Scientific Papers of the University of Pardubice,
Series A; Faculty of Chemical Technology
27 (2021) 201–226.

Human Reliability Assessment: Incorporation of organizational and cultural impacts into human performance

Radim Doležal and Miloš Ferjenčík*

Institute of Energetic Materials,
The University of Pardubice, CZ‒532 10 Pardubice, Czech Republic

Received: February 23, 2021; Accepted: April 9, 2021

The starting point for this article can be summed up in two sentences – "Human reliability can be assessed and predicted" and "The safety culture impact on human reliability can by modelled by performance influencing factors". In this paper, both points are justified before the authors address the question in the title. They recapitulate the latest developments both in the Human Reliability Assessment (HRA) and in the evaluation of safety culture. A combination of three tools (the IDHEAS method for HRA, 10 traits and 40 attributes for safety culture assessment, and the projection of results of assessment into the HRA with the use of Bayesian network models) was used in order to create a new model incorporating cultural and organizational aspects into the human reliability equation. The application of the model is illustrated by an example using the analyses available from the Tokai–Mura accident.

Keywords: Human error probability; Safety culture; Bayesian network

Introduction

The impulse for the paper presented herein was an article [1] which incorporates socio-technical elements into the estimates of Human Error Probabilities (HEPs) and applies the model to the accident that occurred in Tokai–Mura, Japan. This paper seeks to propose a more appropriate method for modelling organizational and cultural impacts on human performance.

---

* Corresponding author, ✉ milos.ferjencik@upce.cz

Section “Fundamentals” retraces the development of models for assessing human reliability and summarizes the basic concepts with regard to organizational and cultural factors. Section “Materials and methods” sums up the previous approaches on how to integrate cultural and organizational aspects into the model of human reliability. It identifies a combination of approaches that currently appears to be the most appropriate for the Human Reliability Assessment (HRA), involving organizational and cultural influences, and illustrates how the proposed combined approach could be applied to a real-life evaluation of human performance. The procedure that had failed in the Tokai–Mura accident was selected as the illustrative example. Section ”Results and discussion” illustrates steps to incorporate safety culture findings into HRA. The first of these steps contains an expert evaluation of the safety culture characteristics for the whole task. Further steps transform task-identified performance influencing factors (PIFs) with dominant safety culture characteristics when using Bayesian networks. The procedure results in a numerical estimation of safety culture’s effect on PIFs and produces numerical values for HEP. The overall probability of a negative event scenario is then calculated in traditional manner.

Fundamentals

The three generations of HRA evaluation methods

Human reliability analysis/assessment (HRA) is intended as a systematic process to evaluate human reliability. HRA tries to predict the probability of human errors that contribute to failures of complex systems. The HRA methods are particularly suitable for predicting errors by control room operators and similar actors.

Since 1975, when the WASH 1400 document [2] was published, we have been able to use the term, and speak about, human reliability analysis/assessment. In the years that followed, what we now know as the first generation HRA methods have been developed. First generation methods tried to be essentially “atomistic” – they encouraged the evaluator to break the task into small parts and assess the potential impact of factors, such as lack of time, device construction, stress, etc. By combining these elements, the evaluator could determine the cumulative probability of human error. The first generation can be represented by the THERP method [3].

In the last decade of the twentieth century, the development had begun of what was later called the second generation HRA methods. The respective methods are characterized by an approach that attempts to take into account the context which the first generation methods overlooked and the search for the so-called error of commission. The most frequently cited methods are ATHEANA [4], CREAM [5], MERMOS [6], CESA [7], and CAHR [8].

Recent developments have led to many comparative studies that tried to identify and improve the weaknesses inherent in HRA methods. Proposals for hybrid methods and an effort to expand the range of factors have been typical features of such attempts. Most recently, a new human reliability analysis method, the Integrated Decision – Tree Human Event Analysis System (IDHEAS) has been introduced in the NUREG-2199 document [9]. IDHEAS is a classic HRA method in the way that it combines both qualitative and quantitative steps:

- Qualitative task analysis leads to documented crew action paths in a Crew Response Tree (CRT).
- For each event in a CRT, applicable Crew Failure Modes (CFMs) are selected.
- Individual CFMs are quantified via so-called Decision Trees (DT).
- For each event, the HEPs are calculated by combining probabilities of the relevant CFMs.

The method allows human failure event dependency analysis and possible recovery actions. Performance Influencing Factors (PIFs) are used to characterize the content of task and probability of occurrence of a CFM. For simplicity, the IDHEAS developers chose to limit the number of PIFs in each DT to four.

Since introduction of the method, attempts have been made to combine modelling of PIFs in decision trees with Bayesian networks. It is a technically simple improvement with some benefits – “The Bayesian Network (BN) model supports practitioners in reasoning about the variables in the model.” [10]. This will be addressed in more detail in section “Materials and methods”.

Safety culture has an impact on human reliability

The title concept was originally introduced without any connection to business-related organizational culture theory. This approach arose after the accident of Chernobyl nuclear power plant in 1986; being developed in [11]. Connection to organizational culture theory started to be present in later International Atomic Energy Agency (IAEA) publications on safety culture, e.g. in [12] where it is stated: “Safety culture is that assembly of characteristics and attitudes in organizations and individuals which establishes that, as an overriding priority, nuclear plant safety issues receive the attention warranted by their significance.”

Another frequently cited definition of safety culture originates from the U.S. Nuclear Regulatory Commission (NRC) [13]. Maybe the greatest advantage of the second definition is its emphasis on behavior which seems to be a much more objective entity than values, attitudes, etc.

Unlike the previous definitions, the American Institute of Chemical Engineers (AIChE) has included its definition of safety culture as a part of complete set of organizational factors. According to [14], the pillars of a safety management system are: 1. Commitment to process safety, 2. Understanding Hazards and Risks, 3. Managing Risk, 4. Learning from Experience. Process safety culture represents a fundamental element of the first pillar. Requirements on safety culture and structured understanding of what they mean are described simply and clearly. Safety culture is set in this concept beside other organizational factors (so called elements).

This is one example, of how safety culture can be part of organizational factors, but this philosophy is widely used across the industry – for instance in aviation [15]. This approach has many advantages compared to today's nuclear practice.

Since the introduction of the safety culture term, there has been general agreement about its contribution to safety and understanding of the importance of the concept. Naturally, the question arises as to how to incorporate culture into the mathematical model of human performance – specifically into the HRA methods. With organizational factors, it is a very similar story, but its roots are much older and less easily traceable. So, from the viewpoint of the HRA, the two terms are treated with a certain portion of scientific confusion.

It cannot be said that organizational factors and safety culture are the same field or terms, but there is a clear overlap of both theoretical backgrounds. In cultural theory, we can find, for example, the so-called interpretative approach which destroys the difference between the concepts involving organization and culture. What we are trying to capture here is precisely this overlap of both terms. In many ways, this is a semantic problem in the area of the human aspects – so here we shall talk mostly about safety culture, but it can be translated (with appropriate caution) into the language of organizational factors.

The article [16] summarizes important ups and downs concerning safety culture practice. Another recent work is the French IRSN report [17]. These two contributions can serve as an introduction for everyone to today's challenges in applying a safety culture approach to everyday problems and seeking tangible evidence. The following text and presented model are also in line with organizational culture concept described e.g. in [18], which becomes a common framework in this field.

Over time, experts in the field of safety culture have struggled with finding evidence of any quantitative impact on safety. The first piece of such evidence, which would connect safety culture with overall safety, was presented in [19]. The article is a follow-up to conclusions of a series of meta-analytic studies published between 2006 and 2010, which significantly advanced the state of safety culture research. The studies included in these meta-analyses measured safety culture using surveys where employees had been asked various questions regarding their perceptions of the extent to which their organization valued safety.

The article [19] contains the following: "…examining the relationship between safety culture and a diverse set of performance measures that focus on the overall operational safety of a nuclear power plant." The significant correlations between overall safety culture and measures of safety performance ranged from −0.26 to −0.45, suggesting a medium effect and that safety culture accounts for 7–21 % of the variance in most of the measures of safety performance examined in this study.

Other and more tangible (but unfortunately only qualitative) evidence showing how safety culture and human reliability are tied together can be found in multiple investigation reports. Across a broad range of industries, with different degrees of precision, investigation reports show clear correlation between a degraded safety culture and its negative impact on event scenarios. In some of the reports, authors are not afraid to talk about direct causality of cultural behavior on event and describe the detailed mechanism. Examples can be found in investigation reports on the Chernobyl accident [20], the Challenger accident [21], the Royal Air Force Nimrod crash in 2006 [22], and the Deepwater Horizon explosion [23].

One conclusion can be drawn from all the above: Safety culture has an impact on human reliability, although a general mechanism of this impact is still rather unclear.

Human performance is modelled in most HRA methods using two basic tools: (1) by examination of the task to be performed and (2) by examination of the influencing factors. Depending on the sophistication of the method and its nomenclature, the influencing factors can be called Performance Shaping Factors (PSFs), Error Producing Conditions (EPCs), Error Forcing Contexts (EFCs), Common Performance Conditions (CPCs) and others. Philosophically, they all have the same function in the various HRA methodologies.

We will keep this approach and formulate the following hypothesis: As all other influencing phenomena, a safety culture impact in HRA can by modelled by performance influencing factors. The remaining part of this article will demonstrate the outcome of this hypothesis.

Materials and methods

Historical attempts to incorporate culture into HRA

Safety culture was difficult to recognize in the early HRA methods, but it was always (at least implicitly) present. First, it was an intuitive cultural background of nuclear industry, with no reason to describe it any way. As time went by, HRA pioneers and practitioners realized the need to pinpoint certain cultural aspects of human performance prediction. One example could be the table 20-15 in the THERP handbook and the section "Organizational Structure and Actions by Others" on pages 3–22 in the same publication [3]. Here, Swain and Guttmann actually talk about some aspects of safety culture and regulatory culture, but with terminology from a different era.

In addition, British HEART [24] considers cultural aspects in its influential factors (EPCs). Early HRA pioneers and practitioners did not have the safety culture concept at their disposal, but cultural aspects were always an implicit part of HRA.

Below we shall describe three attempts on which we have built our own conclusions concerning this topic.

The model from [25] represents a combination of THERP and CREAM and "how to be pessimistic" in prediction, in the presence of an alarming number of negative cultural indicators. As a theoretical concept, it is demonstrated in the example event but does not really show how it can be used if we try to apply it in a prospective way.

The model from [26] is a very clever and precise piece of work. The hybrid framework presented by the so-called Socio-Technical Risk Analysis (SoTeRiA) is detailed, and it is very hard to disagree with anything that is proposed. Nevertheless, this work has made a great leap in its sophistication and practical difficulty for field analysts. Even the best tools have to prove that the difficulty of working with them (including accommodation of knowledge of proper use) balances the benefits in terms of the results. This work may eventually be better appreciated in future.

The model produced in [1] tried to continue the work from [25]. The authors took as a starting point the THERP and CREAM methods. Then, they took a set of human factor / organizational characteristics from the Oil and Gas industry (OGP model). They added a group of ten experts for a semi-quantitative assessment of the importance of human factor characteristics. Afterwards, they used a previously known way to change original HEPs using new PSFs to combine all of these previous approaches together. This combination was then applied to the example event. The problem is that such a composite model works only if you are in the Oil and Gas industry and you have historical HRA data. Even though, in the introduction, the authors of article borrowed the quote from [27], they created a very "engineering" way to cope with the problem. With respect to their new way of thinking about safety introduced in [28], it can even be said that there is not a step forward, but just another step outside the way.

Other point is that the model from [1] certainly responds to a demand that exists. It is a clear approach based on familiar tools and proven methods, which considers the current feelings that practitioners and managers can experience. They see different organizations operating similar technology with different organizational and cultural characteristics or see a change in these characteristics within a single organization over time. And they want to see, how these differences are reflected in HRA predictions.

As mentioned above – the authors of historical HRA methods did not have the necessary terminology at the time, and especially the motivation, to examine safety culture. But some safety culture already existed in the 1960s, when the THERP tables had been compiled. And clearly, some safety culture characteristics were incorporated during the creation of CREAM and of NARA [29], which use [26], etc. What problems can be seen in historical attempts to incorporate culture into HRA?

Ethnocentricity is applied in methods and practical assessments of cultural and organizational factors in other organizations. “Ethnocentrism is the tendency to view the world through one’s own cultural filters” [30]. It is very disturbing to see how certain authors (to varying degrees) address the problems of others, while not critically applying some important knowledge in their own field. This manifests itself, mainly, in the very negative evaluation of some factors, or in the idealization of the past, other industry field habits, national influences, organizations etc.

A too vague description of the mechanism of organizational and cultural impact on human performance. There are many statements in the conclusions of major accident investigations in which a mechanism of the impact of culture (organizational factors) on safety is described. But overwhelmingly it is often only vaguely formulated without any deeper background. With such vagueness, we cannot avoid the question of whether the safety culture is not just another “usual suspect” or “scapegoat”. The problem of “another scapegoat” is discussed e.g. in the IRSN report [17].

Examples of application are not very helpful for executive or regulatory purposes. HRA with extended incorporation of cultural factors could be very beneficial in the aviation or nuclear industries. Nevertheless, it is necessary to be aware that these industries are heavily regulated and quite conservative. The conservative approach is typical both for regulators and also for workers in the field. For instance, NUREG-2165 [13] describes one of the attributes of healthy safety culture in the nuclear industry as “DM.2 Conservative Bias: Individuals use decision making practices that emphasize prudent choices over those that are simply allowable”. Using any tool which is not based on well-known and proven approaches can be very difficult in such an environment. Every new method has to be based on proven procedures and indicators and has to prove its usefulness at every important step.

Combined approach to the modelling of organizational and cultural impacts

Description of organizational and cultural factors

Our approach follows the views and philosophy contained in the Safety-II concept expounded by [28]. This philosophy, along with similar arguments in the article [16], simply says that the classic engineering position, taken from a technological safety point of view, has many limitations when it comes to applying it to people, and may often be even counterproductive.

If we talk about organizational culture, we cannot avoid mentioning the model by Schein [18]. His model has three distinct levels in organizational cultures with three different types of indicator. The deeper we go into levels – the more indirect and circumstantial evidence of culture indicators we find. As yet, this straightforward philosophy has not been successfully challenged and, especially in technical industries, has practically no intellectual competitor.

Based on the model [18], many simple concepts have been created capable of being understood by engineers without any psychological or sociological education, and still being sufficiently sophisticated not to be an obstacle for the other party and its demands for clear hierarchy of terms and the proper scientific research. Improving the Schein model for high-risk industry may not be complete and all the possibilities for finding new practical insights have probably not yet been exhausted. But it is not to be expected that this basic cultural thinking framework will be replaced soon. For this reason, one of our criteria for a cultural indicator framework is coherence with Schein model.

If we want to have reasonable confidence in our findings, we need a set of safety culture indicators, which are based on an accepted framework. Compiling safety culture indicators should be diversified – both methodologically and through the depth of any dependency. This requirement has been adequately described in [26] when using two basic principles: “Principle (L2): There are three different measurement methods: (i) objective (e.g. audit), (ii) subjective (e.g. perceptions/survey), and (iii) hybrid (as a combination of objective and subjective). Principle (L3): There also are three kinds of measurement bases: (i) direct (e.g. capturing organizational safety output, frequency of system accidents etc.), (ii) indirect (e.g. accounting for safety enablers or the safety causal factors, safety climate, safety practice etc.), and, again, (iii) hybrid (a combination of direct and indirect).” Safety culture indicators should also accept and cope with multiple levels of grouping people and the existence of subcultures.

We believe that the approach which meets all these criteria is described in NUREG-2165 [13] as 10 traits and 40 attributes for a healthy safety culture. This approach was not originally intended for incorporation into HRA. However, the reasons why we think that such incorporation is both possible and suitable are as follows – clarity of terms, scalability for different groups of people, subgroups, subcultures, international recognition and acceptance in nuclear industry, its generality, allowing various sources of hybrid information. In addition, recent research [19] shows that 10 traits and 40 attributes can reasonably measure and describe the complex and collective nature of culture in risk industries.

New HRA developments as our way forward

Recent developments in HRA methods show some promise for both confident incorporation of cultural aspects and reasonable results. For finding the new ways of overcoming historical HRA concepts, the IDHEAS method is a natural candidate. The method is described exhaustively in NUREG-2114 [31] and in the application guide NUREG-2199 [9]. It is based on a comprehensive literature review. In addition, IDHEAS uses the so-called "Macro-cognition Model for HRA", which can become a bridge to cultural findings.

We are aware that the IDHEAS method was originally developed specifically for use in analysis of at-power events in nuclear power plant control rooms. Using the quantification tables from IDHEAS to other types of events as we plan to doing so in our application example is not a valid use of the data. Nevertheless, we do not want to relinquish our idea to compare our approach with other ones to the same accident. In this paper, it is much more important to illustrate the structure of the model and procedure how it can be used than to get numerical results. This is why we keep on applying the IDHEAS method, a structure of which suits presented approach. We assume that the development of data tables suitable for exemplary case is possible within the structure of IDHEAS model. Also, a new development of IDHEAS-ECA [32] method clearly shows that this method can be used with proper cautions outside of nuclear power plant control rooms.

Other activities indicating a promising way of thinking will be used as a complementary tool. Authors of [10] show how to use Bayesian network models in HRA or in Probabilistic Risk Assessment (PRA). They do not offer too sophisticated a model, such as that proposed in [26], but show evolution in thinking using clearly understandable arguments.

In addition, international data collection activities offer an insight into human performance, and the necessary data to describe it, when dealing with complex engineered systems. The lack of a causal HRA structure and quantitative traceability is being addressed through the advanced modelling efforts.

The new model is based on the latest hybrid HRA/PRA methods utilizing classic binary event trees with a combination of Bayesian methods for PIF modelling. The BN models (also called Bayesian Belief Networks) have become increasingly popular within HRA as a means of addressing former shortcomings. Its ability to model cause and effect explicitly, combined with the ability to incorporate information from different sources is highly appreciated [10]. Also, a new development of this approach [33], involving transforming HRA using SACADA database data, shows real potential in this field.

The new model is rather evolutionary than revolutionary. If one looks at the recent development in this field [34], many people could have had this idea how to incorporate culture findings into HRA. Adding yet another layer of information which reflects this interest into BN is from this point of view another logical step in this endeavor.

Also new development in SACADA database [35] shows the increasing interest in cultural area in the field of data-collection activities. As a part of information from simulator training, data about leadership are now collected. Also, other part of so-called S.M.A.R.T / Operator Fundamentals has some cultural overlap. All this shows that development is alive.

Basic concept of our combined approach is shown in Figure 1.

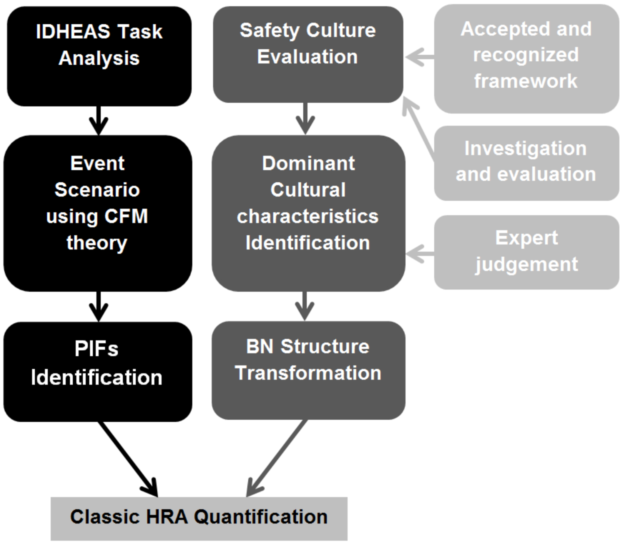

Fig. 1 Idea of combined approach to the modelling of organizational and cultural impacts

The resulting flow chart of our new approach to incorporating safety culture findings into HRA has these general steps:

a) Expert evaluation of the safety culture for the whole task (all actions in the event scenario) using accepted industry framework (10 traits and 40 attributes of healthy safety culture).
b) For all actions in the event scenario:
    b1) Transformation of identified actions into CFMs according to the IDHEAS methodology.
    b2) Identifying the relevant PIFs according to IDHEAS.
    b3) Causality connection of identified PIFs with dominant safety culture characteristics using BN according to the approach of [10].
    b4) Numerical estimation of the safety culture effect on PIFs and the resulting HEP numerical values.
c) Calculation of overall probability of a negative event scenario.

Practical details will be explained in the following example.

Results and discussion

Start of the illustrative IDHEAS task analysis

This section demonstrates the use of safety culture findings as a basis for HRA quantification. For illustrative reason and brevity, we choose the same scenario as that in [25], which is also used in [1]; see Fig. 2. We will use their event tree with a combination of IDHEAS and the 10 traits of healthy safety culture characteristics.

Only a part of the complete scenario analysis will be reproduced here. However, the scope of the example is sufficient to be used as a template to complete the analysis of the whole scenario.

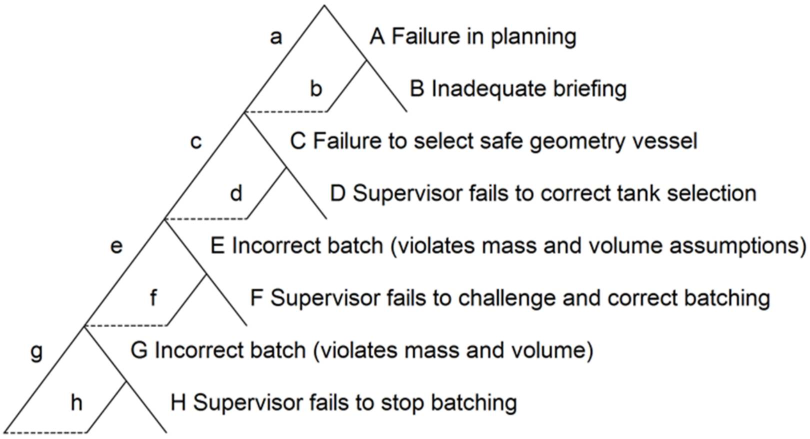

Fig. 2 Tokai–Mura event tree [25]

Safety culture evaluation

We have chosen evaluation of two different groups of operators. The first group consists of the JCO Co, Ltd. (JCO) operators who were at the scene of the Tokai–Mura accident. Safety culture findings for this evaluation are based on our knowledge about the Tokai–Mura accident, mainly from an excellent information summary [36]. In Table 1 below, we give our expert judgment assessment about the safety culture involved in this accident in the subculture of three workers involved in the homogenization process and pouring a solution into a precipitation tank. Errors in this procedure lead to a critical accident. For brevity, Table 1 reproduces assessments of the first 12 attributes while the original table has been prepared for the assessment of all 10 traits and 40 attributes according to [13]. The table also includes behavior patterns for other workers involved in the JCO accident report [36], mainly the company's management and supervisory team – particularly the chief technician for nuclear fuel, who had failed to stop the continuation in spite of being contacted before the job. It also includes non-specific managers who have created organizational deficiencies, mostly by duplicating some important safety positions (Production/Planning Group Chief, Quality Assurance/Safety Management Group Chief).

The second group comprises the hypothetical average nuclear power plant Control Room (CR) operators. This hypothetical group tries to represent what we believe is the average control room operator's characteristics across different nations and companies.

Our evaluation is given in Table 1. Two reasons made the evaluation difficult. First: Although the basis for making expert judgments of the Tokai–Mura accident has been a historical documentation, there is no basis provided for the ratings of the generic group of control room operators. An example below Table 1 shows the way of how knowledge from Tokai–Mura analyses has been applied. In case of control room operators, we assume that control room safety subculture is one of the healthiest safety subcultures across all nuclear power plant workers. For many of its characteristics and attributes, we assess it to be highly positive. For others, we assess it to be neutral and, in some attributes, we even admit that it could be negative. We are aware that validity of expert ratings is questionable, but we hope that for illustrative purpose in this paper they are satisfactory.

Second: A comparison between the JCO operators and Average CR operators seems to be an inappropriate mixing of contexts. The Tokai–Mura accident is a retrospective analysis, whereas making inferences about a generic control room is the prospective analysis. We hope that this mixing, although being difficult, may be useful for our illustration. It may show how different can be evaluation of safety culture of JCO operators that reflects historical experience and may be subject of hindsight bias from the would-be realistic prospective evaluation. Such a way illustrated by the results in two columns of Table 1 represent a possible range of expert estimates of safety culture and draw attention to how strongly the results of analyses involving safety culture may depend on expert judgments.

As you can see, it is very difficult to find anything positive in safety (sub) culture involved in the Tokai–Mura processing plant. In fact, using our reference framework, we did not find any positively assessed attribute. Communication of the work team shows some aspects of good behavior but does not involve any proper safety priority.

22 of 40 attributes were not assessed – either there had not been enough information to assess whether the attribute is negative, positive or neutral; or the attribute was not applicable in our example. In both cases, it has the same quality/value for us.

18 of 40 attributes were assessed negatively. For each of these assessments we find one or more pieces of evidence in the Tokai–Mura report [36]. Here is one example of such evidence:

For attribute LA.2 Field Presence: "Leaders are commonly seen in working areas of the plant observing, coaching, and reinforcing standards and expectations. Deviations from standards and expectations are corrected promptly."

On page 188 in [36] we can find these observations giving an argument for negative assessment:

"There were two people who could stop the three workers from using the precipitation tank. One was the workers' supervisor," also: "His other job was doing a round inspection during work in the conversion building and checking work progress, etc. According to the investigation, he did the inspection at least once a day until 29 September.

However, he failed to detect their pouring of some 16 kg of uranium into the precipitation tank." Last piece of evidence: "He did not stop them using the precipitation tank. He was quoted as saying he had confused jobs in the first and second fabrication facilities and the conversion building."

Table 1 Expert judgement of the Tokai–Mura crew subculture characteristics
First 12 of 40 attributes

LA.1 Resources: Leaders ensure that personnel, equipment, procedures, and other resources are available and adequate to support nuclear safety.	JCO operators: **Negative**	Average CR operators: **Positive**
LA.2 Field Presence: Leaders are commonly seen in working areas of the plant observing, coaching, and reinforcing standards and expectations. Deviations from standards and expectations are corrected promptly.	JCO operators: **Negative**	Average CR operators: **Positive**
LA.3 Incentives, Sanctions and Rewards: Leaders ensure incentives, sanctions, and rewards are aligned with nuclear safety policies and reinforce behaviors and outcomes that reflect safety as the overriding priority.	JCO operators: **Not assessed**	Average CR operators: **Neutral**
LA.4 Strategic Commitment to Safety: Leaders ensure plant priorities are aligned to reflect nuclear safety as the overriding priority.	JCO operators: **Negative**	Average CR operators: **Neutral**
LA.5 Change Management: Leaders use a systematic process for evaluating and implementing change so that nuclear safety remains the overriding priority.	JCO operators: **Negative**	Average CR operators: **Positive**
LA.6 Roles, Responsibilities, and Authorities: Leaders clearly define roles, responsibilities, and authorities to ensure nuclear safety.	JCO operators: **Negative**	Average CR operators: **Positive**
LA.7 Constant Examination: Leaders ensure that nuclear safety is constantly scrutinized through a variety of monitoring techniques, including assessments of nuclear safety culture.	JCO operators: **Negative**	Average CR operators: **Neutral**
LA.8 Leader Behaviors: Leaders exhibit behaviors that set the standard for safety.	JCO operators: **Not assessed**	Average CR operators: **Neutral**
PI.1 Identification: The organization implements a corrective action program with a low threshold for identifying issues. Individuals identify issues completely, accurately, and in a timely manner in accordance with the program.	JCO operators: **Negative**	Average CR operators: **Neutral**
PI.2 Evaluation: The organization thoroughly evaluates problems to ensure that resolutions address causes and extent of conditions, commensurate with their safety significance.	JCO operators: **Not assessed**	Average CR operators: **Neutral**
PI.3 Resolution: The organization takes effective corrective actions to address issues in a timely manner, commensurate with their safety significance.	JCO operators: **Negative**	Average CR operators: **Neutral**
PI.4 Trending: The organization periodically analyzes information from the corrective action program and other assessments in the aggregate to identify programmatic and common cause issues.	JCO operators: **Not assessed**	Average CR operators: **Neutral**

As a result – altogether five dominant safety culture characteristics were identified for the first group:

- Leadership Safety Values and Actions (LA): 6 of 8 attributes were assessed negatively.
- Problem Identification and Resolution (PI): 2 of 4 attributes were assessed negatively.
- Work Processes (WP): 4 of 4 attributes were assessed negatively.
- Questioning Attitude (QA): 2 of 4 attributes were assessed negatively.
- Decision making (DM): 2 of 3 attributes were assessed negatively.

Every model must, in principle, commit a simplification of reality. The use of the only dominant characteristics in further modelling process is one of them. We are aware that other identified characteristics have an impact on human performance, but we expect it to be relatively small. Therefore, we do not model them. In general, it is possible to include these "weaker" characteristics. But it would only increase the time required to work with the model and computational demands leading to loss of practicality.

For the second group, the hypothetical average control room operators, we have the following arguments on which we based our evaluation:

Attributes are assessed mostly positively. This correlates with our experience that control room operators, by the nature of their work, simulator training, etc. have high safety values and attitudes. Their partly isolated workplace helps them to create a shared collective “operator identity” [37].

Our expert judgement is based on central-European experience, but we assume it can be extrapolated to an international level. “There appears to be a relatively homogenous “operating culture” existing at all the nuclear power plants.” [37].

From our assessment of hypothetical CR operators – we found again five dominant safety culture characteristics:

- Leadership Safety Values and Actions (LA): 4 of 8 attributes were assessed positively.
- Personal Accountability (PA): 2 of 3 attributes were assessed positively.
- Work Processes (WP): 3 of 4 attributes were assessed neutral/negatively, 1 attribute positively.
- Questioning Attitude (QA): 3 of 4 attributes were assessed positively, 1 negatively.
- Decision making (DM): 2 of 3 attributes were assessed positively.

The next sections will show how these findings can be transformed into an HRA quantification model using BN.

Transformation of identified actions into CFMs

The IDHEAS method uses the concept of CFMs. For the first action, originally called by [1] and [25] “Failure in planning”, we choose the CFM “Choose inappropriate strategy”. As IDHEAS is originally intended only for a nuclear power plant’s control room in the full power mode, it needs some creativity to adjust it for the Tokai–Mura accident. In [9] is CFM described: “For this CFM, the crew has entered the correct procedure presented with more than one alternative for how to proceed. The crew chooses the wrong alternative, leading to the human failure event. This CFM assumes the crew has the correct mental model for the scenario up until this point.” Another description is “if the crew has a strong preference to choose an inappropriate option for the scenario over the appropriate alternative”. Also “This CFM is applicable where the crew has choices at a particular point in a procedure for how to execute their response and corresponds to a lower level of strategic decision-making. Furthermore, it assumes that a deliberate choice is made. This CFM also covers cases where there is judgment left to the operator (e.g., external events, implementation of severe accident management guidelines (SAMGs)” [9].

We believe this CFM is the best choice from all the IDHEAS method offers and represents properly the strategy, of the Tokai–Mura crew. The crew should know what to do in the homogenization process (correct mental model), the workers were anxious to finish the job at the conversion building and they decided for use of precipitation tank instead of the buffer column [36] as a strategic decision to do it.

Identifying the relevant PIFs

Decision tree for the Crew Failure Mode “Choose Inappropriate Strategy” is reproduced in Fig. 3. This CFM has two dominant PIFs: 1. Preference for appropriate strategy, 2. Advantages to appropriate strategy. Table 2 shows data for this CFM. They originate from NUREG-2199 [9].

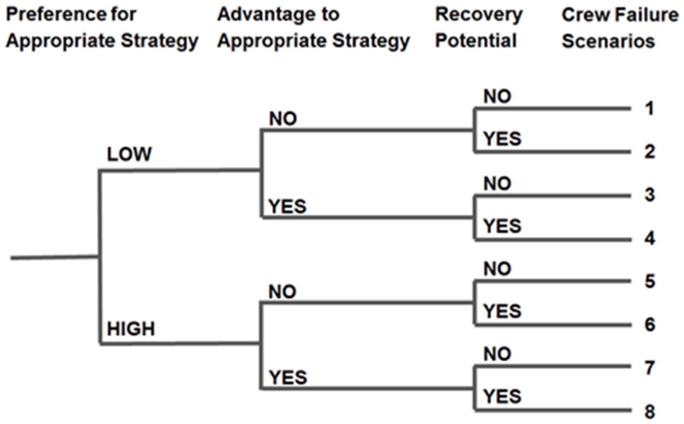

Fig. 3 Decision tree for the CFM “Choose inappropriate strategy” [9]

Table 2 Probabilities table for CFM “Choose inappropriate strategy” (RP-2) [9]

DT path	PIFs (DT branch point) Preference for appropriate strategy	 Advantage to appropriate strategy	 Recovery potential	5%-tie	50%-tie	99%-tie	Mean
1	Low	No	No	$6.0{\cdot}10^{-2}$	$5.0{\cdot}10^{-1}$	$9.0{\cdot}10^{-1}$	$5.2{\cdot}10^{-1}$
2	Low	No	Yes	$5.0{\cdot}10^{-3}$	$5.0{\cdot}10^{-2}$	$5.0{\cdot}10^{-1}$	$8.2{\cdot}10^{-2}$
3	Low	Yes	No	$5.0{\cdot}10^{-3}$	$1.0{\cdot}10^{-1}$	$7.0{\cdot}10^{-1}$	$1.4{\cdot}10^{-1}$
4	Low	Yes	Yes	$5.0{\cdot}10^{-4}$	$1.0{\cdot}10^{-2}$	$7.0{\cdot}10^{-2}$	$1.4{\cdot}10^{-2}$
5	High	No	No	$2.0{\cdot}10^{-3}$	$2.0{\cdot}10^{-2}$	$2.0{\cdot}10^{-1}$	$3.3{\cdot}10^{-2}$
6	High	No	Yes	$2.0{\cdot}10^{-4}$	$2.0{\cdot}10^{-3}$	$2.0{\cdot}10^{-2}$	$3.3{\cdot}10^{-3}$
7	High	Yes	No	$1.0{\cdot}10^{-5}$	$3.0{\cdot}10^{-3}$	$1.0{\cdot}10^{-1}$	$9.3{\cdot}10^{-3}$
8	High	Yes	Yes	$1.0{\cdot}10^{-6}$	$3.0{\cdot}10^{-4}$	$1.0{\cdot}10^{-2}$	$9.3{\cdot}10^{-4}$

BN causal connection of PIFs with dominant safety culture characteristics

Figure 4 incorporates our safety culture dominant characteristics into this CFM. Our dominant characteristics make another layer over the two dominant PIFs for this particular CFM.

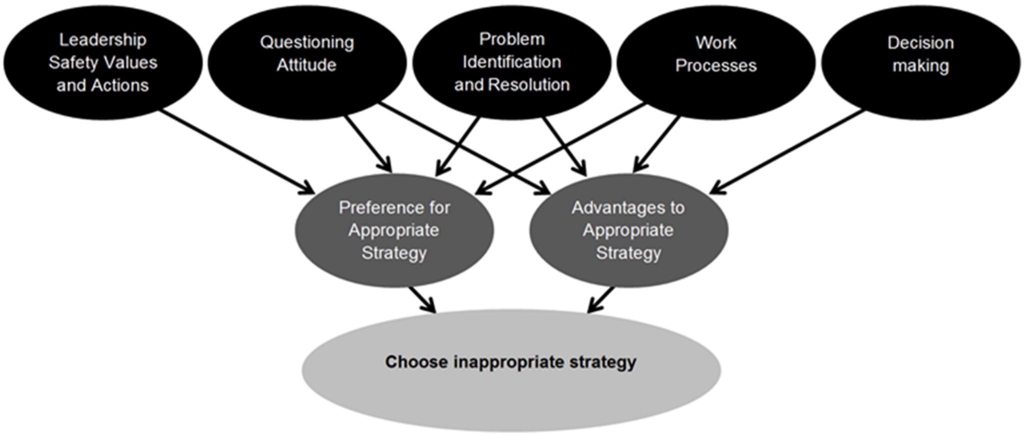

Fig. 4 BN for the CFM “Choose inappropriate strategy” that corresponds to the original DT model from [9]

Decision on which safety culture attributes are related to which performance influencing factors in the IDHEAS method decision trees is another expert judgement input. Such a “mapping” is one of the major hurdles to incorporating safety culture into HRA. There is insufficient data regarding the causal pathways that link safety culture attributes to specific aspects of human performance. For this reason, we connected maximum of all dominant characteristics with proposed

PIFs. Only when we did find an argument, for which this connection would not be reasonable, we didn't consider it in model. As a result – in our model is at least 60% of dominant safety culture characteristics connected to all performance influencing factors.

Example of above-mentioned arguments relates to missing connection of Leadership Safety Values and Actions (LA) and Decision making (DM) dominant safety characteristics to some PIFs. Both characteristics are related to management commitment to safety and overlapping. But LA characteristic can be much more abstract or symbolic. In some cases, we can say, that Leadership Safety Values and Actions lead workers to desirable behavior, Decision making creates conditions how to implement this behavior. For this reason, we used for some PIFs only one of them, based on judged level of abstraction management values are reflected on specific PIF.

The example illustrating the identification of dominant cultural characteristics is shown in the next section.

Dominant cultural characteristics identification

Figure 5 shows the result of the first step of transforming the usual PIF representation for HEP calculation into BN in the IDHEAS HRA method by [10]. Figure 6 interprets BN model with causal details after node reduction – there is more to find in the original article.

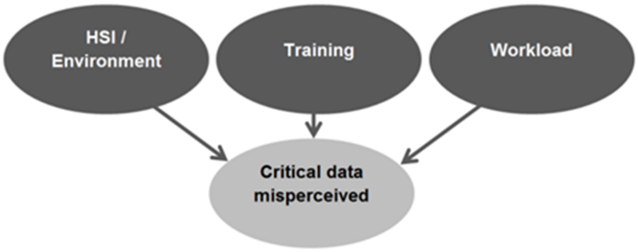

Fig. 5 Bayesian Network for the CFM "Critical data misperceived" that corresponds to the original decision tree model [10]

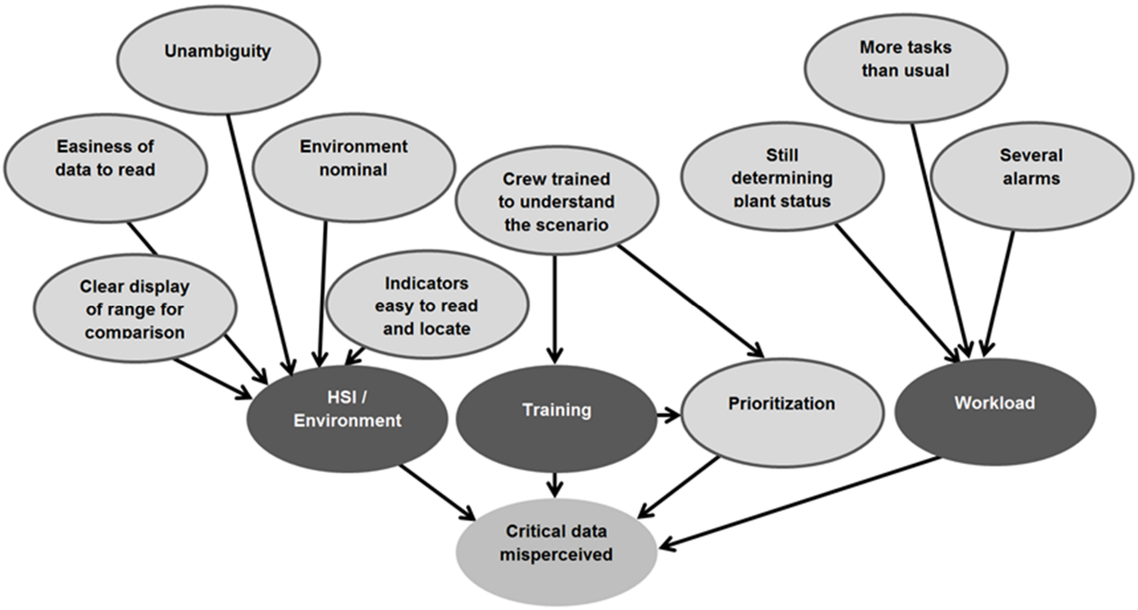

Fig. 6 Expanded BN for the CFM "Critical data misperceived" [10]

The advantage of this approach is that it is fully prepared for integration of safety culture/organizational factors. This integration into a BN network can be done with multiple levels of decomposition. Figure 7 is an exemplary case of one way of how the integration could result.

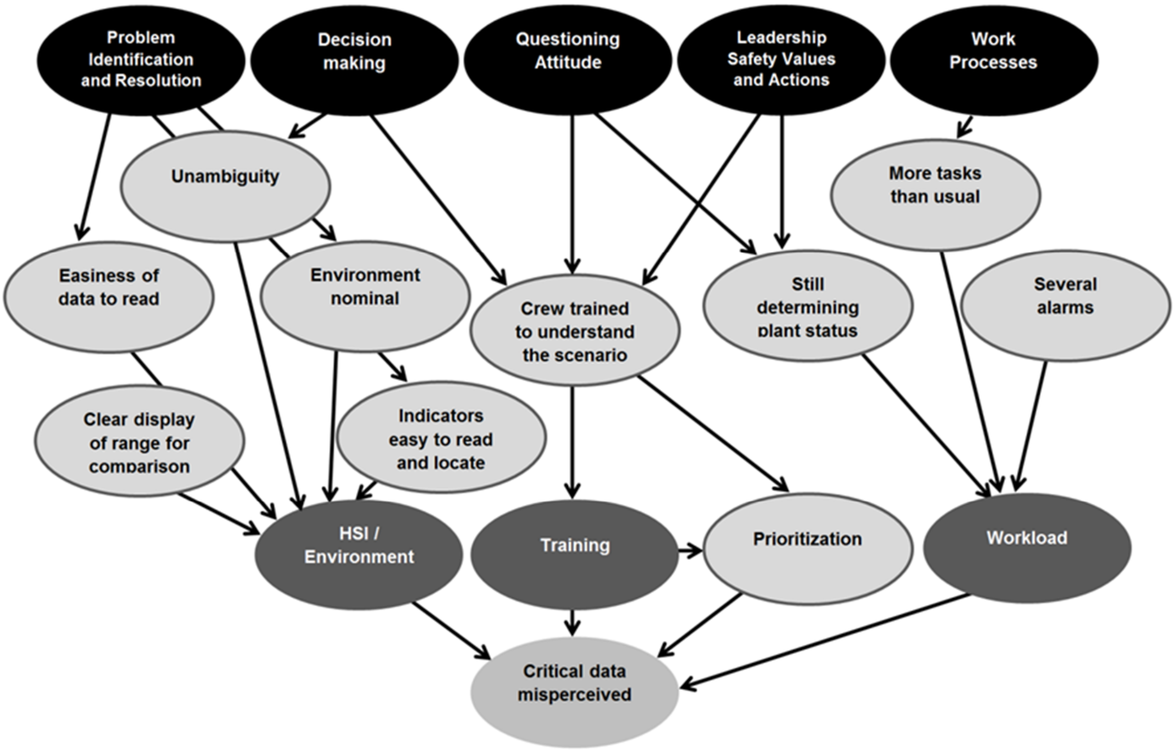

Fig. 7 Illustrative example of safety culture indicators incorporated into BN.

This incorporation could be even easier if we build new a BN layer only above the identified PIFs. In this case, the proposed model is very simple, but we lose the detailed causal relationships and an interesting interpretation mechanism. An illustrative example is shown in Figure 8.

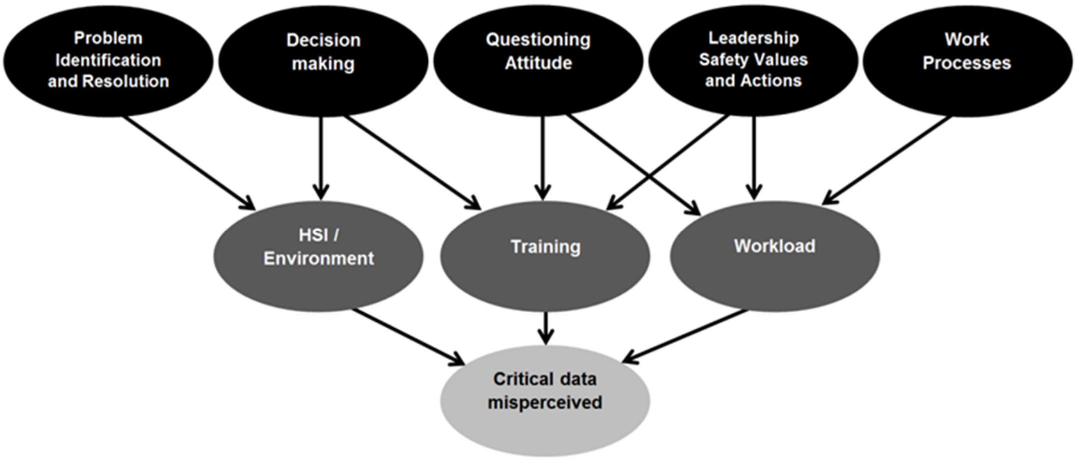

Fig. 8 Illustrative example of simple incorporation of safety culture indicators into BN.

In these examples, we used safety culture characteristics from NUREG-2165 [13] as a general summarization of 40 particular attributes. It will be shown in the next section, how to treat mathematically practical questions of safety culture indication quantification into BN.

Numerical estimation of the safety culture effects

The IDHEAS method provides a table for quantification of HEP of the CFM. It is reproduced in Table 2. Data from this table can be directly used in the BN software (4 lines without recovery potential).

Then, we have to quantify the effect of negative attributes on the overall safety culture characteristics. For illustrative reasons and brevity, we used a simple proportional (linear) formula where the probabilities of a negative effect of a safety culture characteristic relates to the number of negative aspects from the list. As Work Processes (WP) include 4 of 4 attributes assessed negatively – the probability of a negative effect is 100%. Questioning Attitude (QA) has 2 of 4 attributes assessed negatively – with same formula we assessed the probability of a negative effect as 50%. The same approach has been done in other cases.

Alternatives (exponential or logarithmic) to such a linear quantification formula can be used. The question of the most suitable formula, and its reasoning, is a topic for further critical exploration.

A key part of quantification lies in quantification of different safety culture characteristics combination to occurrence of PIF. This quantification may take the form of a table with all the possible influencing characteristics combinations – in our example case both PIFs have four influencing inputs – giving 16 (2^4) unique combinations for each PIF. Quantification of the effect on these combinations on the probability of PIF is, in our model, pure expert opinion. Of course – if you assume that all the characteristics in their numeric effect are equal, the values in the table can be calculated using a single formula. This should be based on expert opinion of the non-linearity of the cumulative effect of negative safety culture characteristics. Table 3 shows an example.

Table 3 Expert opinion on the cumulative effect of negative safety culture characteristics on PIF probability

Leadership Safety Values and Actions	Questioning Attitude	Problem Identification and Resolution	Work Processes	High [%]	Low [%]
Negative	Negative	Negative	Negative	30	70
Negative	Negative	Negative	Positive	50	50
Negative	Negative	Positive	Negative	50	50
Negative	Negative	Positive	Positive	80	20
Negative	Positive	Negative	Negative	50	50
Negative	Positive	Negative	Positive	80	20
Negative	Positive	Positive	Negative	80	20
Negative	Positive	Positive	Positive	95	5
Positive	Negative	Negative	Negative	50	50
Positive	Negative	Negative	Positive	80	20
Positive	Negative	Positive	Negative	80	20
Positive	Negative	Positive	Positive	95	5
Positive	Positive	Negative	Negative	80	20
Positive	Positive	Negative	Positive	95	5
Positive	Positive	Positive	Negative	95	5
Positive	Positive	Positive	Positive	100	0

As you can see – our expert judgment in this case can be interpreted as a slightly S-curved non-linear function with a maximum of 70% negative effect of safety culture on PIFs probability. It can also be translated into a graph as depicted in Figure 9.

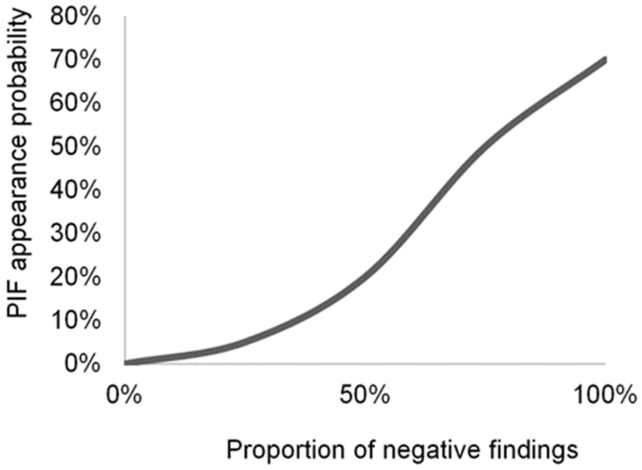

Fig. 9 Cumulative effect of negative safety culture characteristics

A very similar S-curved non-linear function for the cumulative effect was actually the result of expert opinion in other quantification tables. Other than presented cumulative quantification judgment can be used. This question and its reasoning are also a topic for further critical exploration.

Figure 10 shows how the incorporation can look in the Netica software tool. Probability of error in this task A is calculated to be 0.214.

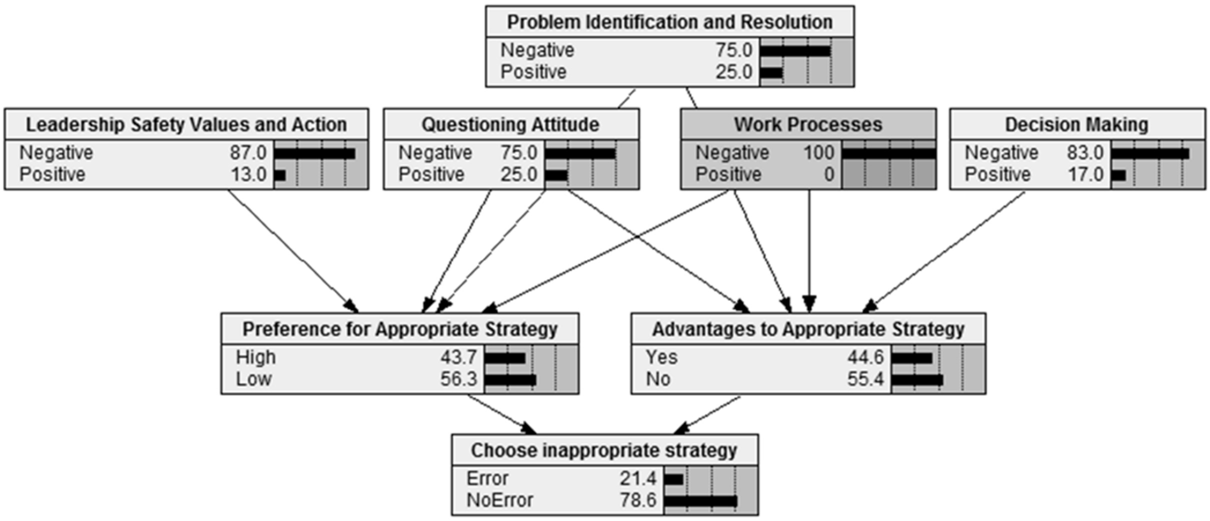

Fig. 10 BN for the CFM "Choose inappropriate strategy" (Failure in planning) in Netica SW for JCO operators

Figure 11 shows how the new model looks for hypothetical and average control room operators. There is a slightly different set of dominant safety culture characteristics and a completely different effect of their negativity/positivity. As a result, the probability of error in this task (task A) is calculated to be 0.0385 (probability decreased more than five times).

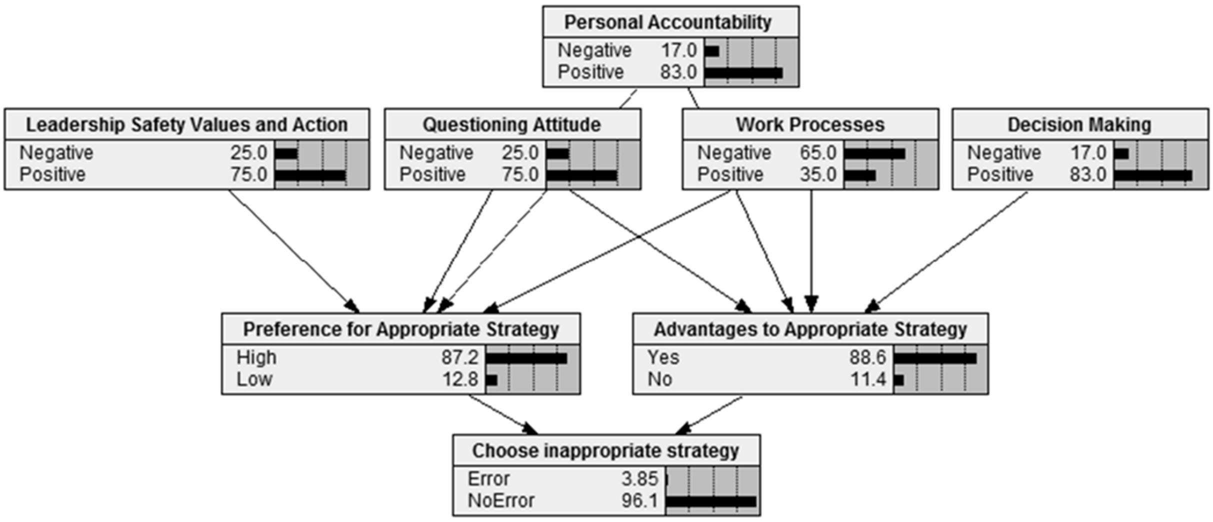

Fig. 11 BN for the CFM "Choose inappropriate strategy" (Failure in planning) in Netica SW for hypothetical control room operators

Calculation of overall probability of a negative event scenario

This procedure can be applied to all actions in the event scenarios. The list of CFMs used for the whole scenario according to Figure 5 is in Table 4.

Table 4 IDHEAS CFMs used for example event tree

Task description by Gertman and Blackman [25]	Closest CFM of IDHEAS method
A – Failure in planning	Choose inappropriate strategy (RP-2)
B – Inadequate briefing	Miscommunication (C-1)/Critical data not checked (E-2)
C – Failure to select safe geometry vessel	Data misleading or not available (SA-1)
D – Supervisor fails to correct tank selection	Critical data misperceived (SA-3)
E – Incorrect batch (violates mass and volume assumptions)	Misinterpret procedures (RP-1)
F – Supervisor fails to challenge and correct batching	Wrong data source attended to (SA-2)
G – Incorrect batch (violates mass and volume)	Misinterpret procedures (RP-1)
H – Supervisor fails to stop batching	Wrong data source attended to (SA-2)

It was not possible to find a single CFM to represent action B – Inadequate briefing – and so a combination of two IDHEAS CFMs – Miscommunication and Critical data not checked – were selected and assigned to action B. Action B, as described in [36], possesses some aspects of both the CFMs. In addition, the decision tree for the Miscommunication CFM was not quantified in the moment in IDHEAS tables. For this reason, we used a combination of both CFMs and PIFs from related decisions trees. Then, we used quantified numerical values from IDHEAS E-2 table. The final calculated probabilities for all nodes in the event tree for the both groups of operators are in the following table.

Table 5 Calculated probabilities for all nodes in the event tree.

Event tree node	JCO operators from real scenario	Hypothetical average control room operators
A	0.214	0.039
B	0.102	0.009
C	0.170	0.036
D	0.304	0.039
E	0.127	0.035
F	0.047	0.014
G	0.127	0.035
H	0.047	0.014

Conclusions

In this article, we have tried to summarize the major problems which we have encountered in our colleagues' previous attempts to incorporate safety culture/ organizational characteristics into HRA. Then, we have explained how we would handle such characteristics with our own approach.

We applied the new approach to the sadly known Tokai–Mura accident. One could get the impression from reading the official investigation reports that it was an inevitable outcome and that the true probability estimation should be close to 100%. However, we expected the total result to be somewhere between 5 and 25 percent. After calibration of the computational model on the control group of hypothetical operators, we have inserted our best judgement and we get the result in which we believe to be defendable.

The approach is based on recognized methods and can therefore be expected to be used in highly regulated sectors. At this moment, it strongly relies on expert judgment. On the other hand – it is scalable and open to updates and improvements using data collection activities. We have tried to show that its strength lies not in its ability to change (and justify) the numerical values of human error probabilities, but rather in its variable, flexible, and self-explanatory philosophy. The new approach also hides a lot of potential for improvement – for instance in moving from the more-or-less general safety culture characteristics to more specific attributes. For the attentive reader this is only the beginning. Of course, any such activity will be struggling with the increasing difficulty of data sources, scientific evidence and mostly the practicality. As human factor specialists, we live at an amazing time. We finally have formal tools enabling capturing the safety culture and drawing up a picture of safety climate. These tools are not perfect, but being good enough to address all the required angles and to provide a common language for understanding each other. HRA methods finally get through the barrier of probability quantification tables dating from the 1960s. Young experts have the passion and mathematical ability to take us to a new level of confidence in the probability prediction. All the tools already exist, but let us connect them in a modern way.

Acronyms

BN	Bayesian Network
CFM	Crew Failure Mode
CR	Control Room
CRT	Crew Response Tree
DT	Decision Tree
EPC	Error Producing Condition
HEP	Human Error Probability
HRA	Human Reliability Assessment
JCO	Japanese company having operated the Tokai–Mura plant
PIF	Performance Influencing Factor
PRA	Probabilistic Risk Assessment
PSF	Performance Shaping Factor

References

[1] Ribeiro A.C., Sousa A.L., Duarte J.P., Frutuoso e Melo P.F.: Human reliability analysis of the Tokai–Mura accident through a THERP-CREAM and expert opinion auditing approach. *Safety Science* **87** (2016) 269–279.

[2] Rasmussen N.: *Reactor safety study. An assessment of accident risks in U.S. commercial nuclear power plants. Executive summary: Main report.* WASH-1400; NUREG-75/014. U.S. Nuclear Regulatory Commission, Washington, D.C. 1975.

[3] Swain A.D., Guttmann H.E.: *Handbook of Human Reliability Analysis with emphasis on Nuclear Power Plant Applications – Final report.* NUREG/CR-1278. U.S. Nuclear Regulatory Commission & Sandia National Laboratories, Washington, D.C. 1983.

[4] Cooper S.E., Ramey-Smith A.M., Wreathall J., Parry G.W., Bley D.C., Luckas W.J., Taylor J.H., Barriere M.T.: *A Technique for Human Error Analysis (ATHEANA).* NUREG/CR-6350. U.S. Nuclear Regulatory Commission, Washington, D.C. 1996.

[5] Hollnagel E.: *Cognitive Reliability and Error Analysis Model (CREAM).* Elsevier Science, 1998.

[6] Meyer P., Le Bot P., Pesme H.: MERMOS : An extended second generation HRA method, in: *2007 IEEE 8th Human Factors and Power Plants and HPRCT 13th Annual Meeting*. IEEE, 2007.

[7] Reer B.: The CESA method and its application in a plant-specific pilot study on errors of commission. *Reliability Engineering and System Safety* **83** (2004) 187–205.

[8] Sträter O.: *Cognition and safety: An Integrated approach to systems design and assessment*. Ashgate, Hampshire 2005.

[9] Xing J., Parry G., Presley M., Forester J., Hendrickson S., Dang V.: *An Integrated Human Event Analysis System (IDHEAS) for Nuclear Power Plant Internal Events At-Power Application.* NUREG-2199. Volume 1. U.S. Nuclear Regulatory Commission & EPRI, Washington, D.C. 2017.

[10] Zwirglmaier K., Straub D., Groth K. M.: Capturing cognitive causal paths in human reliability analysis with Bayesian network models. *Reliability Engineering and System Safety* **158** (2017) 117–129.

[11] INSAG: *Safety culture. A report by the International Nuclear Safety Advisory Group.* Safety Series No.75, INSAG-4. International Atomic Energy Agency, Vienna 1991.

[12] IAEA: *Regulatory Oversight of Safety Culture in Nuclear Installations.* IAEA-TECDOC-1707. International Atomic Energy Agency, Vienna 2013.

[13] Keefe M., Frahm R., Martin K., Sigmon R., Shoop U., Morrow S., Sieracki D., Powell R., Shaeffer S., Rutkowski J., Ruesch E.: *Safety Culture Common Language.* NUREG-2165. U.S. Nuclear Regulatory Commission. Washington, D.C. 2014.

[14] Grossel S.S.: *Guidelines for Risk Based Process Safety.* AIChE Center for Chemical Process Safety, Wiley-AIChE, New York 2008.

[15] ICAO: *Safety Management Manual (SMM). Third Edition.* International Civil Aviation Organization, Montréal 2013.

[16] Strauch B.: Can we examine safety culture in accident investigations, or should we? *Safety Science* **77** (2015) 102–111.

[17] Gisquet E., Levy E., Jeffroy F.: *Addressing Cultural Aspects of Organisations in High Risk Industries.* Report No. PSN-SRDS/SFOHREX No. 2017-0005, 2017. http://www.irsn.fr/EN/publications/technical-publications/Documents/IRSN-PSN-SRDS-SFOHREX-2017-005_cultural-aspects-of-organisations-in-high-risk-industries.pdf (accessed on 16 Oct 2017)

[18] Schein E.H.: *Organizational Culture and Leadership, Second Edition.* Jossey Bass Publishers, San Francisco 1992.

[19] Morrow S.L., Kenneth Koves G., Barnes V.E.: Exploring the relationship between safety culture and safety performance in U.S. nuclear power operations. *Safety Science* **69** (2014) 37–47.

[20] INSAG: *The Chernobyl Accident, Updating of INSAG-1: A Report by the International Nuclear Safety Advisory Group.* INSAG-7. International Atomic Energy Agency, Vienna 1992.

[21] Rogers Commission: *Report of the Presidential Commission on the Space Shuttle Challenger Accident.* United States Government Printing, Washington, D.C. 1986.

[22] Haddon-Cave C.: *The Nimrod Review: An independent review into the broader issues surrounding the loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006.* Stationery Office, London 2009.

[23] USCG: *Report of Investigation into the Circumstances Surrounding the Explosion, Fire, Sinking and Loss of Eleven Crew Members Aboard the Mobile Offshore Drilling Unit Deepwater Horizon in the Gulf of Mexico, Volume I.* MISLE Activity Number: 3721503. U.S. Coast Guard, Washington, D.C. 2011.

[24] Williams J.C.: HEART – A proposed method for achieving High Reliability in Process Operation by means of Human Factors Engineering Technology, in: *Proceedings of a Symposium on the Achievement of Reliability in Operating Plant, Safety and Reliability Society.* NEC, Birmingham 1985.

[25] Gertman D.I., Blackman H.S.: Sensitivity analysis case study: Incorporating organizational factors in HRA, in: *Proceedings of the international workshop building the new HRA: Errors of commission from research to application.* NEA/CSNI/R (2002) 3. Nuclear Energy Agency, Issy-les-Moulineaux 2002, pp. 195–207.

[26] Mohaghegh Z., Kazemi R., Mosleh A.: Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: A hybrid technique formalization. *Reliability Engineering and System Safety* **94** (2009) 1000–1018.

[27] Hollnagel E.: *Barriers and accident prevention.* Ashgate, Hampshire 2004.

[28] Hollnagel E.: *Safety–I and Safety–II, The past and future of Safety Management.* Ashgate. Hampshire 2014.

[29] Kirwan B., Gibson W.H., Edmunds J., Kennedy R.J., Cooksley G.: *Technical basis for NARA, A method of Human Error Quantification.* Report CRA-BEGL-POW-J032 R1, Issue 7. EDF Energy Nuclear Generation, Barnwood 2011.

[30] Matsumoto D., Juang L.: *Culture and psychology. Third edition.* Wadsworth Publishing, Belmont 2003. ISBN 0534535917.

[31] Whaley A.M., Xing J., Boring R.L., Hendrickson S.M.L., Joe J.C., Le Blanc K.L., Morrow S.L.: *Cognitive basis for Human Reliability Analysis (NUREG-2114).* U.S. Nuclear Regulatory Commission & Sandia National Laboratories, Washington, D.C. 2016.

[32] Xing J., Chang Y.J., DeJesus J.: *Integrated Human Event Analysis System for Event and Condition Assessment (IDHEAS-ECA)*. RIL-2020-02, U.S. Nuclear Regulatory Commission, ADAMS Accession No. ML20016A480, Washington, D.C. 2020.

[33] Groth K.M., Tedeschi J.D., Moradi R.: Transforming HRA using SACADA data, Bayesian methods, and DBNs, in: *NRC Human Reliability Analysis Data Workshop*, 15.–16. March 2018. https://www.nrc.gov/docs/ML1807/ML18079A162.pdf (accessed on 28 Jan 2019)

[34] Groth K., Smith R., Moradi R.: A hybrid algorithm for developing third generation HRA methods using simulator data, causal models, and cognitive science. *Reliability Engineering and System Safety.* **191** (2019) 106507.

[35] Chang J.Y., Bley D., Criscione L., Kirwan B., Mosleh A., Madary T., Nowell R., Richards R., Roth E., Sieben S., Zoulis A.: The SACADA database for human reliability and human performance. *Reliability Engineering and System Safety.* **125** (2014) 117–133.

[36] Furuta K., Sasou K., Kubota R., Ujita H., Shuto Y., Yagi E.: Human factor analysis of JCO criticality accident. *Cognition, Technology & Work* **2** (2000) 182–203.

[37] Rollenhagen C., Westerlund J., Näswall K.: Professional subcultures in nuclear power plants. *Safety Science* **59** (2013) 78–85.