Markus Schöbel, Anne Klostermann, Ruth Lassalle, Johannes Beck,
Dietrich Manzey

Digging deeper! Insights from a multi-method assessment of safety culture in nuclear power plants based on Schein's culture model

Schöbel, M., Klostermann, A., Lassalle, R., Beck, J., & Manzey, D. (2017). Digging deeper! Insights from a multi-method assessment of safety culture in nuclear power plants based on Schein's culture model. Safety Science, 95, 38–49. https://doi.org/10.1016/j.ssci.2017.01.012

Digging deeper! Insights from a multi-method assessment of safety culture in nuclear power plants based on Schein's culture model

Authors: Markus Schöbel[a1], Anne Klostermann[b], Ruth Wagner[b], Johannes Beck[b] & Dietrich Manzey[b]

Affiliations:

[a] Center for Economic Psychology, Department of Psychology, University of Basel, Missionsstrasse 62a, 4055 Basel, Switzerland.

[b] University of Technology Berlin, Department of Psychology and Ergonomics, Work, Engineering & Organizational Psychology, Marchstr. 12, D-10587 Berlin

[1]**Corresponding author:**

Dr. Markus Schöbel
Center for Economic Psychology
Department of Psychology
University of Basel
Missionsstrasse 62A
4055 Basel, Switzerland
Phone: +41-61-2070586
m.schoebel@unibas.ch

Abstract

This article describes the development and application of a safety culture assessment approach for nuclear power plants based on Schein's culture model. The developed multi-method approach aimed to unfold deeper levels of culture, and, at the same time, to be applicable by practitioners and transparent in producing meaningful results. In Study 1 we describe the development of the method and its application in a German nuclear power plant. Study 2 presents a cross-validation of the approach in a second German nuclear power plant. The evaluation results of the approach reveal adequate validity with regard to the obtained results and its perceived linkage to safety culture in both plants. It becomes evident that the approach allows for deriving basic assumptions of plant members and demonstrating their significance for safety performance. Moreover, it can be shown that it is worthwhile to go beyond the assessment of artifacts and espoused values in understanding cultural dynamics on a plant level. Finally, insights and limitations of the developed approach are discussed and reflected.

Keywords:

Safety Culture, Nuclear Power Plants, Basic Assumptions, Safety Performance

Funding:

This work was supported by the German Federal Ministry for Economic Affairs and Energy [grant number 1501340].

1. Introduction

The concept of safety culture is a worldwide accepted and popular perspective on optimizing collective human behavior in nuclear power plants. Developed and promoted shortly after the nuclear accident in Chernobyl (INSAG-4, 1991), safety culture has become a prominent vehicle in explaining and evaluating human performance in high-reliability settings. International agencies and regulators, as well as operating companies, develop guidelines, policies, assessment methods, and interventions aimed at the optimization of culture, and therefore safety. Stressing a systemic view of safety, the concept involves individual, organizational, and inter-organizational factors influencing the availability and safety of nuclear power plants. Recently, this perspective has been highlighted by an International Atomic Energy Agency (IAEA) report (2014) of an international expert meeting in response to the accident at the Fukushima Daiichi nuclear power plant. Reflecting on the causation of such a large-scale accident, the report promotes a holistic view of safety culture, claiming that "the interaction of human, organizational and technical factors across all stakeholder organizations and between different levels inside each organization must be evaluated and understood for each phase of the nuclear facility life cycle"(p. 30).

In contrast, scientific debate on safety culture reveals a critical, somewhat alarming picture of the concept and its application in the industry. For instance, Rollenhagen (2010) identifies a misuse of the concept in hiding shortcomings of technological design. Hopkins (2002) claims practitioners to see culture primarily as a matter of individual mindsets. As a consequence, blame is attributed to individual worker attitudes in causing accidents instead of focusing on latent work conditions and their contributing factors. Accordingly, Reiman and Rollenhagen (2014) come to the conclusion that safety culture is not sufficiently integrated into classical engineering principles and concepts. They disqualify the concept from its assumed system-oriented perspective on safety.

These conflicting foci on safety culture are often drawn from a conceptual confusion surrounding definitions and operationalization of safety culture (e.g., Hale, 2000; Guldenmund, 2000; 2007; Silbey, 2009; Myers, Nyce and Dekker, 2014): How deep should one dig to assess culture? Is safety culture quantifiable? And who is responsible if something has gone wrong? The answers to those questions allow for a huge diversity in interpretation due to the holistic nature of the concept. Thus, dealing with safety culture is a double-edged endeavor. On the one hand, it is one of the most acknowledged approaches in strengthening human performance in complex technological systems by capturing a broad range of factors, especially in nuclear power plants. On the other hand, the danger of misuse and disappointment is still present when bringing a scientifically vague concept into practice.

This article describes the results of a research project aimed at the development, application, and evaluation of a safety culture assessment method in nuclear power plants. It aims at contributing to a more common and positive understanding of the concept in bridging practical and conceptual challenges. Moreover, we wish to share our research experiences in developing a method tackling deeper levels of safety culture in the field, i.e., examining basic assumptions that may drive safety-oriented behavior in nuclear power plants. In the following, the applied background and methodological challenges of assessing safety culture are illustrated by focusing on Ed Schein's model of organizational culture (1985; 1990). Next, the development of the method, its application, and evaluation in a reference nuclear power plant (Study 1), and its validation in a second plant (Study 2) is presented. Finally, insights and limitations of the developed approach are discussed.

2. Challenges of assessing safety culture in the nuclear power industry context

Since the nuclear accident at the Chernobyl plant, a vast number of initiatives for optimizing safety culture have been introduced within the nuclear power industry. Starting with the INSAG-4 publication (1991), the concept mainly serves as an explanatory model, assumed to capture the opaque interplay between primarily non-technical parts of an organization as accident-contributing factors. Subsequent efforts put the concept into practice by focusing on the definition of safety culture and elaborating its facets. These initiatives were converted into guidelines (e.g., INSAG-15, 2002), policy statements (e.g., NRC-2010-0282) or regulations (e.g., KTA 1402). At the same time, safety culture has become a prominent issue in peer consultations at national (e.g., VGB, 2013) and international levels (e.g., SCART-Guidelines, 2008).

One crucial assumption underlying these initiatives is that there is an ideal safety culture, which has been extensively elaborated in several frameworks (e.g., INSAG-4, 1991; SCART-Guidelines, 2008; IAEA, 2009). These frameworks support the assessment of the status quo of a given plant and allow the derivation of strategies for improvement. However, interpreting safety culture as a normative concept stands in stark contrast to the scientific understanding of the concept (Schein, 2014). Going back to Edgar Schein's (1990) initial (still among the most popular academic) definition, organizational culture represents a set of shared basic assumptions learned by a group to cope with its problems of external adaptation and internal integration that has worked well enough to be considered as valid (p. 111). Here, culture is understood as a dynamic concept developed from the learning experiences of its members. Contrasting an organization's structures, systems, and members' behavior against general templates of an ideal safety culture might underweight the unique learning experiences of a given organization (Wilpert &

Schöbel, 2007). A similar argument has been put forward by Reiman & Rollenhagen (2014) regarding the use of questionnaires for assessing safety culture. They state that confronting organizational members with preselected dimensions does "not necessarily uncover the issues that the personnel neglect to consider or that they consider insignificant for safety" (Reimann & Rollenhagen, 2014, p. 11; see also: Guldenmund, 2007). From this it follows that one may conclude that an important requirement for assessing safety culture is a focus on the uniqueness of a given culture with regard to the learning experiences of its members.

Another challenge for tackling safety culture is the question of what elements of culture one should focus on. Whereas first definitions of the concept have remained relatively vague (e.g., "characteristics and attitudes in organization and individuals", INSAG-4, 1991), subsequently developed nuclear power plant-specific definitions clearly link safety culture to human behavior and corresponding psychological variables; e.g., "amalgamation of values, standards, morals and norms of acceptable behavior" (IAEA, 1998, p. 3) or "the core values and behaviors resulting from a collective commitment by leaders and individuals to emphasize safety over competing goals to ensure protection of people and the environment" (NRC, 2011, p. 34773). These understandings converge to a more scientific definition of the concept, stressing the importance of deeper or "intangible" levels of culture such as internalized values or norms. At the same time, they leave open room for interpretation with regard to the elements of safety culture. In applied nuclear power contexts, a behavior-focused understanding of culture is favored, because behavior "is what people do and say, and it is a mean to an end — the results. It is observable and measurable" (IAEA, 2013, p. 4). Safety culture is conceptualized as "instrumental in achieving alignment within the facility and in promoting the desired behaviors for overall performance improvement" (IAEA, 2013, p.17). Less attention is given to deeper levels or the so-called core of culture. Going back Schein's (1990; 2004) definition, safety performance is driven by basic assumptions, which are taken for granted and the ''deepest, often unconscious part of a group'' (Schein, 2004, p. 14). Thus, tackling "symptoms" of culture by focusing solely on (desired) behavior clearly risks ignoring and underrating cultural dynamics (Schein, 1990; Hale, 2000). Bringing safety culture into practice is therefore challenged by the necessity to tackle those deeper levels of culture and to demonstrate their potential sequels for daily operations. In the following, we will discuss conceptualizations of deeper levels of culture and potential ways of gathering insights into these levels.

3. Deeper levels of safety culture

As stated above, Schein (1985, 1990, 2004) conceptualizes the core of an organizational culture as shared basic assumptions, which are implicitly taken for granted and not challenged by members of a given culture. These assumptions define valid ways of perceiving and reacting to problems in organizations and develop along six dimensions, i.e., the *nature of reality and truth*, *time*, *space*, *human nature*, *human activity* and *human relationships*. They manifest themselves in *artifacts* and *espoused values.* Whereas artifacts describe observable elements in an organization (e.g., behavior, documents, systems), espoused values are their verbal counterpart derivable from asking members questions "about the things the organization values" (Schein, 1999, p. 17). However, neither artifacts nor espoused values are easily attributable to deeper levels of culture. Both are subject to uncertainty with regard to the meanings that members of a given culture attach to them, especially when inconsistencies between manifestations on the level of artifacts and espoused values exist (e.g., when management communicates the importance of rule compliance but, on the other hand, accepts rule-violations). According to Schein (1999), these inconsistencies provide clues to important deeper levels of thought and perception (i.e., to basic assumptions), which are assumed to drive overt behavior. Importantly, these inconsistencies often signal problems stemming from dysfunctional cultural dynamics, which clearly demands tackling those deeper levels when assessing culture.

However, the majority of existing methods seldom adopt this perspective for going beyond the assessment of artifacts and espoused values. For instance, questionnaires, which naturally address only both upper levels, are still a very popular method of diagnosing safety culture in nuclear power plants (e.g., Lee & Harrison, 2000; Garcia-Herrero et al., 2013). Yet, uncovering deeper layers of culture is a difficult undertaking, which comprises the application of qualitative or anthropological methods mainly via fieldwork and participant observation (or using an iterative clinical approach, Schein, 1990). There are only a few papers on safety culture that have adopted such approaches. One of them is a study by Brooks (2008) using an ethnographic research method to describe the organizational culture of a small manufacturing business. He conducted open-ended interviews and observations to collect observational data about artifacts and espoused values. Based on Schein's (1985) taxonomy of cultural assumptions, he moved through the data collection in an iterative process ("Hermeneutical Canons of Interpretation") to better understand the meaning and therefore the assumptions defined by the members of culture. Guldenmund (2010a) conducted a case study in a service company combining qualitative (e.g., focus groups, interviews) and quantitative (e.g., questionnaires) methods for the diagnosis and description of safety culture. The qualitatively obtained data provide the context within which the quantitative data are interpreted. In an afterword to this study, Guldenmund (2010a) lists a preliminary set of deciphered basic assumptions structured according to the six dimensions of Schein (1985).

Based on this, it becomes evident that the proper application of both methods, i.e., qualitative and quantitative, might advance the revelation of deeper cultural levels. However, thus far, almost no common knowledge and experiences are available about ways to identify relevant basic assumptions of members of an organization.

In the following, we present a multi-method approach of safety culture assessment based on Schein's culture model. The method provides the possibility to identify relevant basic assumptions that are assumed to drive the safety-oriented behavior in the organization. The development and application of this approach was realized in two nuclear power plants (NPPs). In Study 1 we describe the development of the approach and its first application in an NPP, which we will refer to as "reference plant" in the following. Study 2 presents the results of cross-validating the approach in a second NPP.

4. Study 1: Development and feasibility test of the assessment approach

The assessment approach is based on Schein´s (1985, 1992, 2004) model of organizational culture. It provides methods for assessing culture on all three different levels, i.e., artifacts, espoused values, and basic underlying assumptions. Following Schein's advice, a focus on inconsistencies between artifacts and espoused values was chosen as the main strategy for compiling three specific assessment steps focusing on deeper cultural levels. The assessment starts with an artifacts analysis (step 1), which is followed by a gap analysis, i.e., an analysis that contrasts assessed artifacts against espoused values (step 2), and is completed by an assumption analysis (step, 3). The basic structure of the approach is presented in Figure 1. Because manifestations on those three levels represent interacting elements of culture, the three assessment steps are closely interlinked. Specifically, the results from one analysis step provide the starting material for the next — deeper — analysis step. As a consequence, the approach is not assumed to provide a complete picture of a given safety culture. Instead, the intention was to develop an assessment approach that allows for gaining insights into important elements of the culture under study (Schein, 1992, p. 28). In the following, we briefly introduce the three developed methods in more detail.

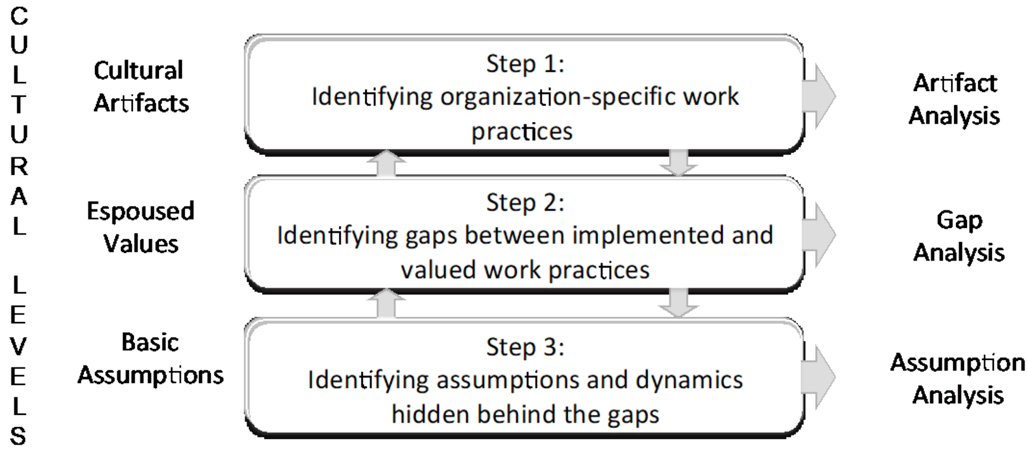

Figure 1. Overview of the assessment steps of the developed safety culture assessment approach

4.1 Method

4.1.1 Development and description of the artifact analysis

The goal of the artifact analysis is to identify those plant-specific artifacts that are relevant to plant members in establishing safe and reliable plant performance. Artifacts represent a broad category comprising everything that "you see, hear and feel as you hang around" (Schein, 1999, p. 15). They reflect learning experiences of organizational members in responding to problems of external adaptation (e.g., how to deal with requirements stemming from regulators) and internal integration (e.g., how to establish safety-directed leadership) with regard to the safe operation of the plant.

In developing a method for the artifact analysis that may feasibly be applied in any NPP, we focused on safety behavior of the members of an NPP by collecting basic safety-relevant work practices. These are specific modes of behavior that members of an NPP apply in their daily work, which are accessible by observation. It can be assumed that such practices represent "...particular ways of conducting organizational functions that have evolved over time" and that "practices reflect the shared knowledge and competence of the organization" (Kostova, 1999, p. 309; see also Wilderom, van den Berg & Wiersma, 2012). Based on this understanding, we considered work practices as patterns of recurrent behavior that reflect manifestations of a given organizational culture (Hofstede, 2001). Furthermore, work practices represent a well-known concept in the nuclear community (e.g., in terms of best practices in IAEA-TECDOC-743, 1994; IAEA-TECDOC-1358, 2003). Thus, focusing on these practices was assumed to be well accepted in this particular industrial domain.

For the artifact analysis, a list of 64 work practices was compiled (see Table 1 for some work practice examples). This compilation is based on (1) an extensive literature search, (2) participant observation by one member of the research team (over a period of one month), and (3) the conductance of 12 semi-structured interviews with employees and supervisors from different plant departments. The collected work practices were then counterchecked and rephrased by two members of the nuclear power plant guaranteeing their comprehensibility. Furthermore, they were structured according to a safety management framework. This framework was developed in accordance with existing safety management conceptualizations (Diaz-Cabrera et al., 2007; Guldenmund, 2007; Haber et al., 1995; Hale, R., Kirwan & Guldenmund, 1999; INSAG-13, 1999; Kennedy & Kirwan, 1998; Weil & Apostolakis, 2001) and allows assigning all collected practices to four management domains, i.e., risk management, task management, management of technical resources, and management of human resources. These four domains were further differentiated into ten facets (see Table 1 for an overview).

Table 1. Safety management framework with corresponding facets and examples of work practices

Management domain	Facet	Practice (example)
Risk management	Monitoring (identification and diagnosis of process deviations)	Employees report safety deficiencies.
	Changing (evaluation and intervention regarding deviations)	Lessons learned from event analyses were communicated to the staff.
Task management	Accomplishing (work tasks)	Employees notice improvement opportunities and actively seek advancements.
	Organizing (structuring responsibilities and processes)	Responsibilities, functions and tasks are clearly conferred.
	Communicating (coordination of information flow)	Employees communicate their experiences to their colleagues.
Human resource management	Planning (requirements planning and competence preservation)	Staff planning considers and counteracts a potential loss of competencies.
	Qualifying (promoting and developing)	Training is evaluated on a regular basis.
	Leading (motivating and controlling personnel)	Supervisors discuss with their staff how deficiencies may be corrected.
Technical resource management	Designing (design of technical components)	Technical reports from suppliers are critically evaluated.
	Maintaining (maintenance and inspections)	Maintenance intervals for inspections of technical components are met.

As target groups for the artifact analysis, 24 middle managers from the reference plant were asked to rate each of the practices with respect to its implementation in daily work. Because middle managers are assumed to receive requirements from top management and pass them to shop-floor personnel, they were supposed to be knowledgeable experts on daily operations in the plant (e.g., Balogun & Johnson, 2004; Zohar & Luria, 2005). The professional background of the middle managers covered all main departments (e.g., production department, maintenance department) directly involved in the safe operation of the plant. Middle managers' mean length of employment was 19.68 (SD=8.5) years.

The ratings were collected for each single work practice by means of seven-stage *Likert*-scales ranging from 1 (= practice not applied) to 7 (= practice fully applied). Furthermore, an additional confidence judgment (i.e., *how confident are you in your rating*?) ranging from 1 (= not at all) to 7 (= completely), and a rating of familiarity with each practice domain (according to the safety management framework facets) was requested, because different work domains of the managers should be taken into account as potentially biasing their ratings. Moreover, an open question was formulated to collect additional relevant work practices that are not featured by the framework.

In order to identify artifacts (work practices) that were sufficiently relevant for the reference plant (i.e., identifying those practices paying tribute to the specific learning experiences of a given culture) the ratings were first screened according to two exclusion criteria. First, only ratings of implemented work practices that could be confidently assessed by the target group were considered (i.e., only ratings that were accompanied by confidence ratings of 4 and above). Secondly, only work practices whose mean implementation ratings were on average 3.5 or higher (with a standard deviation below 1) were considered as relevant for the plant and included in the next step of assessment, i.e., the gap analysis.

For evaluation purposes, a short questionnaire was developed, which was used to evaluate the artifact analysis method by members of the target group. Three items assessed the face validity of the selected practices with regard to their relatedness to safety culture (e.g., "*The practices are an important element of safety culture*"). Four items were developed to rate the usability of the rating method (e.g., "*The formulation of the practices is comprehensible*"). Each of the seven items had be rated on a *Likert*-scale ranging from 1 (= strongly disagree) to 7 (= strongly agree).

4.1.2 Development and description of the Gap Analysis

The second step of our assessment approach aims at contrasting artifacts against the espoused values of organizational members. This approach is known as *gap analysis* (e.g., Boglarsky & Kwantes, 2004; Kilmann & Saxton, 2007). Whereas the artifact analysis identified actually implemented plant-specific safety practices, the gap analysis should assess the organizational members` conception of ideal safety behavior. This is based on the assumption that shared espoused values are reflected in what members of an organization perceive as ideal behavior, ensuring proper process safety and productivity. To identify possible gaps between artifacts (observable behavior) and espoused values (ideal behavior), a questionnaire was designed and distributed to all staff members of the plant. Thirty eight percent of all members of the reference plant, which stemmed from all major departments of the plant, filled out the gap questionnaire. From this, it emerged that 60% of the respondents had worked for more than ten years in the plant.

The gap questionnaire contained the work practices that had been identified as relevant by the artifact analysis. First, each of these work practices had to be rated according to their perceived *actual* implementation in the plant (by means of seven-stage *Likert*-scales ranging from 1 = *practice not applied* to 7 = *practice fully applied*). After that, each practice had to be rated with regard to what extent the practice should ideally be implemented in the plant (by means of seven-stage *Likert*-scales ranging from 1 = *practice should not ideally be applied* to 7 = *practice should ideally be applied*). Both questions allow us to directly contrast the "actual" (level of artifacts) and "ideal" ratings (level of espoused values) for each practice.

With the gap analysis, two goals are pursued: The gap analysis should allow identification of (1) best practices of a given plant culture and (2) inconsistencies between the perceived actual and ideal implementation. With regard to the identification of best practices it was assumed that these give hints to those basic assumptions that are consistently represented in shared values and corresponding working behavior of plant members. To identify best practices, we focused on work practices with high ratings on both their actual and ideal implementation. For this purpose, the practices were rank ordered according to practices with the highest mean ratings of actual and ideal implementation. Next, practices with the highest ranks on both dimensions were selected as best practices.

As a second criterion in case of indifference, practices with the lowest standard deviations accounting for the sharedness of perceived best practices were marked as best practice.

On the other hand, the gap analysis should identify inconsistencies between the perceived actual and ideal implementation of the practices. According to Schein (2004), these inconsistencies provide clues to important basic assumptions, which drive behavior that is not in accordance with the espoused values of organizational members. It is of particular importance to know and understand these basic assumptions because although they are not expressed in the officially espoused values they are nevertheless important drivers of everyday working behavior.

To identify inconsistencies (i.e., cultural gaps) between the perceived actual and ideal implementation of a practice, we defined a gap between actual and ideal implementation of a work practice, reflected in individual ratings, as meaningful when there was a difference of two scale points between the actual and ideal rating of the same work practice (cf. Kilman & Saxton, 2007 for a similar approach). Although this difference may occur in both directions, we specifically focused on practices that were (at least two scale points) higher rated according to their ideal than their actual implementation. In a next step, the frequencies of individual gaps per practice were counted, and the percentage of gaps in relation to the whole sample was calculated. Furthermore, an identified gap was considered as meaningful and relevant for the overall organization when the individual ratings of at least 40% of all respondents perceived this gap.

In addition, a second questionnaire was introduced for evaluating the gap analysis. It was the same questionnaire as used for evaluating the artifact analysis except that we asked additionally for the ease of evaluating the practices according to their ideal implementation.

<u>4.1.3 Development and description of the Basic Assumptions analysis</u>

The final assessment step aims at analyzing the identified gap-practices in more detail to understand and yield the underlying cultural dynamics and corresponding basic assumptions that have led to and maintain these gaps. The gap- and best practices identified in the former assessment step (i.e., in the gap analysis) provide the starting material for this assessment step. Fostering a system view of analyzing the dynamics behind the identified gaps, two methods were developed, both of which are based on the system dynamics approach (e.g., Marais, Saleh & Leveson, 2006): a cultural dynamic interview and a consecutive group workshop.

Cultural dynamic interviews

Motivated insiders from different departments of the plant were selected as the target group for the interviews and workshops. These are members of the plant who had participated in former assessment steps, had profound knowledge of the plant's daily operations, and signaled their willingness to cooperate in developing the assessment approach. In total, 15 cultural dynamic interviews were conducted. Each interview lasted approximately one hour.

The cultural dynamic interview started with an introduction to the "language" of the system dynamics approach. More specifically, the interviewees were inducted into three elements: reinforcing loops, balancing loops, and delays. According to Marais et al. (2006) a reinforcing loop displays a structure "that feeds on itself to produce growth or decline", whereas a balancing loop is a structure "that attempts to move a current state to a desired or reference state through some action" (p. 566). A delay is used to model the time that elapses between cause and effect.

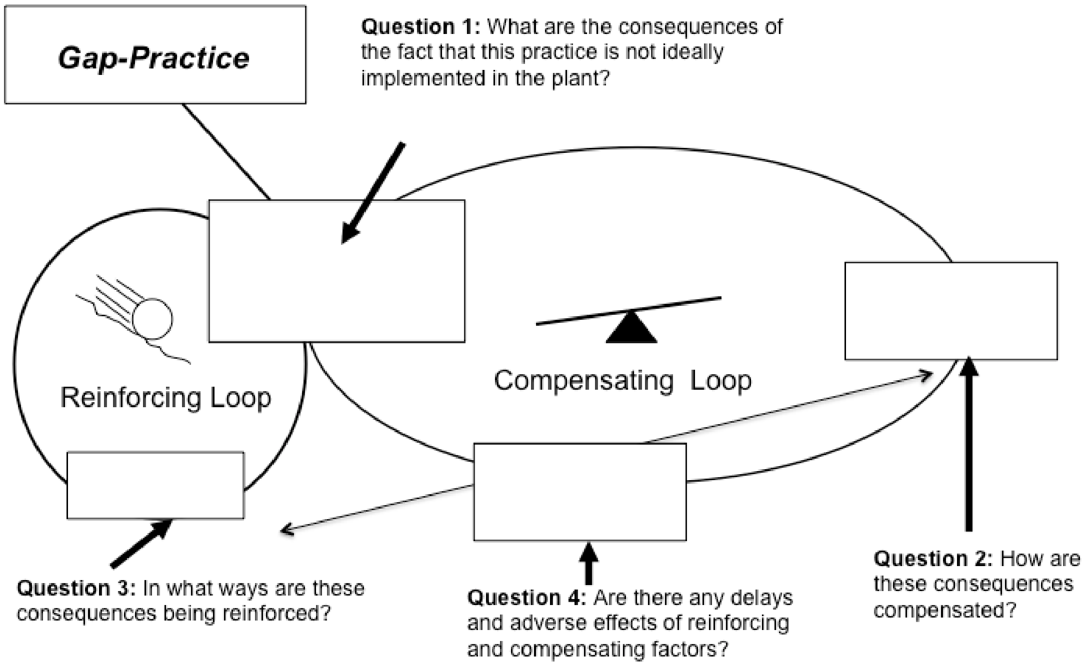

Figure 2. An image template for structuring and guiding the cultural dynamic interviews

The idea was to familiarize the interviewees with these elements and enable them to illustrate potential cultural dynamics underlying the gaps. Moreover, the system dynamic language forced the interviewees to adopt a system view of cultural dynamics within their plant. After the introduction and familiarization phase, four questions were asked (depicted in an image template shown in Figure 2):

1. What are the consequences of the fact that this practice is not ideally implemented in the plant?
2. How are these consequences compensated?
3. In what ways are these consequences being reinforced?
4. Are there any delays and adverse effects of reinforcing and compensating factors?

The interviewees were allowed to sketch or verbally communicate their answers. Based on this, the obtained data (i.e., consequences of the gaps, reinforcing and balancing loops) were aggregated by the project team according to the frequency of mentioning specific elements in so-called "cultural dynamics images"(CDIs), which were discussed and further elaborated in group workshops.

Group workshops

Workshop participants were recruited from the same subject pool as for the cultural interviews, representing all major departments of the NPP. Two workshops with 13 participants took place. Both workshops lasted one day.

The group workshops aimed at (1) validating the CDIs, (2) identifying basic assumptions triggering these dynamics, and (3) finding solutions to problems depicted by the CDIs. Therefore, a workshop-concept was developed containing four modules. The workshops started with an introduction of Ed Schein's culture model and the system dynamics approach to the workshop participants. After briefly describing the previous assessment process and corresponding results, the aggregated CDIs based on the cultural dynamic interviews were presented. The participants' task was to validate the CDIs by discussing them in small groups and, if necessary, rearranging and adding elements. After that, participants assigned potential basic assumptions to (1) compensating and reinforcing loops of the CDIs and (2) the best practices (identified by means of the gap analysis). Whereas workshop participants had no problems in assigning assumptions to the best practices (i.e., no gap-practices), a former pilot study of the workshop with members of the project team revealed that eliciting assumptions for a complete CDI (i.e., the gap-practices) was a quite difficult task. Therefore, workshop participants developed and assigned assumptions to single reinforcing and compensating loops in the CDIs. In the last module, the CDIs and the elicited assumptions were rated according to their dominance in the given culture (up to three assumptions from the CDIs could be named by each participant) and possible interventions to improve the safety culture (i.e., to close the gaps) were generated.

At the end of the workshop, participants filled out an evaluation questionnaire containing questions regarding the validity of the CDIs (7 Items) and the evaluation of the workshop method (6 Items). Again, each of the items was rated on a *Likert*-scale ranging from 1 (= strongly disagree) to 7 (= strongly agree).

4.2 Results

The presentation of application results of the approach and its evaluation in the reference plant will be structured in two parts. In the first part, we will provide a general overview of the application and the evaluation of our approach by the members of the reference plant who were involved in the pilot application on the three different stages of the safety culture assessment. In the second part, an exemplary case of the full assessment across all steps of analysis is presented illustrating the feasibility and potential of the approach as well as providing an insight into what a typical result looks like. A more complete pattern of results of the pilot application cannot be provided due to concerns of the management of the reference plants concerning a full publication of all results.

4.2.1 General Results:

Evaluation of the approach. Mean evaluations of different aspects of the assessment approach for all three levels are provided in Table 2, together with the results of *t*-tests contrasting each mean with the (neutral) midpoint of the evaluation scales. It becomes evident that, on a descriptive level, all three stages of assessment and the methods used were generally positively evaluated with respect to their understandability, ease of use, and the transparency of the assessment. The results of Bonferroni corrected *t*-tests revealed significant positive differences of all mean ratings from the midpoint of the different scales, with only one exception (i.e., the ratings about the perceived goals of the workshop).

Table 2: Evaluation results of the three assessment steps in the reference plant with corresponding results of *t*-tests (testing for deviations from the scale midpoint of 3.5)

Items				
Artifact analysis	**M (SD)**	***T***	***Df***	***p***
Comprehensible phrasing of the practices	5.33 (0.96)	9.33	23	.0001
Temporal adequacy	5.63 (1.40)	7.39	23	.0001
Clarity of the instructions	5.96 (0.99)	12.1	23	.0001
Effortlessness of the method	6.00 (0.89)	13.8	23	.0001
Ease of evaluating the implemented practices	4.79 (1.10)	5.37	23	.0001
Relatedness of practices to safety culture	6.25 (0.89)	15.0	23	.0001
Relevance of the practices for safety	5.71 (0.13)	17.3	23	.0001
Gap analysis	***M (SD)***	***T***	***Df***	***p***
Comprehensibility of the practices	5.08 (1.20)	14.8	125	.0001
Temporal adequacy	4.23 (1.89)	4.32	126	.0001
Clarity of the instructions	5.47 (1.48)	14.9	126	.0001
Effortlessness of the method	5.14 (1.59)	11.6	126	.0001
Ease of evaluating ideal implementation of the practices	4.69 (1.42)	9.43	125	.0001
Ease of evaluating actual implementation of the practices	4.58 (1.41)	8.54	123	.0001
Relatedness of practices to safety culture	6.05 (0.96)	29.8	124	.0001
Relevance of the practices for safety	5.61 (1.10)	20.8	118	.0001
Assumption analysis	***M (SD)***	***T***	***Df***	***p***
Practicability of the system-dynamic approach	4.85 (0.80)	6.06	12	.0001
CDIs as description of the safety culture	4.77 (1.09)	4.19	12	.0001
Transparency of the CDIs	5.08 (0.64)	8.87	12	.0001
Relevance of the CDIs	4.85 (0.81)	6.06	12	.0001
Relatedness of the CDIs to safety culture	5.00 (1.00)	5.41	12	.0001
Importance of CDIs for change	6.00 (0.71)	12.7	12	.0001
Implications for change	6.08 (0.86)	10.8	12	.0001
Structure of the workshop	5.23 (0.93)	6.73	12	.0001
Goals of the workshop	3.92 (1.50)	1.01	12	.3290
Temporal adequacy	5.38 (1.20)	5.69	12	.0001
Clarity of the instructions (validating)	4.92 (1.32)	3.89	12	.0001
Clarity of instructions (weighting)	5.77 (0.43)	18.7	12	.0001
Fun at the workshop	6.00 (1.08)	8.35	12	.0001

Bonferroni corrected α –levels: α = .007 (artifact analysis), α = .006 (gap analysis), and α = .004 (assumption analysis)

Artifact Analysis. In total, 64 work practices were rated by the sample of N=24 middle managers. Applying both exclusion criteria, a total of 54 practices were selected as being relevant and kept for the next step of assessment. No additional practices were named by middle managers of the reference plant.

Gap Analysis. Applying the 40% criteria (see above), eight practices were identified as gap-practices. Moreover, three practices were identified as best practices based on their high correspondence of rated actual and ideal implementation.

Assumption Analysis. All identified eight gap-practices were fed back to plant members and analyzed by means of cultural interviews trying to uncover the dynamics lying behind the gaps. The cultural interviews revealed that interviewees had hardly any problems applying the system dynamics language to describe potential consequences of a given cultural gap with corresponding reinforcing and compensating factors. Most of them favored sketching the cultural dynamics depiction rather than only naming them. As a result, eight CDIs were developed.

The subsequent workshops with selected staff members of the plant generally confirmed the eight CDIs. Most of the improvements elicited from the workshops included a better wording and elaborating the CDIs. Based on this, a total of 21 basic assumptions underlying the three best practices and the eight CDIs (with regard to reinforcing and compensating loops) were identified and defined. Final ratings of the dominance of single assumptions in the culture reduced the number of assumptions to six.

4.2.2 Case Study: Cultural dynamics and assumptions underlying reporting behavior

In the following, a case from the reference plant is presented, which aims at illustrating the potential of the assessment approach in focusing on deeper basic assumption levels of safety culture. This case relates to a general and longstanding problem in high-reliability organizations that have implemented reporting systems for incidents and near misses and try to motivate employees to report. One major recurrent problem here refers to the avoidance of blame, which is seen as a crucial factor discouraging employees from reporting (e.g., Reason, 1998).

Based on the results of the artifact analysis, the practice "*Employees are not blamed for reporting errors*" was rated as highly relevant for the reference plant. However, the gap analysis revealed that 41% of plant members perceive a gap between the actual and ideal implementation of this practice. Results from subsequent cultural dynamics interviews showed that there was a high consensus among the interviewees with regard to the consequences arising from this gap. The majority of the interviewees identified a declining frequency of error reports as a main consequence. Additionally, they pointed out that a higher frequency of undetected errors in the system, a strong demotivation in taking over extra tasks *("When I do not do anything, I will not make any error")*, and a higher occurrence of errors in the long run might also result from the perceived gap. As compensating mechanisms, the interviewees relied on the design and functioning of the reporting system itself. They focused on elements of its formalization (assurance of anonymity and exemption from sanctioning) and the behavior of the employees responsible for running the system (in-depth analysis with a focus on contributing to work conditions and not focusing on individuals committing errors). With regard to the reinforcing mechanisms, the cultural dynamics interviews did not reveal a clear picture. The interviewees mentioned several factors such as supervising behavior, the difficulty to assure anonymity and high workload in general. However, it became evident that a more subtle type of blame from colleagues and managers is a decisive factor, which is not directly related to the perception and functioning of the reporting system.

In the consecutive workshops, the results of the interviews were fed back by means of a CDI with one reinforcing loop solely displaying the factor "subtle blame". Here, workshop participants highlighted the role of the daily morning meetings ("Morgenrunde"), where plant top management discusses occurrences of the previous day and current work tasks. A crucial issue in these meetings is the discussion of errors that have led to system disturbances and were not reported. In contrast to the standards of the reporting system, single individuals and departments are blamed in those meetings, emotionally triggered by the fact that these errors were not reported (although it is open to what extent they were reportable or could be attributed to single individuals). Moreover, the reporting system, with its focus on organizational learning on the one hand and the daily morning meetings with a focus on daily task management on the other, was perceived as only indirectly related or two separate systems with different goals. Consequently, plant management underestimated the potential effects of blaming in one system on another system.

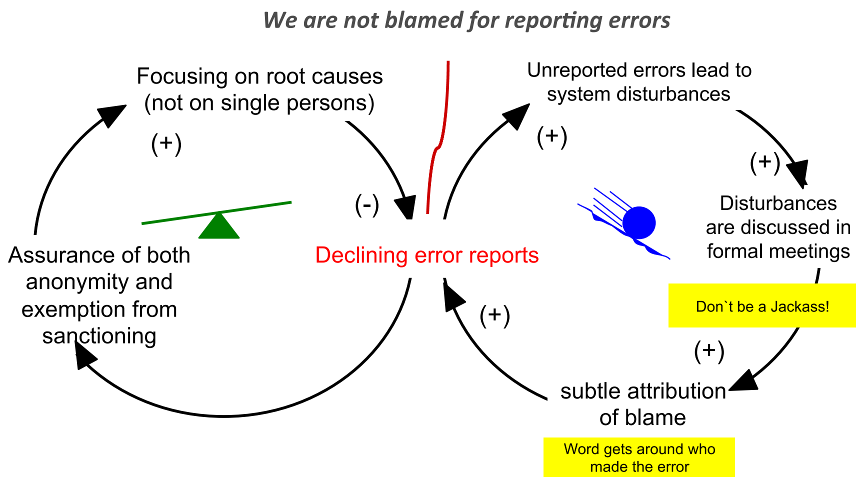

Figure 2. CDI with assigned assumptions (yellow) for the practice "We are not blamed for reporting errors" depicting (1) the consequence of its not ideal implementation, (2) a reinforcing loop, and (3) a compensating loop

After elaborating the reinforcing "subtle blame" loop in more detail, workshop participants assigned underlying assumptions to the elements (i.e., reinforcing and compensating loops) of the CDI. In Figure 2 we present the final CDI together with the assigned basic assumptions after its elaboration and validation in two workshops. In developing suggestions to improve the dysfunctional cultural dynamics displayed by the gap (i.e., to close the gap), workshop participants primarily focus on changing the perception and running of both systems as separate units. One implication was to put a new item on the agenda of the daily morning meeting, where new insights and lessons learned from the reporting system (e.g., stats, specific cases) are discussed and communicated. This implication was quite different from previous efforts to optimize reporting behavior in the plant, because these mainly targeted solely plant members' perception of the reporting system and its underlying conditions.

5. Study 2: Validating the approach in a second nuclear power plant

After developing and applying the assessment approach in the reference plant, a second assessment was conducted in another nuclear power plant (i.e., the validation plant). This allowed us to test our assessment approach (with slight modifications based on the feedback and the results of the assessment in the reference plant) in a second plant. Moreover, the approach could be validated with a second sample that was not involved in the development of the method.

5.1 Method

The assessment in the validation plant passed through the same assessment steps as in the reference plant i.e., the methods and corresponding exclusion criteria were the same. Only minor modifications were made. They are reported in the subsequent sections. All three assessment steps were evaluated by the same questionnaires as in the reference plant.

5.1.1 Artifact analysis

The existing list of 64 safety-relevant work practices served as a starting point for the artifact analysis. However, after reflecting the application of the artifact analysis and integrating the feedback from the middle managers in the reference plant, the project team decided to modify the artifact analysis slightly. First, the rating method was applied in a group context from the very beginning. After rating each of the 64 practices, now a group of middle managers was instructed to name and discuss potential new and shared practices that were not part of the initial framework. Secondly, because some practices were not fully implemented but strongly supported by the managers, each practice (besides the actual implementation ratings and corresponding confidence and familiarity ratings) also had to be rated regarding to what extent plant members discuss its implementation. Accordingly, this allows plant management to evaluate the degree of implementation of practices that were actually promoted by them.

As target groups for this artifact analysis, 17 middle managers from the validation plant filled out the rating questionnaire. Middle managers' mean length of employment was 22.76 (SD=9.4) years.

5.1.2 Gap analysis

The gap analysis was conducted in the same way as in the reference plant. The relevant practices (identified by the artifact analysis) were fed back to members of the plant by means of a questionnaire. Forty four percent of all members of the validation plant completed the gap questionnaire; i.e., they rated the actual and ideal implementation of the practices. Sixty eight percent of the respondents had worked for more than 10 years in the plant.

5.1.3 Assumption analysis

The assumption analysis was conducted in the same way as in the reference plant. Twelve interviews with motivated insiders (who had participated and signaled their willingness to participate in former assessment steps) were conducted. Each interview lasted approximately one hour. In a second step, one day-long workshop was conducted with seven participants in the validation plant.

5.2 Results

In the following, we present the results obtained from the assessment and the evaluation questionnaires in the validation plant. Again, plant management had strong concerns regarding the confidentiality of the results. Thus, the results section is structured into two parts: First, a general overview of technical aspects of the assessment and its evaluation is given. Second, we depict some insights and experiences derived from comparing the assessment results of both plants.

5.2.1 General overview: Results from assessing safety culture in the validation plant

Evaluation of the approach. Similar to study 1, the three stages of the assessment in the validation plant were accompanied by three evaluation questionnaires. Table 3 displays the evaluation results of the validation plant. It becomes evident that the assessment approach revealed adequate validity; the usability of the developed methods is also evaluated as adequate, except for the description of safety culture by the CDIs, which was not judged significantly higher as the scale midpoint.

Table 3: Evaluation results of the three assessment steps in the validation plant with corresponding results of *t*-tests (testing for deviations from the scale midpoint of 3.5)

Items				
Artifact analysis	***M (SD)***	***t***	***df***	***p***
Comprehensible phrasing of the practices	5.82 (0.73)	13.2	16	.0001
Temporal adequacy	6.06 (0.83)	12.8	16	.0001
Clarity of the instructions	6.24 (0.93)	12.5	16	.0001
Effortlessness of the method	5.88 (1.36)	7.20	16	.0001
Ease of evaluating the implemented practices	5.06 (1.35)	4.78	16	.0001
Relatedness of practices to safety culture	6.35 (0.61)	19.4	16	.0001
Relevance of the practices for safety	5.69 (0.70)	12.4	15	.0001
Gap analysis	***M (SD)***	***t***	***df***	***p***
Comprehensibility of the practices	5.47 (1.15)	26.8	243	.0001
Temporal adequacy	4.48 (1.85)	8.26	243	.0001
Clarity of the instructions	5.74 (1.42)	24.7	243	.0001
Effortlessness of the method	5.29 (1.64)	17.0	243	.0001
Ease of evaluating ideal implementation of the practices	4.96 (1.47)	15.5	243	.0001
Ease of evaluating actual implementation of the practices	5.04 (1.28)	18.9	243	.0001
Relatedness of practices to safety culture	6.13 (1.03)	39.7	243	.0001
Relevance of the practices for safety	5.80 (1.10)	32.7	243	.0001
Assumption analysis	***M (SD)***	***t***	***df***	***p***
Practicability of the system-dynamic approach	5.17 (0.75)	5.42	5	.003
CDIs as description of the safety culture	4.83 (0.75)	4.34	5	.007
Transparency of the CDIs	5.67 (0.52)	10.3	5	.000
Relevance of the CDIs	5.33 (0.52)	8.70	5	.000
Relatedness of the CDIs to safety culture	5.83 (0.75)	7.59	5	.001
Importance of CDIs for change	5.50 (0.55)	8.94	5	.000
Implications for change	6.50 (0.55)	13.4	5	.000
Structure of the workshop	5.83 (0.75)	7.59	5	.001
Goals of the workshop	5.83 (0.41)	14.0	5	.000
Temporal adequacy	5.83 (0.98)	5.81	5	.002
Clarity of the instructions (validating)	5.67 (1.03)	5.14	5	.004
Clarity of instructions (weighting)	5.83 (0.41)	14.0	5	.000
Fun at the workshop	5.50 (0.55)	8.94	5	.000

Bonferroni corrected α –levels: $\alpha = .007$ (artifact analysis), $\alpha = .006$ (gap analysis), and $\alpha = .004$ (assumption analysis)

Artifact analysis. By means of the modified version of the artifact analysis, 51 practices (from 64) were rated as relevant for plant performance. Eleven new practices that were not part of the initial list of practices were named by the middle managers.

Gap analysis. Forty four percent of members of the validation plant evaluated the selected 62 practices according to their actual and ideal implementation. The results showed that in the validation plant the mean percentage of individual gaps per practice was somewhat lower than in the reference plant (23% vs. 28.4%). Therefore, the threshold was set to 35% (of individual gaps per practice) in determining relevant gap-practices, resulting in three gap-practices for further analysis. Moreover, three best practices could be identified.

Assumption analysis. Based on the results of the gap analysis, out of the three gap-practices three CDIs for the validation plant were developed by means of the cultural interviews. In the ensuing workshop the three CDIs and three best practices were validated and 10 assumptions were assigned to them. The cultural dominance ratings led to a reduction of these assumptions to five.

5.2.2 Further assessment results: Comparing the results of both assessments

In the following, we summarize some results derived from comparing the assessment results of both plants. On the level of artifacts, the majority of rated work practices (49 out of 64 practices) were identified as relevant in both plants. Although we found no major differences in the implemented work practices in both plants, they differ with regard to the absolute ratings, i.e., the averaged absolute ratings of implemented work practices differed between both plants (5.8 vs. 5.1).

The results of the gap analysis revealed a similar picture. Although the absolute percentage of gaps found differed between both plants, the identified gap-practices and best practices, which were included for further analysis, do not differ at all. The three gap-practices identified in the validation plant were also part of the eight gap-practices that were identified as relevant for further (basic assumption-) analysis in the reference plant. And two out of three identified best practices were identical in both plants.

On the level of basic assumptions we found some consistent results between both plants with regard to the assumptions assigned to the best practices. Members of both plants identified assumptions that refer to the positive effects of formalization on system safety (e.g., *"If it is formalized, it works"*, or *"Conferring responsibilities improves safety"),* and were rated as dominant assumptions in both cultures. Thus, in nearly all CDIs of both plants a compensation loop was integrated, which rests on those assumptions driving overt behavior such as formalizing and further written specification of work processes. However, in most of the CDIs these compensation loops did not solely show intended positive effects, because higher degrees of formalization either result in (delayed) secondary effects such as a higher workload or stronger safeguarding behavior, or do not solve the problem at hand (as shown in Study 1, where the compensation loop did not directly target at the reinforcing dynamics).

From a methodological viewpoint, another interesting result refers to the same three gap-practices that were identified in both plants. Although results of the gap-questionnaires revealed a gap in the same practices in both plants, assumption analysis uncovered different cultural dynamics contributing to the gaps. For instance, the practice "staff planning considers and counteracts a potential loss of competencies" was identified (based on the frequency of individual gaps) as a gap-practice in both plants. However, assumption analysis revealed that in one plant the consequences of the gap (i.e., a loss of competencies) were reinforced by assigning extra tasks to new members instead of working them into the job of their predecessors. This reinforcing loop was based on the assumption "We help each other". In the other plant, the gap was reinforced by problems of hiring external experts based on the assumption that "expert knowledge cannot be easily transferred and has to be bought". What is crucial here is that based on the questionnaire results both plants seem to have the same problem. Without digging deeper cultural levels, developed interventions (e.g., optimizing the "vocational adjustment" process for new members) may fail (at least in one plant), because underlying cultural dynamics call for different solutions to — considered on a superficial level — the same problem.

6. Discussion

The goal of our research was to develop a safety culture assessment approach for nuclear power plants based on Schein's culture model. Specifically, the developed approach aimed at unfolding deeper levels of culture, and, at the same time, being applicable by practitioners and transparent in producing meaningful results. Therefore, a three-step assessment approach was developed, applied, and evaluated in two nuclear power plants. The evaluation results in both plants revealed that the developed approach has adequate validity with regard to the obtained results and their perceived linkage to the safety culture. The application of the method was predominantly rated as transparent and comprehensible. Moreover, it became evident that providing plant members with a language for describing complex cultural dynamics facilitates the derivation and discussion of basic assumptions and their influence on safety-related behavior. Besides the attainment of these intended goals, the application of our assessment approach also revealed secondary beneficial effects as stated by the top management of both plants. The results of the gap analysis were regarded as important feedback for plant management, because they allowed re-evaluating the implementation and the perceived value of specific practices that had been promoted by the top management in the past. Furthermore, management emphasized the importance of providing a "balanced" view of safety culture, i.e., feeding back positive (best practices) and dysfunctional effects (CDIs) on safety culture and associated basic assumptions, which seemed to be a necessary precondition to obtain acceptance of the approach and its results.

The developed approach clearly promotes a system view of safety culture. Therefore, it meets the criticism of several safety scientists (e.g., Reimann & Rollenhagen, 2014; Hopkins, 2002) disqualifying the concept from its assumed system-oriented view of safety. Due to guiding plant members in eliciting basic assumptions by means of the system dynamics approach (e.g., Marais et al., 2006), we were able to stimulate sense-making processes on a system level, where plant members automatically focused on management behavior (as prime driver of culture according to Schein, 2014). It became evident that one of the main barriers in adopting a system view of culture was missing capabilities by plant members to express and describe complex cultural dynamics. The system dynamics approach provides a simple-to-learn language (i.e., reinforcing and compensating loops) to describe complex processes and therefore allows communicating and resolving system issues.

On a theoretical level, the results of both our studies refer to a long-standing paradox in research on safety culture. According to Pidgeon (1998), culture acts simultaneously as a precondition for safe operations as well as an incubator for hazards or, in Silbey's (2009) words, "unusual events and accidents are generated by the same cognitive processes that enable the ordinary, routine interactions of daily life" (p. 357). The results of the assumption analysis in study 1 showed that the elicited assumptions "Word goes around who made the error" and "Don't be a jackass" may also have the potential to motivate safety-directed behavior; for instance, these assumptions can foster behavior such as communicating errors ("Word gets around who made the error") or expressing a strong preference for an error-free work environment ("Don't be a jackass"). However, with regard to the functioning of a reporting system these assumptions triggered dysfunctional dynamics promoting blame and underreporting of errors. Similarly, the assumption "*If it is formalized, it works",* which was derived from identified best practices in both plants, seems to be an important and necessary pre-condition for safe work in nuclear power plants. However, this assumption and its associated behavior may also result in safety-critical behavior, e.g., when working conditions become more stressful and individual workload is high leading to dysfunctional behavior such as non-compliance and withholding important information. One may tentatively conclude that the disclosure of basic assumptions of a given culture is not enough. Especially when it comes to the optimization of safety culture, a directed change of assumptions (if this is possible at all) may also foster non-intended side effects. Thus, a culture assessment should also include an identification of contextual factors, in which safety behavior stemming from dominant assumptions in a given culture is embedded. In line with this, one may tentatively suggest that rather than changing assumptions, the focus should be on changing the system to minimize the influence of dominant assumptions in triggering dysfunctional dynamics.

Moreover, comparing the results of both studies clearly suggests that it is worthwhile to dig deeper levels of culture. Whereas the artifact- and gap analysis identified similar safety-relevant issues in both plants, the assumption analysis revealed different underlying dynamics and assumptions contributing to these. Thus, changing the culture based solely on diagnosing manifestations on upper cultural levels (e.g., by means of questionnaires) runs the risk of ignoring important contributors to dysfunctional dynamics.

Besides the successful development and application of our assessment approach, there are also several limitations. The first limitation refers to the artifact analysis, which aimed at the identification of practices that are perceived as relevant in a given culture based on its unique learning outcomes. Both applications of this method revealed that middle managers had problems in identifying actual implemented practices. It became evident that it is a hard task to rate and differentiate the practices according to the actual and a somewhat wishful implementation. We tried to meet this challenge in study 2 by asking middle managers for practices that are often discussed in the plant. However, we assume that a rating method might not be the best way to exclude potential desirability and missing knowledge on the implementation degree of a work practice.

A second limitation refers to the conducted gap analyses. Initially it was planned to gather a full set of socio-demographic data of the participants (e.g., age, position, department affiliation, etc.), which allows for statistically identifying potential sub-groups (and therefore sub-cultures) of specific gap perceptions. However, members of the works council had objections concerning the loss of anonymity and confidentiality of the obtained data. Thus, our assessment unfolded an integrative perspective on the assessed cultures — in contrast to a differentiating and a fragmented perspective (Martin & Frost, 2004) — by focusing on characteristics shared by all or the majority of plant members. However, during the assessment (especially during the third assessment step) it became evident that some of the elicited assumptions were shared only by a specific subgroup (e.g., the operational shift personnel), but these were dropped when workshop participants from all departments of the plant rated their dominance in their culture. Thus, future applications of our approach should focus more strongly on the identification of subcultures and their specific assumptions.

A final limitation refers to the missing knowledge about the successful implementation of safety-directed changes based on our assessment results. During our assessment in both plants, there were external influences on the development of interventions to optimize safety culture in both plants. After the German government decided to phase-out nuclear power industries in 2001, a new government reversed this decision in 2010 and extended the life spans of the country's nuclear plants until at least 2036. But in June 2011, mindful of the public's concerns about nuclear safety after the Fukushima disaster, the government made a U-turn and decided to shut down the eight oldest plants immediately, and the nine remaining plants within the next 11 years. As a consequence, participation and willingness to implement interventions collapsed, i.e., planned interventions derived from the assessment were not pursued. We learned from this that although culture is supposed to be stable and hard to change, there are circumstances (in this case, political decisions) that have an enormous and immediate impact on culture and corresponding promotion of culture change, independently of internal change efforts. Thus, in future applications of our approach its potential in guiding and implementing cultural change needs to be proven.

References

Balogun, J., & Johnson, G. (2004). Organizational restructuring and middle manager sensemaking. *Academy of management journal*, *47*(4), 523-549.

Boglarsky, C. A., & Kwantes, C. T. (2004). Ideal and actual culture: How different is too different. In *65th Annual Conference of the Canadian Psychological Association, St. John's, Newfoundland*.

Brooks, B. (2008). The natural selection of organizational and safety culture within a small to medium sized enterprise (SME). *Journal of Safety Research, 39*, 73-85.

Díaz-Cabrera, D., Hernandez-Fernaud, E., & Isla-Díaz, R. (2007). An evaluation of a new instrument to measure organisational safety culture values and practices. *Accident Analysis & Prevention*, *39*(6), 1202-1211.

García-Herrero, S., Mariscal, M. A., Gutiérrez, J. M., & Toca-Otero, A. (2013). Bayesian network analysis of safety culture and organizational culture in a nuclear power plant. *Safety science*, *53*, 82-95.

Guldenmund, F. W. (2000). The nature of safety culture: a review of theory and research. *Safety Science 34*, 215-257.

Guldenmund, F. W. (2007). The use of questionnaires in safety culture research – an evaluation. *Safety Science 45(6)*, 723-743.

Guldenmund, F. W. (2010a). *Understanding and exploring safety culture*. TU Delft, Delft: University of Technology.

Guldenmund, F. W. (2010b). (Mis)understanding safety culture and its relationship to safety management. *Risk analysis*, *30*(10), 1466-1480.

Haber, S., Shurberg, D. A. & Hofmann, D. (1995*). Safety culture management: The importance of organizational factors*: United States Government.

Hale, R., Kirwan, B. & Guldenmund, F.W. (1999). Capturing the river: multilevel modelling of safety management. In J. Misumi, B. Wilpert and R. Miller (eds.), *Nuclear Safety: A Human Factors Perspective*, Taylor & Francis: 161-181.

Hale, A. R. (2000). Culture's confusions. Editorial for the Special Issue on safety culture and safety climate. *Safety Science, 34*, 1-14.

Hofstede, G. R. (2001). *Culture's consequences* (2nd ed.). London: Sage Publications.

Hopkins, A. (2002). *Safety Culture, Mindfulness and Safe Behaviour: Converging ideas?* Paper prepared for the Jim Reason Festschrift. Canberra: NRCOHSR.

INSAG - International Safety Advisory Group (1991). *Safety Culture*. Safety Series No. 75-INSAG-4. Vienna: International Atomic Energy Agency.

INSAG – International Safety Advisory Group (1999). *Management of Operational Safety in Nuclear Plants*, INSAG – 13. Vienna: International Atomic Energy Agency.

INSAG – International Safety Advisory Group (2002). *Key Practical Issues in Strengthening Safety Culture*, INSAG – 15. Vienna: International Atomic Energy Agency.

IAEA (1994). ASCOT Guidelines. IAEA-TECDOC-743. Vienna: International Atomic Energy Agency.

IAEA (1998). *Developing safety culture in nuclear activities: Practical suggestions to assist progress.* Safety Reports Series No. 11. Vienna: International Atomic Energy Agency.

IAEA (2003). Means of Evaluating and Improving the Effectiveness of Training of Nuclear Power Plant Personnel. IAEA TECDOC No. 1358. Vienna: International Atomic Energy Agency.

IAEA (2008). *SCART Guidelines: Reference report for IAEA Safety Culture Assessment Review Team (SCART)*. IAEA Service Series 16. Vienna: International Atomic Energy Agency.

IAEA (2009). *The Management System for Nuclear Installations*. Safety Guide No. GS-G-3.5. Vienna: International Atomic Energy Agency.

IAEA (2013). *Managing Human Performance to Improve Nuclear Facility Operation*. IAEA NG-T-2.7. Vienna: International Atomic Energy Agency.

IAEA (2014*). IAEA Report on Human and Organizational Factors in Nuclear Safety in the Light of the Accident at the Fukushima Daiichi Nuclear Power Plant* (International experts meeting Vienna, 21-24 May 2013). Vienna: International Atomic Energy Agency.

Kennedy, R. & Kirwan, B. (1998). Development of a hazard and operability-based method for identifying safety management vulnerabilities in high risk systems. *Safety Science 30*(3): 249-274.

Kilmann, R. H. and M. J. Saxton (2007). Culture-gap survey - Assessing actual versus desired cultural norms. Newport Coast, CA, Organizational Design Consultants.

Kostova, T. (1999). Transnational transfer of strategic organizational practices: A contextual perspective. *Academy of Management Review 24(2)*, 308-324.

KTA - Kerntechnischer Ausschuss. (2012): *KTA 1402 Integriertes Managementsystem zum sicheren Betrieb von Kernkraftwerken, Fassung 2012-11*. Salzgitter: KTA-Geschäftsstelle. http://www.kta-gs.de/d/regeln/1400/1402_r_2012_11.pdf

Lee, T. and K. Harrison (2000). Assessing safety culture in nuclear power stations. *Safety Science 34(1-3)*: 61-97.

Marais, K., Saleh, J. H., & Leveson, N. G. (2006). Archetypes for organizational safety. *Safety science, 44(7)*, 565-582.

Martin, J. & Frost, P. (2004). The organizational culture war games: a struggle for intellectual dominance. In S. Clegg, C. Hardy, W. Nord & T. Lawrence (eds.), *Handbook of Organizational Studies* (pp. 599-621). London: Sage

Myers, D. J., Nyce, J. M., & Dekker, S. W. (2014). Setting culture apart: Distinguishing culture from behavior and social structure in safety and injury research. *Accident Analysis & Prevention*, *68*, 25-29.

NRC - Nuclear Regulatory Commission. (2011). Final Safety Culture Policy Statement (NRC-2010-0282). *Federal Register*, *76*(114), 34773.

Pidgeon, N. (1998). Safety culture: key theoretical issues. *Work & Stress*, *12*(3), 202-216.

Reason, J. (1998). Achieving a safe culture: theory and practice. *Work & Stress*, *12*(3), 293-306.

Reiman, T., & Rollenhagen, C. (2014). Does the concept of safety culture help or hinder systems thinking in safety? *Accident Analysis & Prevention*, *68*, 5-15.

Rollenhagen, C. (2010). Can focus on safety culture become an excuse for not rethinking design of technology? *Safety Science*, *48*(2), 268-278.

Schein, E.H. (1985). *Organizational culture and leadership (First edition)*. San Fransisco, Jossey-Bass.

Schein, E.H. (1990). Organizational Culture. *American Psychologist, 45(2),* 109-119.

Schein, E.H. (1992). *Organizational culture and leadership, (Second Edition)*. San Fransisco, Jossey-Bass.

Schein, E. H. (1999). *The corporate culture survival guide : sense and nonsense about culture change*. San Francisco, Calif., Jossey-Bass.

Schein, E.H. (2004). *Organizational culture and leadership (Third Edition)*. San Fransisco, Jossey-Bass.

Schein, E.H. (2014). *National and Occupational Culture Factors in Safety Culture*. Paper presented at IAEA meeting, April 9, 2014.

Silbey, S. S. (2009). “Taming Prometheus: Talk About Safety and Culture.” *Annual Review of Sociology* 35: 341-369.

VGB - European technical association for electricity and heat generation. (2013). *VGB PowerTech Annual Report* (Juli 1, 2012 to June 30, 2013). Essen: VGB PowerTech e.V., August 2013. https://www.vgb.org/annual_report.html?dfid=64376

Weil, R. and G. Apostolakis (2001). Identification of important organizational factors using operating experience. Safety culture in nuclear power operations. In B. Wilpert and N. Itoigawa, *Nuclear Safety: A Human Factors Perspective,* 139-168. London, Taylor & Francis:.

Wilderom, C. P., van den Berg, P. T., & Wiersma, U. J. (2012). A longitudinal study of the effects of charismatic leadership and organizational culture on objective and perceived corporate performance. *The Leadership Quarterly*, *23*(5), 835-848.

Wilpert, B.,& Schöbel, M. (2007). *Challenges and Opportunities of Assessing Safety Culture*. Paper presented at the OECD/CCA Workshop on Human Factors in Chemical Accidents and Incidents, Germany.

Zohar, D., & Luria, G. (2005). A multilevel model of safety climate: cross-level relationships between organization and group-level climates. *Journal of Applied Psychology*, *90*(4), 616.