Assessing and Treating Risks in Mechanised NDT: A Human Factors Study

Marija BERTOVIC
E-Mail: marija.bertovic@human-factors-ndt.com

Abstract. Reliability of NDT is affected by human factors, which have thus far received the least amount of attention in the reliability assessments. With increased use of automation, in terms of mechanised testing (automation-assisted inspection and the corresponding evaluation of data), higher reliability standards are believed to have been achieved. However, human inspectors, and thus human factors, still play an important role throughout this process and the risks involved in this application are unknown. The aim of this study was to explore for the first time the risks associated with mechanised NDT and find ways of mitigating their effects on the inspection performance. Hence, the objectives were to identify and analyse potential risks in mechanised NDT and devise measures against them. To address those objectives, a risk assessment in form of a Failure Modes and Effects Analysis (FMEA) was conducted. This analysis revealed potential for failure during both the acquisition and evaluation of NDT data that could be assigned to human, technology, and organisation. Since the existing preventive measures were judged to be insufficient to defend the system from identified failures, new preventive measures were suggested.

Introduction

NDT has to provide reliable results. Only that way it can achieve its goals and serve as a contributor to assuring safe operation of complex organisations with high reliability and safety demand. The potential variability in the inspection results observed in manual NDT presents a risk that NDT may not provide reliable results. In turn, unreliable NDT may not be successful in achieving its purpose and may result in unwanted consequences. As suggested by the Modular Reliability Model [1], human and organisational factors play an important role in reliability, however still to a large extent an unknown one.

This is especially true for mechanised testing, considered by many to be more reliable than manual testing and far less prone to the possibility of human error. This assumption is largely based on experiences from the field and generally higher technical reliability (expressed in terms of probability of detection). However, the understanding of potential influencing factors, their interactions, and the potential risks that can arise during mechanised testing is still missing.

Contemporary human-automation interaction research suggests that increased automation is not only related to benefits, but also to costs—a paradox frequently dubbed as the automation ironies [2] or automation surprises [3]. Those ironies and surprises refer to those elements in the interaction between human users and automation that were not considered by the automation designers, but which can fundamentally change the responsibilities of the human operators of systems and the nature of the cognitive demands, e.g. the need for new skills, retention of old skills for problem solving, loss of situation awareness, different nature of workload, reliance on automation, etc. (e.g. [4], [5]).

Every complex socio-technical system with high safety and reliability demand invests effort into the avoidance of adverse effects that could affect people and the environment. Since erroneous actions are unavoidable—they have happened and will continue to happen—the means have to be found to identify them, to determine the consequences and the seriousness of those actions, to assess the likelihood of their occurrence, and to find ways to reduce either the actions themselves, or their consequences. To do that, Hollnagel [6] suggests a combination of theory of human action, appropriate methods for risk and reliability analysis, and a set of strong principles for the man-machine system design. In other words, to tackle the risks in NDT, we need to a) understand the mechanisms underlying potential failure associated with human operators, i.e. human error, b) utilise methods to identify and analyse potential risks, and c) invest efforts into changing faulty practices by appropriate design, therein applying the knowledge of human factors and man-machine design.

The aim of this study was to identify risks associated with mechanised NDT and generate methods for preventing them. Furthermore, this work was meant to serve as a foundation for further empirical work.

Human error and its contribution to failure

When thinking about risks of potential failure of NDT to detect all critical defects, it is impossible not to mention the notion of human error. As stated by Hollnagel [6]: “To err is human; to understand the reasons why humans err is science” (p. 1). The commonly accepted and widely used definition of human error is that of James Reason [7], who defined it as “all those occasions in which a planned sequence of mental or physical activities fails to achieve its intended outcome, and when these failures cannot be attributed to the intervention of some chance agency” (p. 9). In simple terms, if an action fails to achieve its intended outcome, we talk about human error. Human error typically refers to mistakes, slips, and lapses.

Whereas cognitive psychologists are concerned with the internal psychological or cognitive mechanisms of the mind that are assumed to explain the action, practitioners look at human error mainly as an exacerbating feature. Hence, the term “human error” is widely used to explain human action or an event that happened (observable failure), the cause of a mishap or an accident, or is seen as a symptom of deeper trouble [6], [8], [9]. Considering human error as a cause and as a symptom do not only constitute only two different views, but also two different eras in the approach towards human error—a difference that is still sometimes blurry to the managers of complex systems.

Underlying the first approach is the tendency to assign “blame” to the operators and inspectors at the sharp end of the line for mishaps, events, and accidents. After all, the errors do become obvious at the hands of the person handling the equipment and making the decisions. People are available to be blamed: Since they are working with the equipment, it is probable that the accident would not have happened had the operator not been present. People also have a temporal and a physical relationship with the outcome [9]. However, this approach is nowadays considered as a traditional approach. By concentrating on the individual origins of error the act is wrongfully isolated from its context and, therefore, important features can be overlooked [10]. First, it is often the best people that make the mistakes, and second, the same combination of circumstances can provoke the same errors, regardless of the people involved. In addition, people in high-reliability organisations are generally motivated to do a good job - what they do generally makes sense to them at the time [8]. Therefore, this view is being replaced by the modern systems approach focusing on the underlying conditions that create possibilities for failure, and view human error as a symptom of problems hidden deeper in the system. Efforts are thus invested into the conditions under which people work and ways to prevent the failures [6], [8], [11], [12]. This is achieved by implementing defences. Hence, when adverse events do occur, the question should not be who failed, but rather how and why the defences failed.

To illustrate the difference between the observable failure at the hands of an operator at the sharp end and the underlying causes in the system that may lead to an organisational accident, [11] introduced the terms active failure (human errors and violations that have immediate adverse effects and, through that, a direct impact on the safety of the system) and latent conditions (e.g. poor design, gaps in supervision, undetected manufacturing defects or maintenance failures, unworkable procedures, clumsy automation, shortcomings in training, or less than adequate tools and equipment). Latent conditions arise from strategic and top-level decisions made by governments, regulators, manufacturers, designers and organizational managers and are present in all systems, being an inevitable part of organizational life. They can be present for years, before being combined with local circumstances and active failures to penetrate the system’s many layers of defences. The impact of these decisions spreads throughout the organization forming a distinctive organizational culture, which then results in the creation of error-producing factors within individual workplaces.

Classifications of human error

The manner by which an error or a failure is observed is called error mode or failure mode. This term describes the way a failure occurs and its impact on the equipment or operation [13]. One of the most common classifications includes that into errors of omission and errors of commission, developed by Swain & Guttmann [14] for the purposes of the Human Reliability Analysis. They refer to those events that constitute incorrect human inputs to the system. They are regarded as errors only if they can result in a consequence that might be undesirable for the system, thereby affecting the system reliability and safety. Looking for a way to describe situational and organisational factors that can contribute to failure, i.e. the latent conditions, Reason et al. [15], identified 11 general failure types (GFTs). Both classifications are presented in Table 1.

Error prevention

Typical methods for the prevention of errors include designing the system so that it is simple and easy to use, training, effective warnings that can anticipate a system state that will likely lead to error, and restricting the exposure of the operator to opportunities for error [16]. The attempts to minimise the occurrence of errors are either proactive or reactive in nature. The proactive approach is based on improving the human-system interface. This is most commonly achieved by creating decision aids, improving the training or the procedures, automating features of the system interface, etc. The reactive approach focuses on eliminating the reoccurrence of already occurred errors. The common term used for these error prevention or minimisation techniques is defences or barriers. Installing defences can sometimes even harm the system, because in spite of their original purpose, they can backfire [8], [9], [11]. The basic premise is that any change could give rise to new risks. Reason [11] refers to them as “defence-related ironies and paradoxes”. The most frequently cited examples include automation [2] and the procedures [17].

In summary, it is human to err and even the best organisations with the best and highly motivated people face a risk of accidents. This is because the state of no-risk is not achievable. Nonetheless, it is something all organisations strive to. In the attempt to prevent adverse effects, one must take measures. To start with, organisations need to stop looking for the one to blame (the person at the hands of which an event happened) and look deeper for the underlying mechanisms that may lead to errors. Problems with inattention, forgetfulness, or distraction can be only partly tackled with, but not exterminated. The conditions, under which people work, on the other hand, can be subject to change and, hence, should be optimised.

Table 1: Selected classifications of human error

Active failures	Latent conditions
Error of omission, i.e. omitting a task or a part of a task (e.g. a step in the task) **Error of commission**, i.e. adding something that should not be there ▪ Selection error, i.e. incorrect choice among a range of options (e.g. selects the wrong control, issues wrong command or information) ▪ Error of sequence, i.e. incorrect sequencing of actions or events ▪ Error of sequence, i.e. incorrect sequencing of actions or events ▪ Time error, i.e. action carried out too early or too late ▪ Qualitative error, incorrectly carrying out an action (e.g. too much, too little)	**Hardware (H)**, i.e. quality and availability of tools and equipment **Design (D)**, i.e. no external guidance by the designer, designed objects are opaque, the designed object does not provide feedback **Maintenance management (MM)**, i.e. safe planning of operations **Procedures (P)**, i.e. quality, accuracy, relevance, availability, and workability **Error-enforcing conditions (EEC)**, i.e. error-producing and violation-promoting conditions related to the individual or to the workplace **Housekeeping (HK)**, i.e. the problem has been there for some time, the organisation was aware of it, but did not deal with it, e.g. insufficient personnel, poor definition of responsibility, bad hardware **Incompatible goals (IG)**, i.e. individual (preoccupation with private issues), group (norms incompatible with safety goals) and organisational goal conflicts (incompatibility between safety and productivity) **Communications (C)**, i.e. communication channels do not exist; necessary information is not transmitted; information is sent, but misinterpreted by the receiver **Organisation (O)**, i.e. organisational structure, organisational responsibilities, and the management of contractor safety **Training (T)**, i.e. failure to understand training requirements; incompatibility between training and the operation, poor mixes of experienced and inexperienced personnel, poor task analysis, inadequate competence, etc. **Defences (DF)**, i.e. failure in detection, warning, personnel protection, recovery, containment, escape, and rescue

Objectives of the study and assumptions

Observations in the field and the communication with experts revealed that the risks associated with human and organisational factors in mechanised NDT are unknown; that the variability in the inspection results is frequently assigned to the inspectors and their working environment, thereby neglecting other potential influences, and that organisations with high safety and reliability demand, such as the management of spent nuclear fuel, rely on reliable NDT methods. Thus, the objectives of this study were to:

- Identify the potential failures that increase the risk that mechanised NDT will not fulfil its objective, i.e., detect all critical defects,
- Analyse the potential failures, with respect to their origin and effects on the execution of the NDT task, and
- Provide countermeasures to minimise the future risk of failure.

In line with the current state of the art in human error research, the following was assumed:

- There is a risk of failure in mechanised testing.
- The sources of failure can be seen not only in the technology but also in the individual and the organisation.
- The currently installed preventive measures are insufficient to prevent failures in the execution of the mechanised NDT inspection task.

Method

Application: Final disposal of spent nuclear fuel in Sweden and Finland

The study of assessing and treating risks in mechanised NDT was carried out between 2009 and 2011 within several projects of BAM Federal Institute for Materials Research and Testing with the Swedish and Finnish companies responsible for the development of the NDT methods to be used in the final disposal of spent nuclear fuel [18], [19]. For that purpose, four mechanised NDT methods were under scrutiny: ultrasonic testing (UT) and radiographic testing (RT), to complement each other in search for defects in the material's volume, and eddy-current (ET) and visual testing (VT) with a remote camera for the defects near or at the surface [19], [20]. In these inspections, the NDT inspection personnel is not replaced by automated systems and asked to monitor them, but the inspector is still actively involved in the setup of the measurement system, and, most importantly, in the evaluation of the collected data (even though not automated or mechanised per se, evaluation is, in this thesis, embedded in the term mechanised testing).

Failure Modes and Effects Analysis (FMEA)

Failure Modes and Effects Analysis (FMEA) is a standard risk assessment tool. Originally developed by the US Armed Forces in 1949 and revised in 1980 [13], FMEA is defined as "a procedure by which each potential failure mode in a system is analyzed to determine

the results or effects thereof on the system and to classify each potential failure mode according to its severity" (p. 4). FMEA is used to identify potential failure modes (the manner by which a failure is observed), to determine their effects (or consequences) on the operation of the system, to identify the mechanisms of failure, and to identify actions to avoid and/or mitigate the effects of the failure on the system [21]. A crucial step in this process is anticipating what might fail or go wrong. Its use is advantageous in examining potential reliability problems early in their development cycle when taking action to overcome these issues is easier, thereby enhancing reliability through design [22]. Algedri and Frieling [23] suggest that a human-oriented FMEA can lead to improvements on personal, ergonomic (optimised working conditions), and organisational level (optimised interaction of the man, technology, materials, and method), which can result in a significant decrease of system failures.

Participants

The FMEAs were carried out at the two nuclear waste management companies: at SKB in Oskarshamn (Sweden) and at Posiva Oy in Helsinki (Finland) in duration of 1-1.5 days per method. Four to five experts took part in the evaluation of each method. All participants were considered experts in their respective NDT methods (even though not certified). They were all involved in the development of the methods to be used for the inspection of the canister components for the storage of spent nuclear fuel either in Finland or in Sweden, and were, therefore, qualified for the participation in the analyses.

Procedure

Altogether six FMEA analyses were carried out: five were carried out to assess risks during the evaluation of data collected with UT, RT, ET and rVT, and one during the acquisition of data with phased array UT. (Since two companies wanted to assess the risks of the methods they use, the FMEA for the data evaluation with UT—a method used by both companies—was evaluated two times.). More attention was given to data evaluation than to data acquisition, for the following reasons: first, due to its higher perceived criticality for the future of the component (whereas errors in data acquisition can be detected during evaluation, it is harder, sometimes even impossible, to detect errors during evaluation), and second, due to a larger involvement of human inspectors in the task.

The FMEA carried out within the scope of this study was adapted to the needs of identifying potential risks in mechanised NDT and conducted using the following steps[1]:

- Decomposition of the task into sub-tasks.
- Definition of aims for the sub-tasks.
- Identification of possible failures/errors.
- Consideration of potential causes and effects of failures.
- Identification of existing preventive measures/barriers.
- Identification of potential preventive measures/barriers.

Results

Even though the potential errors were analysed separately for each NDT method, and are, therewith, method specific, some similarities in the way each method is applied can be found. It is for this reason that it was possible to combine the results for evaluation of data with different methods, with the aim of reaching general conclusions about the process. Because the NDT methods under study are under development, the results refer to *possible* or *potential* risks, causes etc.

Thirty-eight tasks in data evaluation and 30 in data acquisition were analysed resulting in the identification of altogether 90 **failure modes** in evaluation and 68 in acquisition. Table 2 and Table 3 contain examples of evaluated potential failure modes during data acquisition and data evaluation, organised according to the sub-tasks.

The potential **causes** of failures are associated with the individual, the technology, and the organisation.

- *Individual:* The individual can be a source of error both unintentionally and through rule violations. Some examples include:
  - Unintentional, e.g. subjective assessment criteria, cognitive biases (confirmation bias, representativeness bias, and availability bias), sensitivity to colours, reduced attention, lapses, over trust in automation, inexperience, and so on.
  - Rule violations, e.g. not following the inspection procedure.
- *Technology*, e.g. image quality, display characteristics, defects' characteristics (e.g. indication "hidden" behind a geometrical indication, too many indications, indications close to one another), equipment malfunction, and so on.
- *Organisation*, e.g. the working environment, organisation of the inspection process (e.g. long working hours), flawed inspection procedures, commercial pressure (i.e. time pressure), and so on.[1]

[1] The assessment of risk priority, which is also a part of the FMEA, has been omitted from this manuscript and can be found in [47]

Table 2: Failure modes associated with the subtasks in data acquisition with UT

Task	Sub-task	Failure modes
Data acquisition	Preparation of the component	Incorrect component or incorrect component orientation
	Preparation of the equipment	Inappropriate choice of the equipment, equipment malfunctioning
	Sensitivity settings (calibration)	Incorrect physical setup of the equipment, incorrect scanning parameters, misinterpretation of the collected data, incorrect evaluation of the data and decision about the quality of the completed process
	Scanning of the component	Altering of the physical conditions between the calibration and the scanning, incorrect scanning parameters, incorrect scanning process, incorrect verification of the data
	Sensitivity check (calibration check)	Change of the physical conditions from the initial calibration, incorrect scanning parameters, misjudging data inconsistency

Table 3: Failure modes associated with the subtasks in data evaluation (UT, rVT, ET, RT)

Task	Sub-task	Failure modes			
		UT	rVT	ET	RT
Data evaluation	Preparation of the software for the evaluation	Incorrect selection of the data file, inspection technique or evaluation area, incorrect settings	Missing image quality check, incorrect starting point, missing initial scanning run	Missing or incomplete data validity check	Incorrect image adjustment, inappropriate image quality and scale calibration
	Identification of indications	Missing indications, false alarms			
	Characterisation of indications	Incorrect defect type, misjudgement of the defect's origin (geometrical indication vs. actual defect)			
	Sizing & localisation	Incorrect size measurement, incorrect location of the indication			
	Decision making	False recommendation (acceptance/rejection of the component)			

The information about potential failures and their assumed causes was taken a step further by classifying the active failures and possible latent conditions that might have led to those failures according to the classifications presented in Table 1. The results of this analysis, presented in Table 4 show that the typical errors in NDT include both omissions and commissions (selection and qualitative errors). With regard to the possible latent conditions that might pave the way to failure, the most frequently suggested conditions include error-enforcing conditions (they include both individual and the environment), design, hardware, training, and the procedure.

The **consequences** of the listed failures were categorised into two levels: direct, which can be directly observed, e.g. invalid calibration, incorrect sensitivity, poor data quality, non-inspected areas, missing defects, false alarms, incorrect sizing of positioning of indications, etc. and indirect consequences, which can occur if the failures leading to direct consequences are not recovered, e.g. repeated inspection, false recommendation or false acceptance or rejection of the component.

Errors in data acquisition could be **detected** through consecutive steps, equipment malfunctioning or through data check. Errors in the evaluation of the calibration data are not always easily detected. In data evaluation, errors could be detected through cross-schecking of the results or through complimentary methods, if available. However, there is a high probability that some errors—e.g. missing inspection areas—may not be detected, thereby increasing the risk of missing defects. Considering the possibility of undetected errors, one has to evaluate the existing preventive measures and generate new ones, if needed.

Table 4: Assignment of incorrect human outputs and general failure types to the failure modes at different steps of the execution of the NDT task

Task	Method	Subtask	Incorrect human outputs					General failure types, GFTs										
			Omission errors	Commission e.				Hardware	Design	Maintenance	Procedure	Error-enforcing	House-keeping	Incompatible goals	Communication	Organisation	Training	Defences
				Selection	Sequence	Time	Qualitative											
Data acquisition	UT	Preparation of the component		✓					✓		✓	✓			✓	✓		
		Preparation of the equipment	✓	✓			✓	✓	✓	✓	✓	✓					✓	
		Sensitivity settings	✓	✓			✓	✓	✓		✓	✓	✓					✓
		Component scanning	✓	✓			✓	✓	✓		✓	✓						✓
		Sensitivity check	✓	✓			✓	✓	✓		✓	✓	✓	✓	✓		✓	
Data evaluation	UT, RT, ET, rVT	Preparation	✓	✓			✓	✓	✓		✓	✓	✓				✓	✓
		Identification	✓	✓			✓	✓	✓			✓		✓			✓	✓
		Characterisation		✓			✓	✓			✓	✓	✓				✓	✓
		Sizing and localisation	✓	✓			✓	✓	✓			✓					✓	✓
		Decision making		✓			✓					✓		✓		✓	✓	

At the time of the analysis, the installed preventive measures, i.e. **barriers** to prevent the errors in data acquisition included relying on operator skill, training, inspection procedures, and, if available, a checklist with steps that needed to be signed off during the task. In addition, having complimentary methods served as a barrier in the evaluation of data. The consideration of **potential barriers** (Table 5) suggested a potential for building *new* ones.

Table 5 The suggested barriers in the data acquisition and evaluation

Data acquisition (UT)	Data evaluation (UT, RT, ET, rVT),
Automation of the component identification and of the choice of proper tools (e.g. probes and cables) using a bar code reader; automatic refill of the coupling when it reaches a certain level, synchronisation of the movement between the UT system and the manipulator, and so on. **Hardware and software solutions**, e.g. redesign of the probe fixture; alarms for inconsistencies, insufficient amount of couplant, or to indicate the incorrect orientation of the component; automatic data archiving; and so on. Improvement of the **inspection procedures** and instructions in terms of quality and information they contain. **Organisation**, e.g. maintenance of the equipment, instructing and training the personnel; providing a disturbance-free environment; ensuring that the process is performed according to the appropriate up-to-date procedure and that all of the correct tools are used; motivation; clear responsibilities of the personnel; and so on. **Human redundancy** in the evaluation of the sensitivity settings (i.e. calibration) and in deciding whether or not to accept the inspection.	**Automation**, i.e. automated detection and sizing of indications (with confirmation by an inspector). **Software solutions**, e.g. software alarms for areas not being inspected, changing the colour scale (e.g. red denotes a high magnitude of the signal (alarm), and green a low one (safe)), defining screen view parameters (resolution, size, distance from the screen, and so on), plausibility checks in reporting. Improvement of the **inspection procedures** and instructions in terms of their content and usability. **Organisation**, e.g. disturbance-free environment; better time management; organisational learning (learning from previous events through event analyses). **Human redundancy,** i.e. evaluation performed by two independent inspectors, e.g. in cases of uncertainty, after a critical defect had been found, or randomly by the supervisors (the frequency of which would depend on the frequency of error occurrence). **Training,** i.e. in terms of introducing human factors training, by e.g. increasing awareness of possible cognitive biases, group effects, mistakes, etc. **Detection and decision aids**, i.e. visual representations of possible known defects (defect catalogue) and further development of detection and sizing aids.

Discussion

The aim of this study was to identify potential risks during mechanised NDT for the spent nuclear fuel management and find methods that can counteract their effects.

The presented results confirmed the assumptions made before the study. First, they showed that there is a chance for failure in mechanised NDT during both the acquisition and the evaluation of data. This was illustrated by a number of identified potential failure modes and their effects. Second, the consideration of the potential error causes showed that next to the technical factors, human and organisational factors play an important role throughout the entire process and could give rise to failure. It also indicated that for sources of error one should look beyond the person carrying out the inspection task—in line with the current safety practices. Third, the list of preventive measures, relying up to the point of the analysis on experience, training, qualification, and inspection procedures, has been expanded following the results of the analysis.

The analysis has shown that NDT is most frequently associated with errors of omission (e.g. missing defects) and commission (errors in selection and errors during the preparation for the inspection and the inspection process itself). Even though the errors of commission might be more prevalent during the single subtasks, a failure to detect a defect (omission) is of major concern. Omissions are frequently associated with failures in maintenance, which NDT is a part of. According to Rasumussen's [24] analysis of 200 event reports, 33% of all omissions in maintenance happen during testing and calibration, i.e. NDT. Positive effects in reducing omissions were shown by providing with cognitive aids such as check-lists and by improvements in the training [11], [25].

The most prevalent identified latent conditions present almost at all stages of the NDT task include error-enforcing conditions, hardware, design, procedures, and training. The focus on technology (hardware, design) and on training is consistent with the stage of development of the NDT methods. The participants, i.e. the developers of the NDT methods and techniques, are expectedly interested in the improvement of the technology and are aware of its current shortcomings. This is, also, where the most improvements are expected, since continuous effort is being invested in developing the procedures and the equipment, as well as considering necessary requirements for the future training. Since failures are difficult to be detected and understood in isolation from their context, some now not so salient latent conditions, i.e. those related to the organisation, may have a greater role when the operation of the fuel management starts. Considering that most errors have their origins in managerial and organisational actions or inactions [26], it is reasonable to assume that organisation, incompatible goals, communication, or housekeeping will merit more attention once the operation starts. Inadequate tools, unworkable procedures, design deficiencies, poor communication, and housekeeping are cited as some of the most influencing local-error provoking factors [25]. Holstein et al. [27] proposed that the reliability of NDT can be affected by not only the internal organisational context (i.e. the business, information, and the delivery processes), but also by the external organisational context (i.e. regulatory practices, technical rules, social/technical rules, safety culture, and the market itself).

The discussion on possible preventive measures yielded a number of ways some of the identified risks could be prevented. Some of the most salient preventive measures include the improvement of the inspection procedure and instructions, implementation of human redundancy, and hardware and software improvements, including further automation of the process.

Critical reflection on the preventive measures

Failures are contained and prevented not only by installing defences, but also by identifying gaps in the defences [25]. As discussed during the FMEA, the existing barriers might be insufficient to prevent the potential failures: For example, the task is to a high extent aided by automation, but errors do occur; inspection procedures and instructions do exist, but are flawed and need optimising, and so on. This indicates that the existing barriers might need improving and that having barriers alone does not necessarily prevent from failure. Considering the potential paradoxes associated with protective barriers [2], [8], [11], three suggested measures will be further discussed: improvement of the inspection procedures and instructions, human redundancy, and automation.

Procedures: Inspection procedures and instructions are some of the most important tools in the everyday working life of an NDT inspector. They are typically written by certified personnel in accordance with standards, codes, or specifications. During the FMEA, the procedures and instructions were identified as a potential error cause, and their optimisation as a potential barrier. Failure was generally assigned to insufficient content or to the inspectors not following the procedure, and the suggestions made for its improvement were focused mainly on its content. Operating experience and research over the years have shown that procedures and instructions are not always used properly and, thus, might need to be optimised. In his analysis of scrams (emergency shutdowns of the nuclear reactor) and LERs (Licensee Event Reports) in Swedish nuclear power plants in the period 1995–1999, Bento [28] reported that 15% of all scrams and 31% of MTO-related scrams as well as 10% of all LERs and 25% of MTO-related LERs occurred due to procedural deficiencies. Of all LERs, 23% were related to testing activities. Deficient procedure content was assigned to 70% of the procedure-related LERs and 85% of the procedure-related scrams, followed by missing procedure and missing updates. Lack of adherence to the procedure was the most important contributing cause of LERs. Procedure-related events were more related to maintenance, testing, and modification tasks (74%) than to operational tasks (20%). In Gaal et al.'s [29] study on human factors influences on manual UT inspection performance, a procedure that was written by a highly experienced and qualified writer, was not entirely understood by the users. After improvements have been made together with the participants, they reported higher satisfaction. The research initiative Programme for the Assessment of NDT in Industry - PANI, revealed that each inspector applies the procedure differently and that the inspectors do not necessarily read the full procedure or apply the procedure as intended by the procedure writers [30]. In the PANI 3 study [31], a review of the procedure from a human factors perspective was completed to identify improvements that may encourage the full use of procedures during inspections. Issues such as length and structure,

content and presentation of information, procedural steps, procedure format, and record keeping were addressed in detail. The author suggested that the inspection procedure is central to a reliable inspection, and as such needs to be written in a way that not only contains all the relevant information but also supports their systematic application. For that purpose, the procedures need to be developed together with the user. With this in mind, it becomes clear that attention should not be given only to the procedure content, but also towards its usability. Hence, the suggested approach to further development of the inspection procedure is to direct focus on understandability of the procedures and on the format in which they are presented to the inspectors, in hope that procedures will be used and will be used appropriately. Adding more procedures has been a frequent engineering approach to deal with human variability. The risk associated with over-specification is that it tends to lead to routine violations. The tendency to violate increases the possibility of an adverse event, especially in the case of breaching safe operating procedures [17]. Hence, NDT community should weigh the advantages of the extent of the procedures and of investing effort in a better training of the personnel. Furthermore, as the procedure is written by any inspector with sufficient qualification, it may be useful to develop a unified approach to its writing and guidelines that would improve its usability [2].

Human redundancy: One of the suggested methods to detect possible errors during data evaluation was to introduce human redundancy. The suggestions include random checks by the supervisors and a repeated inspection/evaluation by another inspector once a critical defect had been found. Although human redundancy is generally used to increase reliability, its implementation can also carry risks, especially when the principles of technical redundancy (i.e. two independent systems that perform the same function) are applied to social systems. One of these risks is social loafing, i.e. the phenomenon of investing less effort when working on tasks collectively than when working alone [32]. Whereas technical systems are assumed to function independently of each other, this scenario is often not the case with regard to social systems ([33], [34]. Swain and Guttmann [14] indicate the checker's familiarity with the inspector, who had already conducted the task, and his or her knowledge of the other inspector's technical level as some of the factors influencing human redundancy. Clarke [34] added that a checker might fail to perceive an error because of a belief in the colleague's competence. In the case of the management of spent nuclear fuel, in which the demand for the inspecting personnel will be low (at least in the first years of operation), independence might be difficult to achieve. In small companies, such as the two investigated in the scope of this study, but also in other inspection companies active in other domains, inspectors are highly likely to know each other and be aware of each other during the inspection. This raises concern with respect to independence. The latest studies have indicated that due to social loafing effects human redundancy might not necessarily be an effective safety measure in working with automated systems [35], [36]. Taking this into consideration, the implementation of human redundancy in NDT requires further consideration with respect to potential negative effects that can outweigh the benefits expected from redundancy [3].

Automation: Further automation of parts of the data acquisition and evaluation tasks was frequently suggested during the FMEA. The benefits of automation in form of a bar-code reader were especially seen in data acquisition, as a result of which mistyping errors or opening of the wrong setup file could be avoided. In data evaluation, automation was identified as a potential aid in identification and characterisation of indications. An example of existing automated aid is the software used for the evaluation of data with ET. This software aids in the evaluation by automatically detecting and sizing the indications, whereas the role of the inspector remains to control the results. The major goal of introducing automation into the working environment is to reduce human error [37]. The advantages and disadvantages of automation have been widely investigated in various industrial applications (e.g. [38]–[40]). Despite its many benefits regarding processing speed and accuracy, and reduction of human error to some extent, automation has shown to lead to new error sources and new risks in ways that are unintended and unanticipated by the designers [2], [5], [41]. This is because automation does not necessarily replace human operators; rather, it changes what they do. Note that automation mentioned here is not meant to replace the human operator completely, but rather aid the operator in carrying out selected tasks. For example, instead of manual control, inspectors now need to cope with the complexity of constantly developing technology and multitasking, by relying on what the equipment tells them. This can occasionally result in overlooking errors with regard to the functioning of the automated system, thereby leading to errors that may compromise safety. An uncritical reliance on the proper functioning of an automated system without recognising its limitations and the possibilities for failures often occurs when the task demands are too high and when the automated system is perceived to be reliable and is trusted [4], [5]. Considering the perceived superiority of automated systems in NDT, i.e. higher perceived reliability of mechanised over manual NDT, uncritical reliance could be one of the automation ironies associated with NDT. Nevertheless, automation offers many advantages, regarding reducing human error. When functioning properly, automation saves time, decreases workload, and generally reduces human error. The key for NDT is to be aware of the potential errors that arise from this interaction and to find means of avoiding them [4].

[2] A study into the usability and understandability of the inspection procedure has been conducted following up this risk assessment and can be found in [48]

[3] The study into the effects of social loafing on UT data evaluation has been conducted following up this risk assessment and can be found in [47]

[4] The study on the use of automation in the evaluation of NDT data has been conducted following up this risk assessment and can be found in [47]

Limitations of the study

The primary limitation of the study is that the analysed NDT methods are under ongoing development, thus lacking field application and experienced participants. The FMEA has shown to be a valuable tool for identifying and evaluating potential failures in NDT. Still, there are restrictions to this method that need to be mentioned. For example, the FMEA is suitable for identifying single failure modes, but lacks the combinations of different failure modes and could be difficult to conduct for multi-layered systems ([21]. Hollnagel [42] points out that explanation for risks cannot always be found in single components of the socio-technical system, such as the operator or the technology, but can also stem from their interaction or normal variability in human performance combined in unexpected ways. Thus, future attempts to assess risk in NDT should also include interactions between different systems. Benefits of the FMEA are seen in the multidisciplinary approach and in its ability to identify failures early in the design. In addition, it is easy to understand, highlights safety critical tasks that require attention, provides with countermeasures to prevent failure in the future, etc. [43]. Numerous alternatives to FMEA exist. It remains to be seen which of those methods is the most suitable for the purposes of identifying risks in mechanised NDT in the management of spent nuclear fuel. The process might differ greatly from the process today and new risks can and probably will arise. Other or new prospective and retrospective approaches can be used, and their suitability should be decided based on the relevant criteria at the time of the analysis. Risk and error management work best if both proactive and retroactive methods are combined [44], [45].

In conclusion, implementing preventive measures is a process that requires detailed consideration. Risk assessment and risk treatment are cyclical in nature. To ensure the highest profit, the FMEA should be repeatedly applied to identify new risks that can arise over time resulting from, e.g. implementing barriers or from the changes in the way NDT inspections are carried out. Risk management is dynamic, iterative, and responsive to change [46]. For it to be successful, the need for risk management has to be recognised, the risks need to be identified, the underlying mechanisms of their effects understood and, finally, measures have to be taken for the risks to be successfully treated. This is most frequently achieved by preventing something unwanted from happening or by protecting the organisation from its consequences [45]. This study raised questions regarding the suggested protective measures, which is why a deeper consideration of the potential implications of their implementation is needed.

Concluding remark

This manuscript is an excerpt from the doctoral dissertation: Bertovic, M. (2016). Human Factors in Non-Destructive Testing (NDT): Risks and Challenges of Mechanised NDT. Doctoral dissertation. Technische Universität Berlin. Berlin: BAM Bundesanstalt für Materialforschung und –prüfung. For the purposes of this publication, the chapter has been shortened, some paragraphs omitted and some slightly edited to provide context.

Acknowledgement

The dissertation has been conducted during the author's employment at BAM Federal Institute for Materials Research and Testing. The study has been conducted within several projects on NDT reliability between BAM, Swedish Nuclear Fuel and Waste Management Co (SKB) and Posiva Oy and owes thanks to all participants, and project partners, as well as Dr. Babette Fahlbruch, Dr. Dietrich Manzey and Dr. Christina Müller for guidance.

References

[1] *C. Müller, M. Bertovic, M. Pavlovic, D. Kanzler, U. Ewert, J. Pitkänen, and U. Ronneteg, "Paradigm Shift in the Holistic Evaluation of the Reliability of NDE Systems," Mater. Test., vol. 55, no. 4, pp. 261–269, 2013.*

[2] *L. Bainbridge, "Ironies of Automation," in New Technology and Human Error, J. Rasmussen, K. Duncan, and J. Leplat, Eds. Chichester, UK: John Wiley & Sons, 1987, pp. 271–283.*

[3] *N. B. Sarter, D. D. Woods, and D. R. Billings, "Automation surprises," in Handbook of Human Factors & Ergonomics, 2. Edition., G. Salvendy, Ed. New York, NY: Wiley, 1997, pp. 1926–1943.*

[4] *D. Manzey, "Systemgestaltung und Automatisierung [System design and automation]," in Human Factors. Psychologie sicheren Handelns in Risikobranchen. 2. Auflage, P. Badke-Schaub, G. Hofinger, and K. Lauche, Eds. Berlin Heidelberg: Springer, 2012, pp. 333–352.*

[5] *R. Parasuraman and V. Riley, "Humans and automation: Use, misuse, disuse, abuse," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 39, no. 2, pp. 230–253, 1997.*

[6] *E. Hollnagel, "The phenotype of erroneous actions," Int. J. Man. Mach. Stud., vol. 39, no. 1, pp. 1–32, 1993.*

[7] *J. Reason, Human error. New York: Cambridge University Press, 1990.*

[8] *S. W. Dekker, The field guide to human error investigations. Aldershot, England: Ashgate, 2002.*

[9] *D. D. Woods, S. W. Dekker, R. Cook, L. Johannesen, and N. Sarter, Behind Human Error. Second edition. Farnham, Surrey: Ashgate, 2013.*

[10] *J. Reason, "Human error: models and management," BMJ, vol. 320, pp. 768–770, Jun. 2000.*

[11] *J. Reason, Managing the Risks of Organizational Accidents. Farnham, Surrey: Ashgate, 1997.*

[12] *J. Rasmussen, "Risk management in a dynamic society: a modelling problem," Saf. Sci., vol. 27, no. 2–3, pp. 183–213, 1997.*

[13] MIL-STD 1629A, "Procedures for performing a Failure Mode, Effects and Criticality Analysis. Military standard." U.S. Department of Defense, Washington, DC, 1980.

[14] A. Swain and H. Guttmann, "Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report [No. NUREG/CR-1278; SAND-80-0200]," U.S. Nuclear Regulatory Commission, Washington, D.C., 1983.

[15] J. Reason, R. Shotton, W. Wagenaar, P. T. W. Hudson, and J. Groeneweg, "TRIPOD, A Principled Basis for Safer Operations," Shell Internationale Petroleum Maatschappij, The Hague, 1989.

[16] T. B. Sheridan, "Risk, Human Error, and System Resilience: Fundamental Ideas," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 50, no. 3, pp. 418–426, Jun. 2008.

[17] J. Reason, "A systems approach to organizational error," Ergonomics, vol. 38, no. 8, pp. 1708–1721, Aug. 1995.

[18] Posiva Oy, "Annual report 2014," Posiva Oy, Eurajoki, Finland, 2015.

[19] SKB, "RD&D Programme 2013. Programme for research, development and demonstration of methods for the management and disposal of nuclear waste [SKB Report TR-13-18]," Svensk Kärnbränslehantering AB., Stockholm, 2013.

[20] J. Pitkänen, "Inspection of Disposal Canisters Components [Report POSIVA 2012-35]," Posiva Oy, Eurajoki, Finland, 2013.

[21] IEC/ISO 31010, "Risk management - Risk assessment techniques," ISO/IEC, Geneva, Switzerland, 2009.

[22] G. Cassanelli, G. Mura, F. Fantini, M. Vanzi, and B. Plano, "Failure Analysis-assisted FMEA," Microelectron. Reliab., vol. 46, pp. 1795–1799, Sep. 2006.

[23] J. Algedri and E. Frieling, Human-FMEA. Munich: Carl Hanser Verlag, 2001.

[24] J. Rasmussen, "What can be learned from human error reports?," in Changes in Working Life, K. Duncan, M. Gruneberg, and D. Wallis, Eds. London: Wiley, 1980.

[25] J. Reason and A. Hobbs, Managing maintenance error: a practical guide. Aldershot, England: Ashgate, 2003.

[26] J. Reason, "Managing the management risk: New approaches to organisational safety," in Reliability and Safety in Hazardous Work Systems: Approaches to Analysis and Design, B. Wilpert and T. Qvale, Eds. Hove: Lawrence Erlbaum Associates Ltd., Publishers, 1993, pp. 7–22.

[27] R. Holstein, M. Bertovic, D. Kanzler, and C. Müller, "NDT Reliability in the Organizational Context of Service Inspection Companies," Mater. Test., vol. 56, no. 7–8, pp. 607–610, Jul. 2014.

[28] J. Bento, "Procedures as a Contributing Factor to Events in the Swedish Nuclear Power Plants [SKI Report 02:63]," Swedish Nuclear Power Inspectorate (SKI), Nyköping, Sweden, 2002.

[29] M. Gaal, M. Bertovic, S. Zickler, B. Fahlbruch, V. Spokoiny, D. Schombach, T. Just, and H.-J. Cramer, "Untersuchungen zum Einfluss menschlicher Faktoren auf das Ergebnis von zerstörungsfreien Prüfungen, Möglichkeiten zur Minimierung dieses Einflusses und Bewertung der Prüfergebnisse," Bundesamt für Strahlenschutz, Salzgitter, Germany, 2009.

[30] B. McGrath, G. Worrall, and C. Udell, "Programme for the Assessment of NDT in Industry, PANI 2," HSE report on CD-ROM, 2004.

[31] B. McGrath, "Programme for the Assessment of NDT in Industry, PANI 3 [Report No. RR617]," Health and Safety Executive, 2008.

[32] S. J. Karau and K. D. Williams, "Social loafing: A meta-analytic review and theoretical integration.," J. Pers. Soc. Psychol., vol. 65, no. 4, pp. 681–706, 1993.

[33] S. D. Sagan, "The problem of redundancy problem: why more nuclear security forces may produce less nuclear security," Risk Anal., vol. 24, no. 4, pp. 935–46, Aug. 2004.

[34] D. M. Clarke, "Human redundancy in complex, hazardous systems: A theoretical framework," Saf. Sci., vol. 43, no. 9, pp. 655–677, Nov. 2005.

[35] J. Marold, "Sehen vier Augen mehr als zwei? Der Einfluss personaler Redundanz auf die Leistung bei der Überwachung automatisierter Systeme," (Doctoral dissertation. Technische Universität Berlin), 2011.

[36] D. Manzey, K. Boehme, and M. Schöbel, "Human Redundancy as Safety Measure in Automation Monitoring," Proc. Hum. Factors Ergon. Soc. Annu. Meet., vol. 57, no. 1, pp. 369–373, 2013.

[37] L. J. Skitka, K. L. Mosier, and M. D. Burdick, "Does automation bias decision-making?," Int. J. Hum. Comput. Stud., vol. 51, no. 5, pp. 991–1006, 1999.

[38] J. E. Bahner, A.-D. Hüper, and D. Manzey, "Misuse of automated decision aids: Complacency, automation bias and the impact of training experience," Int. J. Hum. Comput. Stud., vol. 66, no. 9, pp. 688–699, 2008.

[39] P. Madhavan, D. A. Wiegmann, and F. C. Lacson, "Automation Failures on Tasks Easily Performed by Operators Undermine Trust in Automated Aids," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 48, no. 2, pp. 241–256, 2006.

[40] E. Alberdi, A. A. Povyakalo, L. Strigini, and P. Ayton, "Effects of incorrect computer-aided detection (CAD) output on human decision-making in mammography," Acad. Radiol., vol. 11, no. 8, pp. 909–918, 2004.

[41] R. Parasuraman and D. Manzey, "Complacency and Bias in Human Use of Automation: An Attentional Integration," Hum. Factors J. Hum. Factors Ergon. Soc., vol. 52, no. 3, pp. 381–410, 2010.

[42] E. Hollnagel, "The changing nature of risks," Ergon. Aust. J., vol. 22, no. 1, pp. 33–46, 2008.

[43] B. Dhillon, Design Reliability: Fundamentals and Applications. Boca Raton, FL: CRC Press, 1999.

[44] K. A. Latorella and P. V Prabhu, "A review of human error in aviation maintenance and inspection," Int. J. Ind. Ergon., vol. 26, no. 2, pp. 133–161, Aug. 2000.

[45] E. Hollnagel, "Risk + barriers = safety?," Saf. Sci., vol. 46, no. 2, pp. 221–229, Feb. 2008.

[46] ISO 31000, "Risk management — Principles and guidelines.," International Organization for Standardization (ISO), Geneva, Switzerland, 2009.

[47] M. Bertovic, Human Factors in Non-Destructive Testing (NDT): Risks and Challenges of Mechanised NDT. Doctoral dissertation. Technische Universität Berlin. Berlin: Bundesanstalt für Materialforschung und -prüfung (BAM), 2016.

[48] M. Bertovic and U. Ronneteg, "User-centred approach to the development of NDT instructions [SKB Report R-14-06]," Svensk Kärnbränslehantering AB., Oskarshamn, Sweden, 2014.


Foto: P. Schüle

Dr. Marija Bertovic ist Trägerin des Wissenschaftspreises 2018 der DGZfP. Von 2006 bis 2017 arbeitete und forschte sie an der BAM zum Thema „Human Factors bei der Zuverlässigkeit zerstörungsfreier Prüfsysteme" in der Arbeitsgruppe von Dr. Christina Müller. 2015 war sie im Rahmen eines Projekts viereinhalb Monate bei der DGZfP Ausbildung und Training GmbH beschäftigt. Bertovic untersuchte, welchen Einfluss menschliche Faktoren wie Zeitdruck, Umgang mit der Automatisierung, menschliche Redundanz oder suboptimale Prüfanweisung auf Prüfergebnisse haben können. 2015 schloss sie ihre Promotion zum Thema "Human Factors in Non-Destructive Testing (NDT): Risks and Challenges of Mechanised NDT" mit der Note "magna cum laude (sehr gut)" ab.

Marija Bertovic stammt aus Kroatien. Von 2000 bis 2006 studierte sie in Rijeka Psychologie. Christina Müller war zu dieser Zeit als Beobachterin an einem Projekt zur Minensuche in Kroatien tätig. Sie lud Marija Bertovic als Gastwissenschaftlerin an die BAM nach Berlin ein. Daraus entstand gemeinsam mit u.a. Dr. Mato Pavlovic und Dr. Daniel Kanzler eine sehr engagierte Arbeitsgruppe zu Fragen rund um die Zuverlässigkeit Zerstörungsfreier Prüfungen an der BAM. Gemeinsam entwickelten sie eine sehr erfolgreiche Veranstaltungsreihe, den "European-American Workshop on Reliability of NDE" weiter, der im vergangenen Jahr bereits zum siebten Mal veranstaltet wurde, mit 56 Teilnehmenden aus Europa, den USA, Südamerika und Asien. Dr. Christina Zanotelli, vormals Müller, ist leider Anfang 2017 verstorben.

Seit Oktober ist Marija Bertovic wieder an der BAM beschäftigt und arbeitet an probabilistischen Zuverlässigkeits- und Sicherheitsanalysen/Human-Factors-Analysen im Bereich der Werkstoffprüfung.