste time and effort on learning and remembering behaviors that does not come natural, in order to operate the equipment. If tasks are fitted to the human, and the human behavior, there is much less room for human error.

There are human factors handbooks that could be used by the designers of various technical equipment (e.g., Bailey, 1989; Grandjean, 1995; Kroemer & Grandjean, 1997; Salvendy, 1997). These handbooks deal both with harnessing man's abilities, and how to avoid unnecessary wear and tear on man (workplace ergonomics). As stated earlier in the present review, there are a number of factors that can affect performance (Wickens, 1992; Pond, Donohoo, & Harris, 1998). Engineers make many decisions about inspections and procedures that bear directly on reliability, and often directly influence the inspection operator. The inspection operator is often unaware of the details (and potential errors) in these decisions, and therefore cannot be expected to compensate for inadequate engineering judgment. It is important to consider human factors early in the design. If applied later human factors considerations may become "too little, too late".

One main problem when applying research results to "real life" problems lie in the fact that research results derive mainly from experiments that have been performed in a context of relatively low complexity. Wiener (1984) thinks that uncritical inferences are made from experimental and empirical research results to "the shop floor". Wiener bases his opinion on the fact that while researchers look at tasks that induce fatigue, the industry uses techniques that decrease fatigue. Examples of such techniques can be; task rotation, scheduling of pauses, or having workers set their own pace. So, applying the above results, on a complex task, such as that of NDT, must be done with great care. Norros (1998) suggests that "instead of interpreting inspection as an abstracted signal detection problem, it should be considered as intentional activity in meaningful physical and social contexts".

Below, some conclusions will be presented, which have implications for inspection performance. First, conclusions with bearing on the operator, Second, conclusions with bearing on the organization are presented. Third, conclusions that concern the working conditions of the operators are presented. Last, conclusions are presented which have bearing on the qualification of operators.

3.1. The operators

Kettunen & Norros, 1996 conclude: " ...no single human or organisational factor is responsible for the NDT performance fluctuations obtained in various reliability studies..."

Even though the importance of human factors for the reliability of inspection is not questioned, no clear correlations have been identified between single human factors and inspection performance (e.g., Dahlgren & Skånberg, 1993, PISC III, 1995). The decision criteria or cues used for assessments under the inspection, is influenced by the training and experience of the operators, more than the procedure used at the time. The ability/status of the operator is much influenced on the physical and mental environment and how he interprets it (Skånberg, 1994). To solve the problems associated with human operators by eliminating the human element might be tempting. However, a total elimination of the human from the inspection process is, for a number of reasons, not yet possible (Dickens & Bray, 1994; Harris, 1990; Taylor et al., 1989. Due to the fact that different operators use different criteria for assessments and decisions, the results vary even if the inspection is mechanized (Skånberg, 1991). Much of the problem still lies in the assessment of the indications, not if they are detectable or not. Therefore, the main approach to take in order to improve performance is to turn to the human factors in order to optimize the diagnostic abilities of the operators.

The operators basing their decision/assessment on less than all of the available information is probably what lies behind the poorer performance of the group not being aided in their assessments, in Harris' (1992) study. Thus, the inspection procedure must be correct, appropriate, and easy to follow. If all operators follow the same procedure, the plants know what they get, even if it means less than optimal inspections.

Findings in the Harris & McCloskey (1990) study suggest that the development of explicit hypotheses as an integral part of the inspection process lead to high performance. This is probably so because it helps the operator to structure the inspection. Thereby the operator avoids reaching a conclusion too early in the inspection process, before all information has been obtained and considered. Bray & McBride mean that the training of operators should be specific and task oriented, so that the operators gain concrete experience. For example, experience in developing explicit hypotheses. Experience can make the operators more confident in themselves as (NDT) decision makers.

When training and qualifying, the operators know that there are defects in the components, in "real" inspection defects are much less frequent. The operators are also aware that their ability is tested, this means that they try to perform their absolute best. Of course they try to perform their absolute best in "real" inspection as well, but then they work day in and day out, and therefore loose in vigilance. A qualification is for a limited number of days, and if successful can lead to years of not having to fully prove ones ability in the discrimination, characterizing, and sizing of cracks. Attitudes, motivation, and criteria for assessments and decisions can also be different for training, qualification, and inspection in the plants (c.f., Wheeler et al., 1986). The criteria for assessments and decisions that the operators use in a situation where they are uncertain can be affected by the different consequences that can follow.

3.2. The organization

Nuclear power plants are required to have NDT inspections performed; the only time when it is practically possible to perform inspections is during outages, when every day means loss of income from electricity production. With this in mind it is easy to see that the plant does not want to make any repairs that are not absolutely necessary. There are two erroneous assessments an operator can make of a component, he can make a *false call* or he can make a *miss*. The feedback he gets from having made a *false call* is relatively immediate. The operators can feel to be met with the attitude that they are the ones that "make the cracks appear". When replacing or repairing the component, the plant can notice that there was no flaw present. When the operator has made a *miss,* the component might still work, at least for a while. Thus, he either receives negative feedback soon (*false call*), or he might get negative feedback later (*miss*). The more distant in time a possible negative consequence is, the smaller importance people tend to give it. In economics this psychological principle is formalized in the discounting of future value. This means that a negative event distant in time is not as dangerous, economically, as the same event now. The discounting of future consequences can result in the operator preferring (or trying to please others, who supposedly prefer) *misses* in front of *false calls.*

Feedback is important for operators to gain experience, and for inspection performance to be good. When performing inspections in nuclear power plants, it is important that the operators are motivated to perform well, and that they receive feedback on their performance. It is most important to receive feedback when the operator is not certain that he is correct. The feedback should be correct, and it should not be delayed too long. By having operators work in teams of two, they are able to receive feedback from their teammate. Another source of feedback is the foreman (see Norros, 1998). Although the foreman is not with the operator all of the time, the operator gets social support and can have the possibility to double-check "tricky" assessments. To improve inspection performance, it is important for the organization (management) to support the operator. That is to make him intrinsically motivated, so that he does a good work for the right reasons.

In the Spencer & Schurman (1995) study, teams of two operators, performed better and worked faster than single operators. However, the time per component, measured in man-ours, was higher for the teams. Consequently, the radiation doses per man would increase when working in radioactive environments. On the other hand, these findings can have the following positive implications:

- More hits
- Less double-checks
- Possibly, less man-hours per hit

Furthermore, the operators in the Wheeler et al. (1986) study considered the opinion of other operators to be the best aid. Inspection in teams would give the operators feedback more direct, and more frequent. Because many NDT activities are not directly supervised, there is a potential for higher number of errors (c.f., Norros, 1998). Furthermore, Berglund & Lindberg tried to improve the accuracy and reliability of inspections by combining the (post facto) results of operators in pairs (Berglund & Lindberg, 1990). The performance of the operators was categorized as good, average, and poor, a total of 55 pairs were combined. Results showed that the performance of operators, whose individual performance was average or poor, improved by being combined with another operator's. The same effect was found even though two operators with poor performance were compared.

According to Norros & Kettunen the results of their (1998) study indicate that the strong division of tasks may promote differences in the attitudes and actions of the operators. To avoid this differentiation between operators Norros & Kettunen (1998) suggest exchange of experience among the operators e.g., in the form of teamwork.

3.3. The work conditions

What is beneficial for the quality of an inspection in theory; is not necessarily beneficial for the operators' work-situation, and thereby not beneficial for the quality of inspection, in practice. For example, by not wearing protective clothing the operators movement is not restricted, this can, in theory, improve the operators performance. However, not wearing protective clothing, in practice, would probably make the operator feel uneasy in radioactive environments, with a subsequent drop in performance. This is hypothetical example is similar to the case reported by Dahlgren & Skånberg (1993), where the performance of the operator was sub-optimal as a consequence of his feelings towards the inspection environment (he left the procedure, and failed to detect several cracks).

There are a number of problems in the work environment of NDT inspectors. They include bulky or difficult to use equipment and any potential source of danger or perceived danger (e.g., radiation, hot surfaces, heights, and confined spaces). As much as possible, the operator should be allowed enough control to make the environment as comfortable as possible. If the task to be performed is simple (i.e. not so demanding in effort/attention), environmental factors do not play as significant a role in performance. Furthermore, with experience, the performance of the operators is not so much affected by the fear of radiation (Wheeler et al., 1986). This is probably not true for high levels of radiation. The radiation, however, puts a limit on the time allowed at each particular component.

However, as the task becomes increasingly more complex, environmental factors affect performance. Manual UT is a very complex task and the performance depends on the sensory, perceptual, cognitive, and often motor skills of the person who is performing it.

3.4. Qualification of operators

The role of qualification is to ensure that (a) the equipment used (hardware and software), (b) the procedure used, and (c) the personnel performing the inspection; are capable of detecting (and determining the size of) relevant defects residing in the material of interest (European Commission, 1996). A qualified system (qualified personnel following a qualified procedure using qualified equipment) does not automatically have a high reliability. That is, just because all parts of the system are as good as possible, does not necessarily result in the system working optimally (c.f., the Gestalt psychology axiom - the organized whole is more than the sum of its parts). To achieve a high reliability in an inspection, other aspects must be considered as well (cf., "-no chain is stronger than its weakest link"). Especially important is the human factors aspect, since humans designed all parts of the system and often are (involved in) the interface between the parts of the system. It is also humans who have the final say if an indication should be assessed as a defect or not. This leaves room for human error in many steps, see Reason (1990) for further details.

"All decisions steps related to combination and interpretation of the results..." "...should be written down in a clear, logical, and traceable manner. This will minimize the extent that the results depend on the experience of the evaluation expert" (ENIQ, 1997 p. 25). In real life, however, an operator's experience clearly does influence his performance. For example, in the case of qualification, the operators know that their performance is being tested. This can make them try their very best. Since they know that there must be defects in the components they inspect under the qualification, the operators do what they can to find them. Wheeler et al., (1986) report that several of the operators indicated that the way of inspecting used to pass the qualification was unrelated to the way of inspecting they would use during an actual inspection.

If the operators can choose to not follow the procedure under a qualification, they might do the same under an actual inspection when they have *prior probabilities*. That is, an operator might find a flaw if he is convinced that there is one present, even if it means that he have to violate the procedure in order to prove it. It also stands to reason that, if an operator does not expect to find flaws he might be less observant on indications of the opposite.

4. Suggestions for further studies

The findings presented in the present review are interesting and sometimes contradicting. With earlier findings in the area of human performance in NDT as a start, and by applying an MTO-perspective, the research in this field can move ahead. As has been said in the research reviewed (e.g., Doctor, 1995 and Harris, 1992), the next step of research require a proper experimental design in order to reach meaningful results. Below, themes for research important to consider in order to improve the quality of inspections in nuclear power plants are presented.

Can the assessment of indications be made easier? By analyzing how operators make assessments, and relating their decision criteria to the inspection performance of the operators, the inspection assessments may be standardized at an optimal level. The importance of following a well-planned procedure have been shown (e.g., Harris, 1992). By analyzing the procedures, and re-designing them applying an MTO-perspective, the procedures might be optimized. As a consequence, the operators may better comply with the procedure. In further studies, examining what decision strategies and cues the operators use to make their assessments, can lead to finding methods on how to guide the decision making of the operators. Furthermore, by storing ambiguous indications for later review, there is room to decrease the mental pressure on the operator, thus, keeping him in "good spirits".

Experience is important for inspection performance, but what separates the experienced operator from the novice? In experiments it is possible to see what makes an inspection good or poor. Furthermore, by putting emphasis on the important information, the level of "noise" presented to the operator is reduced, which should improve inspection performance. Also, training and education can be improved by utilizing findings of such studies.

Findings presented in the present review suggest that the inspection performance is affected by the physical environment (e.g., Taylor et al., 1989; PISC III, 1994). However, no correlations with single environmental factors have been reported. Pond et al. (1998) suggest a shift in emphasis away from research on PSF (e.g., heat, humidity) and towards organizational factors. However, the present authors believe that by applying knowledge from the area of workplace ergonomics to the specific situation of the operators, studies may reach further results. The special circumstances for this particular occupation (NDT inspectors) require proper experimental studies to also examine the importance of PSFs on inspection performance.

The operator's *ability* to perform an inspection is proven in a qualification. In order to perform a high quality inspection, the operator also requires the *opportunity* to do his job well. Both *psychosocial* and *physical* opportunity is required. In the present review the importance of the operator being (intrinsically) motivated has been stressed. Further studies on the importance of the psychosocial environment are needed.

Findings presented in the present review imply that the use of ergonomically designed equipment may enhance inspection performance, especially in the field of manual UT. Typically, all equipment that require the operator to focus on that particular equipment in order to use it, takes attention resources away from the primary task i.e., performing the inspection). It is probable that non-ergonomical equipment can lead to unnecessary fatigue in the operators. *Mentally* by depleting the attention resources, and *physically* by straining the hand and arm. Further studies can determine the effect non-ergonomical equipment have on inspection performance.

The importance of organizational characteristics reflected in attitudes and motivation of the operators have been shown to play a role for the inspection performance (e.g., by Norros, 1998). Also, the role of the management and the organization in motivating and supporting the operators has been discussed. Furthermore, findings presented in the present review, suggest that implementation of teamwork can result in an overall improvement of inspection performance. In further studies it is important to find out what implications organizational characteristics (e.g., teamwork) have on inspection performance.

Finally, by hiring the personnel most apt for inspection, the quality of inspection might be secured. Personnel can be selected in a number of ways, and with the help of a number of tests, both general and specific. For example, the dynamic inspection aptitude test (DIAT) that tests the potential of becoming a good (manual UT) operator (Harris & Spanner, 1998). According to Jack Spanner Jr. (personal communication), the DIAT predicts with a high degree of accuracy, who has the potential of becoming a good (manual UT) operator. In further studies, it is very important to study what implications different instruments for the selection of personnel have on the quality of inspection.

References

Allnut, M.F., Human factors in accidents. *British Journal of Anaesthesia,* **59**, 856-864. 1987.

Ammirato, F., Jeong, P., Lucas, S., Scott, A., & Sanders, B., NDE of cast stainless steel in nuclear plants. *Nondestructive examination.* March. 1989.

Bailey, R.W., *Human Performance Engineering: Using Human Factors/Ergonomics to Achieve Computer System Usability.* Englewood Cliffs, NJ: Prentice Hall. 1989.

Berglund, U. & Lindberg, S., *Human Reliability of Non Destructive Testing.* SKI-report 901207/ub/sl. 1990.

Bray, D.E. & McBride, D., *Nondestructive Testing Techniques.* New York: John Wiley & Sons, Inc. 1992.

Bray, D.E. & Stanley, R.K., *Nondestructive Evaluation: A Tool in Design, Manufacturing, and Service.* Florida: CRC Press 1997.

Christensen, J.M., The Human Factors Profession. In G. Salvendy (Ed.) *Handbook of Human Factors.* John Wiley & Sons; New York 1987.

Dahlgren, K. & Skånberg, L., *Felaktigt utförd provning av rörböjar i system 321.* Utredningsrapport, SKI. 1993.

Dickens, J.R., Human factors in nondestructive testing. In Bray, D.E. & McBride, D., *Nondestructive Testing Techniques.* New York: John Wiley & Sons, Inc. 1992.

Dickens, J.R. & Bray, D., Human Performance Considerations in Nondestructive Testing. *Materials Evaluation.* Vol. 51, No. 9 1033-1041. 1994.

Drury, C.G., & Fox, J.E. (Eds.). *Human reliability in quality control.* New York: Halstead, 1975.

ENIQ, *Report No. 2. EUR 17299 EN* 1997.

Enkvist, J. Edland, A., & Svenson, O. *Unpublished manuscript.* 1998.

European Commission, *Common position of European regulators on qualification of NDT systems for pre- and in-service inspection of light water reactor components. EUR 16802 EN* 1996.

Eysenck, M.W., Arousal, learning, and memory. *Psychological Bulletin,* **83**, 389-404. 1976.

Festinger, L., *Conflict, Decision and Dissonance.* Stanford University Press, California 1964.

Firth, D., Demonstrating Effective Inspection. In W. Gardener (Ed.) *Improving the Effectiveness and Reliability of Non-Destructive Testing.* 1992.

Frankehaeuser, M., A Psychobilogical Framework for Research on Human Stress and Coping. In M.H. Appley & R. Trumbull (Eds.) *Dynamics of Stress: Physiological, Psychological, and Social Perspectives.* pp. 101-116. Plenum Press: New York. 1986.

Grandjean, E., *Fitting the Task to the Man: A Textbook of Occupational Ergonomics.* Taylor & Francis 1988.

Halmshaw, R., *Non-Destructive Testing.* Edward Arnold: London 1991.

Harris, D.H., Effect of Human Information Processing on the Ultrasonic Detection of Intergranular Stress-Corrosion Cracking. *Materials Evaluation.* **48** April. pp. 475-480. 1990.

Harris, D.H., *Effect of Decision Making on Ultrasonic Examination Performance.* Electric Power Research Institute: Palo Alto, California. EPRI Report TR-100412 1992.

Harris, D.H., *Dynamic Inspection Aptitude Test, User s Guide.* EPRI TR-106304 Palo

Alto, CA: Electric Power Research Institute 1996.

Harris, D.H & McCloskey, B.P. *Cognitive correlates of ultrasonic inspection performance*. EPRI Report NP-6675, 1990.

Harris, D. H. & Spanner, Jr., J.C. *Dynamic Inspection Aptitude Test (DIAT): A Measure of Aptitudes Required for Nondestructive Testing*. 1998.

Hoogstraate, H., EN 473/45013 Certificates are not the Ultimate Answer to Questions about Non-Destructive Testing. *Proceedings 7th. ECNDT Copenhagen May 26-29* 1998.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.)., *Judgments under uncertainty: Heuristics and biases*. New York: Cambridge University Press. 1982.

Karimi, S.S., *Human Factors Affecting the Performance of Inspection Personnel in Nuclear Power Plants*. Electric Power Research Institute: Palo Alto, California. EPRI Report NP-6000. 1988.

Kecklund, L. *Samspelet Människa-Teknik-Organisation (MTO) och säkerhet i tekniska system*. (Manuscript) 1998.

Keinan, G., & Friedland, N. Dilemmas concerning the training of individuals for task performance under stress. *Journal of Human Stress*, **10**, 185-190. 1984.

Kettunen, J., *Beliefs conserning the reliability of nuclear power plant in-service inspections*. STUK-YTO-TR 121. 1997.

Kettunen, J. & Norros, L,. *Human and organizational factors influencing the reliability of non-destructive testing. An international literary survey*. STUK-YTO-TR 103, 1996

Kroemer, K. & Grandjean, E., *Fitting the Task to the Human: A Textbook of Occupational Ergonomics*. Taylor & Francis: London 1997.

Mahan, R.P., Continuous Performance and Operator Judgment Implications for the Automated Work Place. In W. Karwowski & M. Rahimi (Eds.) *Ergonomics of Hybrid Automated Systems II*. Amsterdam: Elsevier. pp. 741-746 1990.

Mordfin, L., Nondestructive Evaluation. In J.F. Young & R.S. Shane (Eds.) *Materials and Processes*. Marcel Dekker: New York 1985.

Norros, L. *Human and organisational factors in the reliability of non-destructive testing (NDT)*. Final report, VTT Symposium No. 190. 1998

Norros, L. & Kettunen, J. *Analysis of NDT-inspectors working practices*. STUK-YTO-TR 147. 1998.

PISC III,. *Human Reliability in Inspection, Final Report on Action 7 in the PISC III Programme, PISC III Report 31*. Nuclear Energy Agency 1994.

Pond, D.J., Donohoo, D.T., & Harris, Jr., R.V., *An Evaluation of Human Factors Research for Ultrasonic Inservice Inspection*. NUREG/CR-6605, PNNL-11797. 1998.

Reason, J.T., *Human Error*. New York: Cambridge University Press. 1990.

Salvendy, G., *Handbook of Human Factors*. John Wiley & Sons: New York 1997.

Schwartz, D.R., & Howell, W.C., Optional stopping performance under graphic an numeric CRT formatting. *Human Factors*, **27**, 433-444. 1985.

Seligman, M.E.P., *Helplessness*. W.H. Freeman: San Fransisco 1975.

Skånberg, L., *Bedömning av tillförlitlighet vid återkommande provning av huvudcirkulationskretsar*. Report (Swedish Nuclear Inspectorate) SKI TR 91:28 October. 1991.

Skånberg, L., *Kvalificering av OFP-sytem*. SKI Rapport 94:25 [*The qualification of NDT systems*. Swedish Nuclear Inspectorate Report 94:25(In Swedish)] 1994.

Spanner, J.C., Badalamente, R.V., Rankin, W.L., & Triggs, T.J., *Human Reliability*

Impact in Inservice Inspection. Vol 1. U.S. Nuclear Regulatory Comission. Washington, USA 1986.

Spencer, F. & Schurman, D., *Reliability assessment at airplane inspection facilities, Vol III: results of an eddy current inspection reliability.* Report No. DOT/FAA/CT-92/12, III

Stokes, A.F., Belger, A., & Zhang, K., *Investigating factors comprising a model of pilot decision making II: Anxiety and cognitive strategies in expert and novice aviators.* Savoy, IL: University of Illinois aviation research lab technical report no. ARL 90-8/SCEEE 90-2. 1990.

Swets, J.A., *Signal detection theory and ROC analysis in psychology and diagnostics: collected papers.* Mahwah, New Jersey: Lawrence Erlbaum Associates. 1996.

Svenson, O., Process Descriptions of Decision Making. *Organizational behavior and human performance.* **23**, 86-112 1979.

Taylor, T.T., Spanner, J.C., Heasler, P.G., Doctor, S.R. & Deffenbaugh, J.D., An Evaluation of Human Reliability in Ultrasonic In-Service Inspection for Intergranular Stress-Corrosion Cracks through Round-Robin Testing. *Materials Evaluation, 47,* 338 - 344. 1989.

Trumbull, R & Appley, M.H., A Conceptual Model for the Examination of Stress Dynamics. In M.H. Appley & R. Trumbull (Eds.) *Dynamics of Stress: Physiological, Psychological, and Social Perspectives.* pp. 21-45. Plenum Press: New York. 1986.

Tulving, E.. Human Memory. In P. Andersen, Ø. Hvalby, O. Paulsen, & B. Hökfelt (Eds.) *Memory concepts - 1993. Basic and clinical aspects.* Elsevier Science Publishers 1993.

Tversky, A., & Kahneman, D., Judgment under uncertainty: Heuristics and biases. *Science,* **185**, 1124-1131. 1974.

Warm, J.S., An Introduction to Vigilance. In J.S. Warm (Ed.) *Sustained Attention in Human Performance* pp. 1-14 1984.

Webster, C., Some Individual Psychological Factors Leading to Error in Ultrasonic Testing. *British Journal of Non-destructive testing,* **31**, 680-682 1989.

Wheeler, W.A., Rankin, W.L., Spanner, J.C., Badalamente, R.V., & Taylor, T.T., *Human Factors Study Conducted in Conjunction with a Mini-Round Robin Assessment of Ultrasonic Technician Performance.* Batelle Human Affairs Research Centers. 1986.

Wickens, C.D., *Engineering Psychology and Human Performance.* : Harper Collins: New York. 1992.

Wickens, C.D., Stokes, A.F., Barnett, B., & Hyman, F., The effects of stress on pilot judgment in MIDIS simulator. In O. Svenson & J. Maule (Eds.), *Time pressure and stress in human judgment and decision making.* Cambridge, UK: Cambridge University Press. 1991.

Wiener, E.L., Vigilance and Inspection. In J.S. Warms (Ed.) *Sustained Attention in Human Performance.* New York: John Wiley & Sons. Pp.207-246 1984.

Wiener, N., *The human use of human beings; Cybernetics and Society.* Boston: Houghton Mifflin Company 1950.

Yerkes, R. M. & Dodson, J D., The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurological Psychology,* **18**, 459-482. 1908.

Zhang, K., & Wickens, C.D., Effects of noise and workload on performance with two object displays vs. a separated display. *Proceedings of the 34th annual meeting of the Human Factors Society.* Santa Monica, CA: Human Factors Society. 1990.

Acknowledgments

The authors would like to thank the Swedish Nuclear Inspectorate (SKI) for funding the study. Furthermore we wish to acknowledge the efforts of Kerstin Dahlgren, Arne Eriksson, Conny Eriksson, Kent Gebertsberger, Maj-Lene Hedenborg, Anna Lekberg, Martin Ryf, Lars Skånberg, Jack Spanner Jr., Henry Stephens, Magnus Strand, Tom Taylor, and Tommy Zetterwall, who have all contributed valuable comments, assistance and information. Finally, we thank all the operators who have taken part in the pilot study at SQC Kvalificerings-centrum AB.

List of abbreviations

DIAT	Dynamic Inspection Aptitude Test
ENIQ	European Network for Inspection Qualification
EPRI	Electric Power Research Institute
ET	Eddy Current Testing
FAA	Federal Aviation Administration
IGSCC	Inter Granular Stress Corrosion Cracking
ISI	In-Service Inspection
KSA	Knowledge Skill and Ability
MTO	Man Technology Organization
NDT	Non-Destructive Testing
NUREG	Nuclear Regulatory Board
PISC	Programme for the Inspection of Steel Components
PNNL	Pacific Northwest National Laboratories
PRS	Perceptual Priming System
PSF	Performance Shaping Factors
SKI	Swedish Nuclear Inspectorate
STSS	Short Time Sensory Storage
UT	Ultrasonic Testing