ABSTRACT

Nuclear power industry has the responsibility to manage and dispose of all radioactive waste from its plants. Finnish Posiva and Swedish SKB are leading in the world in the development of disposal of spent nuclear fuel. Their method consists of encapsulating spent nuclear fuel in copper canisters and depositing them in the bedrock at a depth of about 500 meters for the next 100,000 years, leaving the radioactivity to decrease naturally through the decay of the radioisotopes in it. The copper canisters, consisting of a copper tube, a lid and a bottom (which make the outer shell) and an insert made of a cast iron, need to be inspected for their structural integrity to ensure no critical defects are present in the materials and welds that could lead to a leakage of the waste into the environment. Data acquired by 4 different non-destructive testing (NDT) methods (i.e. UT, ET, RT and VT with a remote camera) are evaluated by skilled human operators and therefore could be subject to human error.

Human Factors approach lies in identifying potential errors made by the human, their causes and ways of preventing them. A customized Failure Modes and Effects Analysis (FMEA) was conducted to anticipate possible human failures during the data evaluation. The results led to designing several experiments (e.g. diffusion of responsibility within the 4-eye principle, over trust in automated systems) which are being experimentally tested in ongoing projects. The results are expected to lead to the optimization of the procedures followed by the NDT operators and consequently to the improvement of the overall NDT reliability.

INTRODUCTION

Nuclear power industry has the responsibility to manage and dispose of all radioactive waste from its plants. Nuclear waste management experts in Finland (Posiva Oy) and Sweden (Svensk Kärnbränslehantering AB, SKB) have been conducting research for several decades to find a long term solution to store the spent nuclear fuel safely, avoiding the risks to the society and the environment. Their 3 barriers-method (KBS-3 method) consists of encapsulating spent nuclear fuel in copper canisters and depositing them in the bedrock at a depth of about 500 meters, and additionally protecting them with a buffer of bentonite clay intended to protect against corrosion and movements in the rock. Stored like this, the canisters should withstand untouched for the next 100,000 years, leaving the radioactivity to decrease naturally through the decay of the radioisotopes in it [1][2]. See Figure 1 for an illustration of the SKB’s deposition plan.
The copper canisters (Figure 2), consisting of a copper tube, a lid and a bottom (which make the outer shell) and an insert made of a cast iron, need to be inspected for their structural integrity to ensure no critical defects are present in the materials and welds that could lead to a leakage of the waste into the environment. This is done using non-destructive testing (NDT) methods, such as ultrasound, radiography, eddy current etc.

The probability that the NDT-system will find all critical defects is defined within the term “reliability”. Reliability, as defined by the IEEE [3], is the ability of a system or a component to perform its required functions under stated conditions for a specified period of time.

According to the modular model [4], reliability of the NDT depends on 3 modules: (1) intrinsic capability of the NDT system, (2) application parameters and (3) human factors.
Human Factors in NDT

The problem of varying performance in non-destructive testing has typically been approached by improving the equipment used and by changing the procedures; but very few resources have been invested in research in the human factors field. In order for an inspection to be reliable the whole system, as well as its parts, needs to be reliable (equipment, procedure and personnel). However, the largest source of performance variation can be found in an operator. After all, it is the operators who interpret the signals provided by the equipment [5] [6]. Therefore, in order to better understand the process of the NDT inspection, and through this, how to improve NDT reliability, the impact and the importance of human factors needs to be addressed.

The human factors approach relies on understanding the properties of human capability and limitations under various conditions and the application of that knowledge in designing and developing safe systems [5].

In this paper a new approach to the human factors in NDT will be presented. The new methodology includes identifying potential errors made by the human, causes of those errors and ways of preventing them. Hypotheses derived from this analysis are being empirically tested in ongoing experiments, and their results are expected to lead to the optimization of the entire NDT-system.

HUMAN FACTORS APPROACH

Within the projects “NDT Reliability V: Human Factors” (SKB-Posiva-BAM) [6] and “POD of Posiva's EB-Weld and the Effect of Human Factors in NDT of EB-Weld” (Posiva-BAM) [2] the task was to study human factors in the application of 4 different NDT methods: UT-ultrasonic testing (UT; of the copper tube (SKB) and the EB-Weld (Posiva)), radiography (RT), eddy-current testing (ET) and visual testing (VT) with a remote camera (Posiva).

A typical automated NDT inspection of the copper canister parts consists of two tasks: collection of the data with an automated system (data acquisition), and evaluating the collected data with the help of a specially designed software, including reporting (data evaluation). The focus of this part of the human factors investigation was put on the data evaluation.

The approach relies on four consecutive steps:

1. Evaluation (of the current condition)
2. Anticipation (of potential errors involving a human operator)
3. Experiments (experimentally testing all the hypotheses that might have come out from the evaluation and anticipation phase)
4. Optimization (using the results of the previous 3 steps as a tool to optimize the NDT system)

Evaluation

Evaluation refers to observing and evaluating the current state of all of the methods concerned. The data evaluation part of 4 different NDT methods was carefully observed. Each step in the procedure was written down which resulted in a Hierarchical Task Analysis, HTA [8] – a thorough description of each step taken during the data evaluation.

Task Analysis is one of the most commonly used human factor methods. It is used to help the analyst to understand and represent human and system performance in a particular task or a scenario. Task analyses are used for understanding the required human-machine and human-human interactions by decomposing tasks or scenarios into component task steps or physical operations. Hierarchical Task Analysis (HTA) describes an activity under analysis in terms of hierarchy of goals, sub-goals, operations and plans. The end result is an exhaustive description of task activity [8].
Even though the HTA helped in the evaluation part, it could not be used as an anticipation tool, since the NDT methods are under development and therefore are constantly being changed. HTA being very complex and time consuming, another way of identifying potential human errors needed to be found. For that purpose a Failure Modes and Effects Analysis [9][10][11] was used.

Anticipation

Anticipation refers to foreseeing, based on expert knowledge and experience, which failures could occur as a result of human error during different stages of the NDT data evaluation.

For this purpose a group of internal NDT experts (4-5 per method), led by external human factors specialists, gathered and conducted a customized Failure Modes and Effects Analysis (FMEA).

Human Factors Customized FMEA

Failure Modes and Effects Analysis (FMEA), widespread in the automotive, aerospace and other safety critical industries, is a classical system safety analysis technique [9]. It is known to be a systematic procedure for the analysis of a system to identify the potential failure modes, their causes and effects on system performance. The analysis is successfully performed preferably early in the development cycle so that removal or mitigation of the failure mode is most cost effective [10].

For the purposes of the two NDT reliability projects [6][2], where human factors were of interest, a customized human-FMEA was used. The method had to be slightly changed due to the fact that the aim of the analysis was to study possible human errors and omissions and not component failures [11].

The analysis was carried out in the following steps:

1. Decomposition of the task into sub-tasks
2. Definition of aims for the sub-tasks
3. Identification of possible failures/errors
4. Consideration of potential causes and effects of failures
5. Identification of existing barriers
6. Identification of potential preventive measurements
7. Assessment of error probability (EP), relevance of effects (R) and detection probability (DP)
8. Calculating of risk priority (RP = EP x R x DP; see above)

Results of the customized human-FMEA

Altogether 5 workshops were organized to complete the FMEA for 4 different NDT methods (FMEA for the UT was done separately at the SKB for the data evaluation of the ultrasonic inspection of the tube, and at Posiva of the EB-Weld (electron-beam-welding, EBW)). The results of all 5 workshops could be summarized due to the similarities of the human tasks in all of them. Table 1 shows possible errors that might happen in different phases of the data evaluation process.
Table 1- Typical errors during different phases of data evaluation summarized for all 4 NDT methods (ultrasound (UT), radiography (RT), eddy current (ET) and visual testing (VT) with a remote camera)

<table>
<thead>
<tr>
<th>SUB-TASKS</th>
<th>UT, RT, ET, VT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation</td>
<td>Choosing wrong area for evaluation; Mistakes in positioning the zero point; Opening the wrong files or reporting sheets; Mistakes in settings (e.g. setting the wrong sensitivity level); Missing or faulty data validity check etc.</td>
</tr>
<tr>
<td>Identification of defects¹</td>
<td>Missing a defect; Missing an indication or an area during dynamic or static observation, etc.</td>
</tr>
<tr>
<td>Characterization of defects²</td>
<td>Incorrect defect type; Misinterpretation, e.g. two indications vs. one indication, etc.</td>
</tr>
<tr>
<td>Sizing &amp; Location</td>
<td>Mistakes in sizing &amp; localization; Not seeing the entire size of the defect; Misplacement of the defect, etc.</td>
</tr>
<tr>
<td>Decision Making</td>
<td>False acceptance/ rejection of the part; False calls; Misinterpretation of the reference indications; Misjudging the origin of the indication (geometry vs. real defect), etc.</td>
</tr>
<tr>
<td>Other</td>
<td>Typing errors; Unstructured or not-understandable reporting, etc.</td>
</tr>
</tbody>
</table>

Proposed causes of the listed errors include:

- **Procedures**, i.e. shortcomings in the procedures, misinterpretation of the procedures
- **Software limitations**, i.e. resolution and size of the screen, different scaling of the different views, speed of dynamic observation
- **Individual differences between inspectors**, i.e. sensitivity to colors, subjective assessment criteria, reduced attention or concentration, confirmation bias³ (premature hypothesis), competency, poor writing skills (for reporting), heuristics⁴, (in)experience
- **Geometry/technology**, i.e. characteristics of the defects, misinterpretation of the geometry, various technology shortcomings and technical difficulties, ergonomics
- **Other**, i.e. confusion, lapses (accidentally skipping areas, choosing a wrong set up etc.), interruptions, monotony, capacity overload (too many information)

The common consequences of occurrence of these errors are:

- Missing inspection areas and/or missing defects
- Wrong decision making: false calls; false rejection/acceptance of the part, false recommendation

The existing prevention methods consist of relying on the experience of the inspectors, improving the training and the procedures.

Most of the listed errors are either hard or almost impossible to detect if one doesn’t suspect he had make a mistake.

1 Finding defects
2 Determining the defect type, distinguishing between a defect and a non-defect indication
3 Confirmation bias is a tendency to seek confirmation rather than disconfirmation of what we already believe [11]. *For example, missing a defect on a preconception that there are no defects present.*
4 A heuristic is a mental shortcut that allows people to solve problems and make judgments quickly and efficiently. These informal, intuitive, speculative strategies sometimes lead to effective solutions, but sometimes not (Sternberg, 2000, in [11])
The following prevention methods (possible barriers) had been suggested during the workshop:

- **Software solutions**, i.e. introduce warnings when an area is not being inspected or when the sensitivity is set too high/low, define screen view parameters (resolution, size, distance from the screen etc.), alarms against mistyping or choosing a wrong start/end point

- **(Semi)Automation**, i.e. semi-automated detection of indications in static mode (+ confirmation by a human operator), automatic transfer of size and location details into an excel sheet, automated cross check of the data on the screen and the data in the report sheet

- **Procedures/instructions**, i.e. increasing the awareness of some potential errors, introducing a B-scan cross-check when setting a channel

- **Training**, i.e. not only in the technical sense, but also in awareness of possible cognitive biases, group effects; mistake awareness

- **4-eye principle**, i.e. evaluation performed by two independent operators (no common norms, knowledge, group processes), introducing random checks by supervisors (frequency would depend on the frequency of error-occurrence)

- **Aids**, i.e. introducing interpretation aids, i.e. a defect catalogue (in form of visual representations of possible known defects)

- **Other**, i.e. shorter time periods of dynamic observation, plausibility check in reporting

Calculating the risk priority (a function of error probability (EP), relevance or significance of effects (R) and probability of detection (DP) of those errors) led to a risk priority ranking of the subtasks. Table 2 shows, according to the opinion of experts, three most critical tasks for each NDT method separately.

<table>
<thead>
<tr>
<th>Tube (SKB)</th>
<th><strong>High risk prioritized tasks</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>UT</td>
<td>Dynamic observation; setting the views and zooming; identification of defects</td>
</tr>
<tr>
<td>Weld (Posiva)</td>
<td>UT Sizing and localization; characterization; identification of defects</td>
</tr>
<tr>
<td></td>
<td>RT Identification; sizing and localization of defects; image adjustment</td>
</tr>
<tr>
<td></td>
<td>ET Characterization of defects (determining the defect type)</td>
</tr>
<tr>
<td></td>
<td>VT Identification; characterization; sizing and localization of defects</td>
</tr>
</tbody>
</table>

It should be noted that the FMEA analysis is a subjective method. The results of these analyses would probably differ if they were to be repeated with another group of experts. Therefore the results, i.e. typical errors, causes of these errors, their consequences, as well as the suggestions for possible preventive measurement and the risk priority ratings can only be applied to SKB’s and Posiva’s NDT practices, and are subdue to change as the methods develop and the procedures change.
Experiments

Some of the results or suggestions from the FMEA have raised questions regarding the best way to implement them. Is the 4-eye principle a good solution to reduce human error? Can the applied software be improved and how? How can the operators learn from errors that have happened? Etc.

As an answer to those and many other possible questions, this approach suggests an empirical study, i.e. experimentally testing hypotheses. Several of these questions are under evaluation in ongoing projects.

Social Loafing within the 4-Eye Principle

First one is raising the question about the implementation of the so-called 4-eye principle, or in other words, having two operators conduct the same task.

Assigning several individuals to solve the same problem consecutively, therewith avoiding failures of single individuals is known as the redundancy principle, which is assigned from technical redundancy. Typically applied in high-reliability organizations, the redundancy principle is used to increase the system’s overall reliability in monitoring the performance of technical processes [13]. Technical components or systems are called redundant if they are independent. However, the redundancy principle is not always as straightforward as it might seem. The main predisposition for the principle to fulfill its purpose is that the individuals involved need to work parallel to and completely independent of each other [14] which is not often the case. According to Sagan [15], individuals in redundant systems are normally aware of each other, and one can reasonably assume that knowing someone else is carrying out the same task might counteract the benefits expected from human redundancy. A number of studies (e.g. [16][17][18]) has shown that individuals often excerpt less effort when working on tasks collectively as compared to working alone, a phenomenon referred to as social loafing. However, under some conditions (e.g. expecting other team members might perform insufficiently), people might actually work harder in order to compensate for others in the group, an effect referred to as social compensation [16].

Occurrence of these effects in high-reliability organizations could represent a great risk to safety. The present study will attempt to investigate these effects during 4-eye principle applied to ultrasonic data evaluation. An experiment had been designed to test: 1) the differences between working alone and working in a team; 2) the difference between the position (working as first as opposed to working as the second operator in the team) and 3) the difference between not knowing who the first operator was as opposed to thinking that the first operator is very experienced. According to Swain and Guttman [19], checker’s familiarity with the operator, as well as his knowledge of the other operator’s technical level are some of the most influencing factors on human redundancy. Clarke [20] adds that a checker might fail to perceive an error because of a belief in the competence of a colleague. Therefore, the hypothesis is that the operator working in a team with a very experienced operator will loaf. If this hypothesis is confirmed, it will lead to the conclusion that the 4-eye principle needs to be implemented in a way that both operators work completely independently of each other in order to profit from the human redundancy.

Automation Bias in the Evaluation of the Eddy Current Testing Data

Posiva’s efforts to help the operators with the evaluation of the eddy current testing (ET) data by aiding the operator to make decisions faster and more reliably have resulted in higher automation of the ET data evaluation software. Though automation has shown to help the decision making process, it brings problems of a different nature.
Along with the increasing capacity and reliability of automated systems, the role of humans as operators of technical systems has changed considerably. Their main task today is what is referred to as supervisory control of automated systems. It involves monitoring and controlling what the automation is doing. Although this has a number of benefits, in terms of speed and accuracy of processing, it has also led to new error sources and new risks, among which is also what is referred to as the “misuse of automation”, usually defined as an uncritical reliance on the proper function of an automated system without recognizing its limitations and the possibilities of automation failures [21]. The behavioral aspect of this error lies in an insufficient monitoring and checking of automated functions. This phenomenon is known as the “automation-induced complacency”, or just “complacency”. Important performance consequence of this behavior includes an elevated risk that operators will fail to detect and manage automation failures in due time. The term complacency has mostly been used for monitoring tasks in aviation, so Mosier and Sitka [21] introduced the term “automation bias” in use of the automation aids. Following an aid (e.g., software) without verification, in cases of contradictory information, is a result of over-trust in the automated system. It has been shown that complacency effects often appear in interaction with automated systems, which are considered as highly and constantly reliable [23], and especially in situations where the overall task demands are high [23]. Most common areas in which automation bias occurs are aviation, nuclear power plants, chemical plants, intensive care units, etc. [25]

Will over-trusting automation (i.e., automation bias) lead to missing a critical defect? The aim of this study is not to criticize automation, but to raise suspicion in the automation. Automation bias could be expected in a condition where participants are led to believe that the system is very reliable. This hypothesis is being tested in an ongoing experiment where the reliability of the automated system (i.e., the software’s reliability) is varied (low reliability vs. high reliability). If this hypothesis is confirmed, awareness of this problem should find its way in the training and in the procedures for the NDT personnel. Even though complacent behavior cannot be completely excluded, Bahner, Hüper and Manzey [26] have shown that raising the awareness of possible automation failures during the training can significantly decrease complacency effects and therefore represents a suitable method to reduce the risk of such failures.

Optimization

The results of the FMEA study, as well as the results from the empirical study are aimed to help in the development of the NDT methods and procedures and serve as an optimization tool. The results of the FMEA have shown to be very useful to our clients and have been used in improving the procedures and the software. It is the wish of SKB and Posiva to continue with the human risk assessment using the customized FMEA for the entire NDT system, and not only the data evaluation.

Another optimization tool that had been proposed comes from the praxis: learning from errors.

Learning from errors: Event (surprises) analysis

Instead of a classical event analysis, a systematic analysis of events (defined as occurrences of unexpected, undesirable system states either with negative consequences or without [27]) or surprises is proposed. Since real events are not to be expected yet in this phase of the projects, a focus can be laid on the so-called “surprises” which happen during the inspections with the purpose of learning from these experiences. An in-depth analysis of this kind could be performed using the Safety through Organizational Learning (SOL) methodology, developed by Wilpert & Fahlbruch in 1997 [27]. This methodology is based on an assumption that organization learns through operational experience. Therefore, all organizational weaknesses and latent failures should be identified by permanent monitoring and by systematical analyses of events to later generate adequate measures for the prevention of comparable events. Basic characteristics of SOL are: description of the situation, identification of contributing factors, and reporting of the event.

By openly reporting about existing problems, the organization can learn how to avoid them. The results of this analysis should serve as a “must follow” optimization tool.
CONCLUSIONS

The customized (human) Failure Modes and Effects Analysis has shown to be a useful tool in evaluating a system, such as the NDT methods - identifying potential errors, their causes and consequences, as well as to speculate about potential preventive measures. Several experiments designed to test the hypotheses drawn from the FMEA are being conducted in ongoing projects and are expected to serve as an additional optimization tool. The goal of this human factors approach is transparency regarding existing or potential human-caused problems, acknowledging them and directing effort into changing current faulty practices. FMEA can be repeatedly used in all cycles of the development of NDT-program, and serve as an evaluation (e.g. at the beginning or after a change had been made) and optimization tool.

REFERENCES


