Submitted for publication in the AIP Conference Proceedings of the *42st Annual Review of Progress in Quantitative Nondestructive Evaluation, 2016.*

Conclusions of the 6th European American Workshop on Reliability of NDE

Christina Mueller[1, a)], Marija Bertovic[1], Daniel Kanzler[1], and Ulf Ronneteg[2]

[1]*BAM Federal Institute for Materials Research and Testing, Unter den Eichen 87, 12205 Berlin, Germany*
[2]*SKB Swedish Nuclear Fuel and Waste Management Co., Box 929, SE-572 29 Oskarshamn, Sweden*

[a)]christina.mueller@bam.de

Abstract. The principles of Open Space Technology (OST) were again applied to discuss burning issues in the field of NDE reliability. The results of the discussions among NDE professionals concerning new reliability methods, human factors and integrated solutions will be presented.

INTRODUCTION

The European American workshops (EAW) have a history of 18 years. The motivation for the first European American Workshop in 1997 was the unsolved discrepancy about the way how to guarantee the quality of NDE processes between the American proposal in terms of the PDI (Performance Demonstration Initiative by EPRI [1])—to do it by practical performance demonstration under field conditions—and the conservative European approach—to do it by the definition of the physical parameters and their tolerances (e.g. physical reasoning by ENIQ). The main achievement of the first Workshop was the conceptual model [2] (later referred to as Modular Reliability Model [3] in terms of the reliability formula which simply says that the total reliability of an NDE system is composed of the intrinsic capability, IC (which stands for the physical principle behind the defect indication and its technical realization as an upper bound), the application factors, AP (describing the realistic circumstances like UT coupling, limited access, noise of the surrounding …) and the human factors, HF, present in each application. While imperfect—e.g. the mutual interactions between the factors were not considered—the conceptual model helped to properly define the potential for performance optimization, and worked also as an assessment tool for the adequacy of open and blind trials.

The main benefit from the second European American Workshop on NDE in 1999 [4], was the clear definitions of a) the NDE system as the procedure, equipment and personnel that are used in performing NDE inspection and b) the NDE reliability as the degree that an NDT system is capable of achieving its purpose regarding detection, characterization and false calls.

The main conclusion from the third European American Workshop in 2002 [5] was: The risks in NDE and demining need to be defined. The fourth European-American Workshop in 2009 furthermore showed the progress in attempts to consider the reliability of NDE on a system level with the goal for integrated solutions in the industrial applications. The limitations of the original empirical methods were shown resorting to advanced and model assisted methods. This was also the first time that an extra *human factors* session was launched resulting in the recognition that human factors are present at all stages of NDE activity and that improvements can be made to the training, the procedure, the calibration, as well as to the inspection and evaluation processes.

The fifth European American Workshop in 2013 [6], [7] had its focus on the question "What is the '*delta*' between our existing reliability evaluation or qualification methods and the real performance under everyday field conditions?". In the open space technology discussion, the information exchange between NDE providers and customers as well as between science and industry—especially concerning shortcomings of mechanized or automated testing—were identified as weak-points.

The specific aim of the current workshop was to explore “How to live holistic reliability in reality?”. From the technical side, it is of interest to come to an agreement on how to deal with real parts and defects versus artificial ones in terms of sophisticated combination of methods. Furthermore, the application of modeling tools came to a higher level of understanding with respect to how to calibrate the signals and the noise, and with respect to the awareness that there is always an “unknown factor” present. Regarding human factors and the organizational context, the issue of the flow of information was further explored and the question of how to balance regulations versus culture was raised. The challenge of the reliability of Structural Health Monitoring (SHM) is met today by well-established programs and concepts.

The Workshop was held for the first time in its history as a part of the RPQNDE in Minneapolis, MN, USA. The connection to the biggest scientific conference on the progress in quantitative NDE opened doors to new mutual inspirations.

OPEN SPACE TECHNOLOGY (OST) WORKSHOP

After the success of the last EAW [6], the principles of Open Space Technology (OST) [8] were used again to discuss the current issues in the field of reliability of NDE (see [7] for the description of the OST and for the results of the previous EAW’s discussions). OST provides an interactive platform in which NDE professionals can be brought together to discuss topics that interest them.

The workshop is carried out by collecting questions on all topics of interest to the NDE community on the wall of ideas and then discussing them in smaller topic-oriented working groups. In the end, the groups are asked to present the results in front of all members. The summaries of these discussions will be henceforth presented.

Three groups were formed: (i) **Technical POD and advanced methods** dealing with physical and mathematical questions related to the POD basics and advanced methods, especially to the MaPOD approaches; (ii) **Human factors** dealing with questions related to skill and motivation of personnel, procedure, operator attention, and the knowledge transfer from general human factors research; and (iii) **Integrated Solutions,** where the POD and human factors principles are applied in industry within a tension or balance between rules and standards and a living safety culture.

Technical POD and Advanced Methods

This group was focused on the statistical requirements of the probability of detection (POD) evaluation and the question, how to calculate a correct POD characteristics. The questions, which came from attendees of the conference, ranged from “how to define the principle inputs” to “the specific use of the POD in the real testing environment”.

The main issue discussed was the definition of “a”, the defect size metric, in cases, in which there are more than one main defect parameter (when the typical relationships — penetrated length in RT or crack depth in ET — are not sufficient or cannot be defined). One exemplary situation is airport security, in which there is no main defect parameter that defines the criticality of the indication. POD approaches for this situation are still under development; e.g. [9]. One approach, which can be used in several cases, is the multi-parametric POD approaches (e.g.[10], [11]). In this case, the POD dependence on a number of different parameters can be taken into consideration. A model that works without the parametric dependence on a defect parameter is the hit/miss analysis that is very often usable for an easy POD approach. Still it should be kept in mind that in the hit/miss analysis a monotonically increasing function is often needed. But this approach requires a larger amount of data and has fewer possibilities to check the evaluation process.

Furthermore, the false alarm rate or probability of false alarm (PFA) of the system under investigation was put in focus. Since the aim is to keep the level of false alarms low, is it not possible to see them in the POD curve. One indication of a high PFA is a non-zero minimum value of the POD curve with the y-axis [12]. But the crossing point is not directly the PFA, rather a sign of not fulfilling the basic requirements of the statistical models. The Spencer model should be kept in mind for the practical evaluation of testing results. This model is still under discussion due to the relationship between the four-parametric model and the PFA.

Another discussed question was the choice of the statistical model. Tools to find correct models include variance tests and likelihood aspects. An interesting discussion went about the fact that one and the same data set might result in substantially different POD curves due to different models applied. Therefore, a discussion will be continued after sighting the data set [13]. The future discussion will be planned in the next European American Workshop in Munich in 2016.

One question that is asked in almost every POD discussion is the amount of samples. The amount of samples is shown in the confidence bounds, which reflects the compliance with the chosen statistical model. The number of samples drives the confidence bounds. The 29/29 model can use only 29 defects and is by definition based on the binomial statistics with the statement for one defect size (one bin of defect sizes). The model-assisted POD (MaPOD) approaches are changing the importance of the number of samples; therefore the confidence bound is not appropriate anymore. One alternative is the usage of a margin band on quantities and variations in inputs and model errors, which are the counterparts of the statistical variations. The main challenge is the accumulation of the model and the data uncertainty to get a significant result. Experiments, on the other side, contain uncertainties that are not included in the confidence bound. Therefore, alternatives, such as quantiles and tolerance bands, can be used [14]. In this way, the MaPOD approaches might find an adequate alternative to the confidence bounds. However, at the same time, they need more information.

The next topic, related to the confidence bands, is the comparison between two POD curves. The comparison between experimental and MaPOD is still a challenge and should be in the focus in the near future. In the comparison of two curves the confidence bounds can intersect for the same system. The main tool is the analysis of variance (ANOVA) to test the fit of the parameters. One important fact is, though, that too small defects are not important for the detection. The connection with the reliability program and the structural integrity is essential. The comparison (especially with the MaPOD) is relevant if the POD curves contain the variability of the same parameters, which is a question of the design of experiments.

Another question often raised in the last discussions was the independence of data points. The main questions that need to be asked here instead are: How do we control variability? How do we sample data to get the POD (doing a good design of experiments)? How do we get feedback to know we are correct? The idea of independence is important and might be used incorrectly, which should be discussed also at the next European American Workshop. These questions are more important as the defined results of independence tests. The feedback and the quality of the results should stand in the focus of the discussions and it might make sense to repeat POD trials to check approaches. Therefore, a network is desired to make sure to have a working approach.

Two practical discussions were also discussed:

1) The combination of MaPOD and human factors: The main question in this case is what a model can do. This is a question the evaluator of the data and the user of simulation should know.
2) Which role does the variability play in simulations and in real experiments? Are there differences? This is a question that is still under investigation and not so easy to answer. Safety factors and the approach of tolerance bounds can help but should be further discussed in the next workshops and further researches.

In the SHM, the POD needs a design of experiments (DOE) that includes all variability - in the same way as should be used for NDE. Sensitivity studies are essential. Furthermore, the varying threshold is a further difference to the normal NDT. Examples for varying threshold can be found in some of the published work (e.g. [9], [15]).

The question of the attendees showed, on the one hand side, that there is a continuing need for discussing the statistical requirements of the POD evaluation. On the other hand, the discussion highlighted that new needs and new fields of application give us—as the community—the task to work on new approaches and to solve the open questions of getting a useful and meaningful POD evaluation.

Human Factors

Human factors again drew plenty of attention of the international NDE community. The first issue of concern was the inspectors' skill degradation over time. If not used, the skills are bound to deteriorate. It was suggested that the retention of those skills could be supported by means of requalification at regular intervals (already practice in some domains, e.g. requalification every 5 years in railway) and/or by a demonstration/practice task, which would give inspectors an opportunity to train their skills immediately before the inspection (also known as the "´just-in-time´ training"). The benefits of skill demonstration/practice task were observed in a study by Bertovic et al. [16], in which the participants were given an opportunity to practice on a test piece to demonstrate whether they have understood the inspection procedure. The participants reported high satisfaction with the demonstration task, as it allowed them to refresh their skills and provided feedback about their performance according to a new inspection procedure.

The second discussed topic was the inspection procedure, i.e. the community was interested in a) how to design procedures that fit both the “new” and the experienced “best” inspectors and b) whether it is possible to quantify how often inspectors follow the procedures. The panel concluded that the procedures need to be designed for the *least experienced* inspector. Only this way it can be ensured that personnel of various experiences can accurately complete the task. However, too detailed procedures may not be read by the experienced personnel. This can be especially critical when changes are introduced into the procedures. This issue could be overcome by highlighting the changes in the procedures, therewith making the new information stand out and draw attention to be read. Another useful strategy could be to develop long and short forms of the procedure (for examples, see [17]. Apart from content, procedure usability and design should also be addressed in order to make the procedures more usable and more likely to be used by all personnel (e.g. [17], [18]). Some lessons could also be learned from other fields, e.g. from the guidelines for the preparation of emergency operating procedures (e.g. NUREG 0899; see [19]). Anonymous survey may provide with useful information on the nature and frequency of the inspection procedures' use. Moreover, adopting a user-oriented approach in the design of the procedures is a useful strategy for the design of the procedures that will a) be applicable to users with different experiences and b) be more likely to be used.

Some NDE tasks are very monotonous and maintaining vigilance in those tasks is a continuous concern for the NDE community. Potential strategies for maintaining vigilance in monotonous tasks may include artificially induced indications, as well as requiring the personnel to report geometrical indications—if present—in tasks, in which the occurrence of defects is very rare. Airport security is a domain in which artificial indications, i.e. artificially-induced threat images of weapons or explosives, are being used in luggage screening as a method for increasing vigilance. However, a question was raised whether a failure to detect those indications should only be used to indicate vigilance decrements or whether repeated failure should also be punished. The members of the discussion group concluded that a repeated failure to detect those indications may indicate a need for further training, which should be a preferred method over punishment.

One important topic discussed during the EAWs is the further development of model-assisted POD (MaPOD). Considering that MaPOD approach “excludes” human factors, the human factors discussion group was asked whether it is possible to incorporate human factors into MaPOD and if yes, how to do it. The discussion group agreed that MaPOD approach may not be able to include human factors as human behavior is difficult to model and valuable information may be lost if human behavior is averaged. In extension, this suggests that MaPOD may not be able to provide a realistic scenario of the field practice and its results should be taken with caution.

The existence of three important gaps, i.e. the gap in knowledge (missing knowledge on human factors), the gap in communication (e.g. missing communication between NDE and the human factors field), and the gap in implementation (the findings are not known and/or not implemented in the practice)—identified during the previous Workshop [7]—was further addressed and the methods for overcoming some of them were discussed. Some efforts in closing the gaps have been acknowledged, e.g. implementing reliability topics into the regular UT level 3 training commenced by the German Society for Non-Destructive Testing (DGZfP)—even though human factors issues are still missing—, and funding new research (Germany and United States). However, the discussed gaps are far away from being closed.

The discussion group suggested that the strengthening of the communication between NDE and the field of human factors may provide the NDE field with valuable knowledge. For example, a reoccurring question on how working conditions or time pressure affect performance in NDE may be answered by a vast amount of existing literature on human performance under those conditions. The transfer of this knowledge could be achieved by theoretical topical reports. Furthermore, the knowledge gap can be filled by further experimental studies. However, acquiring a sufficient amount of participants required to have a sound statistical basis for the conclusions can be hard and costly. An acceptable alternative could be to use students or trainees (as is practice in various other domains of human factors research). For that purpose, an appropriate simulation of the NDE task needs to be developed, that would provide an interface for which participants do not need to undergo typical NDE training, but can still carry out tasks which resemble the demands of NDE in the practice. With such a platform and this approach, scientists can obtain knowledge on human behavior in NDE, which— combined with the field studies — can expand existing knowledge and lead to generation of optimizing strategies that would lead to higher reliability, and possibly safety of organizations that NDE serves to protect.

Integrated Solutions

The aim of the group named *Integrated Solutions* was to discuss (i) the interplay of POD and fracture mechanics/structural integrity and (ii) the human factors and organizational issues and how this in combination results in safe processes and products.

POD-related Questions

One topic that was thoroughly discussed was concerned with the purpose of calculating POD curves. One answer was that it is done simply because it is required by law or the authority. The deeper reflection on this question makes clear that the relationship between the structural integrity and the POD results is vital for the safety of components and processes. The gain in using POD curves was discussed instead of only calculating the probability of the presence of relevant flaws. Three aspects were seen: i) the structural engineer has first to define the types and sizes of critical flaws; ii) the requirement to distinguish between the needs for the input to establish a proper POD curve and the actual safety critical flaw parameters e.g. a specific geometry of the flaw; and finally iii) the group members agreed that the probability of flaw occurrence would be desirable to know but is in almost all cases highly uncertain.

The next question covered POD and damage tolerance and what might be the hidden relationship between them. Examples were itemized from different branches of industry for the first three questions:

- *US air force*: The a90/95 value sets the inspection intervals based on the risk calculations for each individual location in the damage tolerance analysis. In these calculations, the known detectability is used to set the inspection interval based on the criteria that if no crack has been detected, any fatigue crack cannot grow to critical size before after two inspection intervals.
- *Nuclear (US):* In the lifetime cycle, in addition to the inspection results, the information on the POD is needed to make sure that when a crack of a certain size is not detected it is actually not there or merely not detected (see Fig. 1).

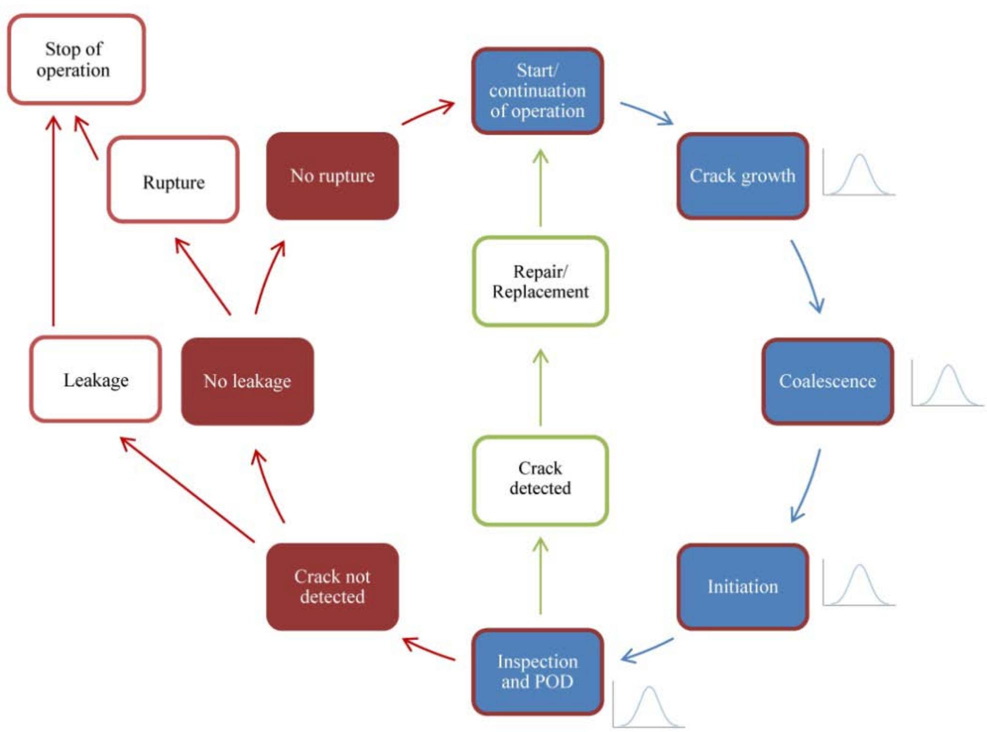

FIGURE 1. Lifetime cycle of a component of a nuclear power plant under operation

- Railway wheel axles: The crack growth speed and POD of the inspection are used more and more to set the inspection intervals.

- Spent nuclear fuel storage: The POD is a part of the technical justification (qualification) of the inspection systems but it is also planned to be a part of the final risk assessment. The determination of the POD is planned to be supported by real defect verifications and modelling.

Another discussed topic was why and how the POD in the field might be low. The group agreed that it is not known whether the POD in the field is low or not and that cases might exist where the POD is actually low due to working conditions. This raised the question if it will ever be possible to achieve realistic quantitative POD for the manual NDE performed in the field. If yes, how should this be done? The answer was that in almost all cases, the POD equals intrinsic capability (POD = IC). AP and HF can be covered only partly (see Fig. 2 for the Modular reliability model [3]). Could maybe a performance demonstration predict the field POD? The group agreed that performance demonstration can only predict the maximum possible performance in the field.

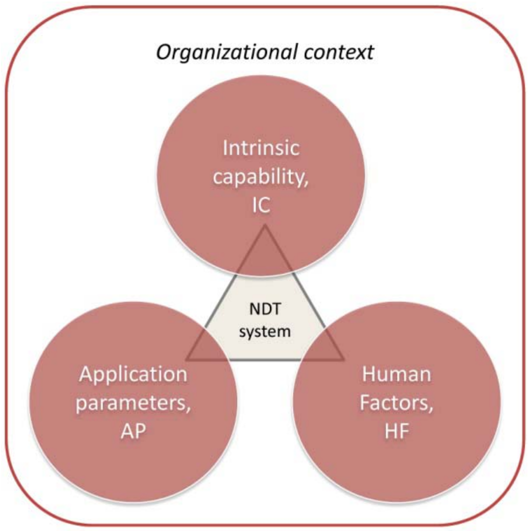

FIGURE 2. Modular reliability model

However, for a realistic picture of the reliability in the field the different factors can in some sense be covered by: handling the organizational context by standardisation procedures; considering the intrinsic capability as the ideal POD; considering application factors in making the POD on representative configurations; investigating the human factors within a duplicate inspection environment; and by striving to minimize the differences and to understand the important factors.

In order to use the POD and the human factors models in the field it is critical to harmonize the several different existing models. The group agreed that there are not only practical problems but also legal ones for the safety critical applications. The proposals to overcome the obstacles were to standardize, on the one hand, the mathematics, e.g. M1823 and, on the other hand, the procedures for qualification (e.g. ENIQ). The need for different specification for different industrial branches, e.g. nuclear, aviation, railway, was also highlighted.

A more practical topic was if it could be possible to practically build the ergonomics of an inspection into a POD by using a transfer function from a "generic" POD how to a "specific" inspection situation. The answers were to make a representative/realistic experiment with different ergonomic conditions and to define a transfer functions from the "data is ideal" POD to "specific situation" POD eventually supported by simulations.

Examples can be found on the MAPOD website [20], describing e.g. the transfer from ideal cracks/notches to cracks with real geometry.

Human and Organizational Factors Related Questions

The discussion group furthermore identified that the knowledge of NDT reliability—in terms of both the POD and the human factors—is rather limited in the NDT community. Based on this it was questioned whether an NDT reliability guideline could be issued in order to spread the knowledge in the NDT community.

Following two oral presentations highlighting the importance of safety culture, the following question was raised: Where on the safety culture ladder is actually the NDE community. Figure 3 depicts the safety culture ladder concept developed by Hudson [21] with the suggested position of NDE organizations.

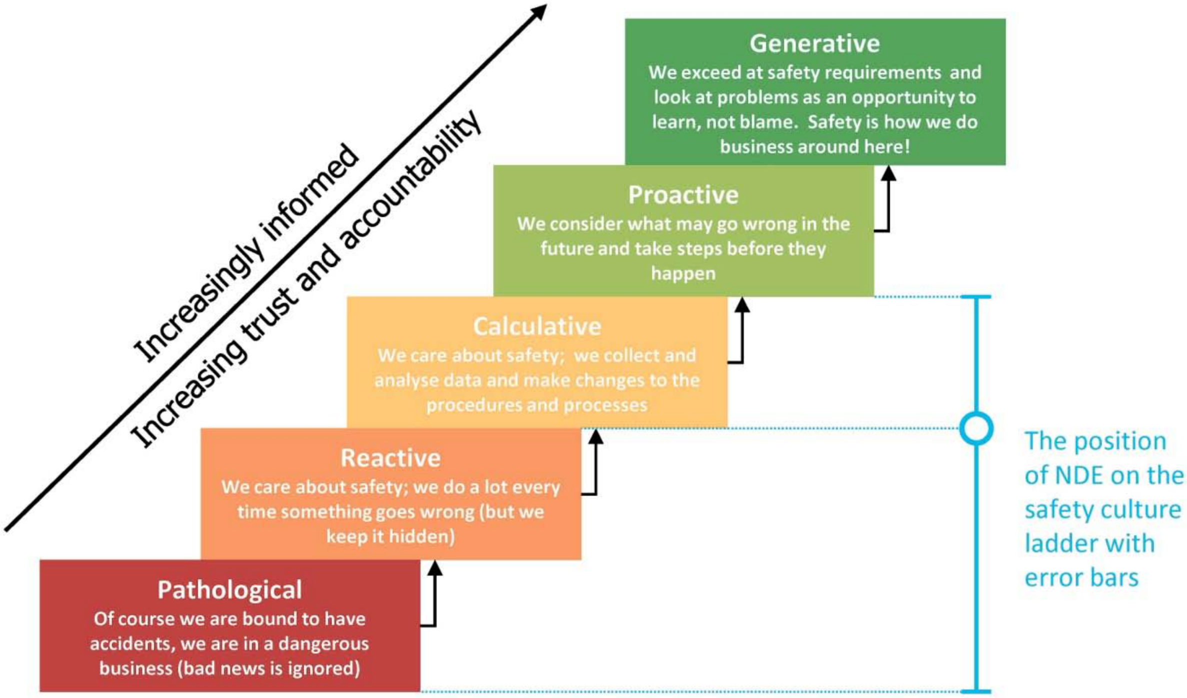

FIGURE 3. The position of the NDE on the safety culture ladder

As seen from the Fig. 3, the group suggested the central position between the levels reactive habit and calculative habit. But the variety in different industries is huge so that the error bars would also reach the pathological habit but at the same time already touch proactive ladder. The question was also raised who is responsible for the implementation of the safety culture. From the list of possible influencers—owner, NDT service vendor, examiner, equipment OEM, regulator, end user—it was agreed that the owner is responsible for the safety culture in the company but that it is finally an interplay of all the parties involved. It was agreed that it should be indicated that the safety culture is adding a value to the product. Safety culture can be in principle mandated but actually an intrinsic motivation is necessary and the regulator can only trigger the process of implementation. Finally, the management and people need to implement the safety culture and be motivated to do this. The financial aspects should be seen on a long term scale.

The final topic that was discussed was concerned with how to achieve a healthy balance between rules/regulations and culture. The group discussed that the regulations provide the tools (e.g. mandatory quality control), but the organisation has to implement it. A good reference is the research from the field of aviation regarding cultural differences (e.g. pilots). The following were identified as the keywords for safety culture: firmness, repetition, and consistency (reliability culture); and the following as the key elements of safety culture: commitment, learning, reporting, and communication. It is important to recognize that NDT reliability is a path of safety and needs interaction with the end user and the vendor.

CONCLUSIONS

While the advanced reliability models become better and better the determination of the reliability of an NDE-system will never be perfect. Some of the variation, that cannot be explained by the pure technical capability could be explained by the application parameters and the human factors (see Modular Reliability Model by [2], [3]). However, even these elements of the reliability model cannot explain the entire variation in the inspection performance, as a number of *unknown* factors that can affect it too. Those unknown parameters represent the uncertainty in the measurement or the residual or even additional influences not yet included in the model. As a consequence the reliability determination must be mastered by the experiment under realistic conditions supported by modeling. For the usefulness of the POD in the field the interplay of POD and structural integrity is vital. The consortium agreed that an optimum in the safety or reliability level could be reached by an appropriate safety or reliability culture in the companies supported by the authorities.

ACKNOWLEDGEMENT

The authors would like to thank all the participants of the 6th European American Workshop for their scientific contributions and the QNDE for an outstanding organization and welcome. Furthermore, we would like to thank all participants of the open space discussion groups for their valuable input, as well as to the group moderators for their guidance.

REFERENCES

1. "ASME Boiler & Pressure Vessel Code, Section XI, Appendix VIII." American Society of Mechanical Engineers, New York, 2002.
2. C. Nockemann and C. Fortunko, "Summary of the workshop," in *Proceedings of the European-American Workshop: Determination of Reliability and Validation Methods on NDE, 18-20 June 1997, Berlin, Germany*, Berlin: DGZfP, 1997, pp. 11–17.
3. 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," *Materials Testing*, vol. 55, no. 4, pp. 261–269, 2013.
4. T. Taylor and C. Nockemann, "Summary of American-European Workshop on NDE Reliability," in *Paper summaries book - American-European Workshop on Nondestructive Inspection Reliability, 21-24 September 1999, Boulder*, Columbus, OH: The American Society for Nondestructive Testing, Inc., 1999, pp. 7–8.
5. R. Singh, M. Anderson, and C. Müller, "Summary & Consensus. Breakout session 5," in *Proceedings of the 3rd European-American Workshop on Reliability of NDE and Demining, 11-13 September, Berlin, Germany. DGZfP Proceedings BB-81-CD*, 2002.
6. C. Müller, R. Holstein, and M. Bertovic, "Conclusions of the 5th European American Workshop on Reliability of NDE," *Materials Testing*, vol. 56, no. 7–8, pp. 599–601, 2014.
7. M. Bertovic, P. Calmon, L. Carter, J. Fischer, D. Forsyth, R. Holstein, B. McGrath, C. Müller, M. Pavlovic, U. Ronneteg, W. Rummel, F. Schubert, and G. Selby, "Summary of the 'Open Space Technology Discussions' on Reliability of NDE," *Materials Testing*, vol. 56, no. 7–8, pp. 602–606, 2014.
8. O. Harrison, *Open Space Technology: Ein Leitfaden für die Praxis*. Schäffer-Poeschel Verlag, 2011.
9. D. Kanzler, U. Ewert, C. Müller, and J. Pitkänen, "Observer POD for radiographic testing," *41st Annual Review of Progress in Quantitative Nondestructive Evaluation. AIP Conference Proceedings*, vol. 1650, pp. 562–570, 2015.
10. M. Pavlovic, K. Takahashi, and C. Müller, "Probability of detection as a function of multiple influencing parameters," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 54, no. 11, pp. 606–611, 2012.
11. M. Carboni, "A critical analysis of ultrasonic echoes coming from natural and artificial flaws and its implications in the derivation of probability of detection curves," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 54, no. 4, pp. 208–216, 2012.
12. F. W. Spencer, "Curve fitting for probability of detection data: A 4 – parameter generalization," *40th Annual Review of Progress in Quantitative Nondestructive Evaluation. AIP Conference Proceedings*, vol. 1581, pp. 2055–2062, 2014.
13. A. Gianneo, M. Carboni, C. Müller, and U. Ronneteg, "PAUT Inspection of Copper Canister: Structural Attenuation and POD Formulation," *42nd Annual Review of Progress in Quantitative Nondestructive Evaluation. AIP Conference Proceedings (current issue)*.
14. C. Annis, J. C. Aldrin, and H. Sabbagh, "What is Missing in Nondestructive Testing Capability Evaluation," *Materials Evaluation*, vol. 73, no. 1, pp. 44–54, 2015.
15. C. Gollwitzer, C. Bellon, A. Deresch, U. Ewert, and G.-R. Jaenisch, "On POD Estimations with Radiographic Simulator aRTist," in *International Symposium on Digital Industrial Radiology and Computed Tomography, 20-22 June, 2011, Berlin, Germany*, 2011.

16. M. Bertovic, M. Gaal, C. Müller, and B. Fahlbruch, "Investigating human factors in manual ultrasonic testing: testing the human factors model," *Insight*, vol. 53, no. 12, pp. 673–676, 2011.
17. B. McGrath, "Programme for the Assessment of NDT in Industry, PANI 3 [Report No. RR617]," Health and Safety Executive, 2008.
18. 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.
19. NRC, "Guidelines for the Preparation of Emergency Operating Procedures [NUREG-0899]," U.S. Nuclear Regulatory Commission, Washington, DC, 1982.
20. Model-Assisted POD Working Group, "Reference documents." [Online]. Available: https://www.cnde.iastate.edu/mapod/Reference Documents.htm.
21. P. Hudson, "Implementing a safety culture in a major multi-national," *Safety Science*, vol. 45, no. 6, pp. 697–722, 2007.