Full paper

Building bridges between safety culture science and practice

Wouter Martinus Petrus Steijn[1], Dolf van der Beek[1] en Jop Groeneweg[1,2,3]

Trefwoorden: Safety, Safety culture, Culture assessment, Culture interventions

Abstract

In the pursuit of further reducing accidents, organizations are implementing safety culture improvement programs. While ideally this safety culture practice would draw upon scientific research for effective interventions or provide the data for further validation of the interventions, there is a noticeable gap in the feedback loop between science and practice. Through interviews with experts active in safety culture science and practice, this study highlights the differences in approaches to safety culture in science and practice. This paper seeks to draw attention to this gap and provides suggestions on how to bridge this gap to empower the safety culture domain towards developing evidence-based interventions.

Samenvatting

Om het aantal ongevallen verder terug te dringen implementeren organisaties programma's om de veiligheidscultuur te verbeteren. Idealiter zou hierbij gebruik gemaakt worden van wetenschappelijk onderzoek voor effectieve interventies en zouden de resultaten gebruikt kunnen worden voor verdere validatie en optimalisatie van de interventies. Er is echter een gat in de feedbackloop tussen wetenschap en praktijk. Door middel van analyse van interviews met experts werkzaam in de wetenschap en de praktijk, belicht dit onderzoek de verschillen in benaderingen van veiligheidscultuur in wetenschap en praktijk. Dit artikel wil de aandacht vestigen op deze kloof en geeft suggesties over hoe deze kloof overbrugd kan worden om zo evidence-based interventies te ontwikkelen om de veiligheidscultuur in organisaties te verbeteren.

Introduction

To support the progress in reducing the number of accidents further, many organizations have put programs in place aimed at improving their safety culture. Safety culture is often seen as the next logical step in improving safety performance after optimizing technology, systems, and human behavior (Groeneweg, 1992; Guldenmund, 2000). As a result, organizations introduce a plethora of programs and interventions through which they attempt to improve their safety culture (van Kampen et al., 2019). However, despite this wealth of data, little is known about the effectiveness of safety (culture) interventions (Nielsen, 2014; van Kampen et al., 2019, Dyreborg et al., 2022). In this paper, we look into ways on how to bring the practice and science of safety culture closer together.

The limited body of evidence is especially striking given the fact that safety culture has been of scientific interest for decades (Guldenmund, 2000; Keenan, Kerr & Sherman, 1951; Arzahan, Ismail & Yasin, 2022; Turner et al., 1989). Overall, only a few (safety) culture change intervention studies exist in the safety literature (DeJoy, 2005; Hale et al., 2010; Nielsen, 2014). Safety interventions are defined by Robson and colleagues (2001) as 'an attempt to change how things are done in order to improve safety'. In their review on safety interventions, Dyreborg and colleagues (2022) included no studies that investigated safety culture as a means to improve safety and reduce accidents since none of the available studies met the selection criteria used for their review (i.e. RCT, quasi randomized and experimental study designs, - single group - controlled before and after study designs and studies utilizing serial measures).

Instead, Dyreborg and his colleagues (2022) looked at interventions to improve safety climate and safety norms since these are widely available. Safety climate is distinct from safety culture as it is more focused on the perceptions, rather than the actual underlying values, attitudes and beliefs associated with culture. Safety climate may be changed through leadership-based interventions, goal setting and performance feedback methods and a few other approaches to modify values or norms related to safety at work (Dyreborg et al., 2022). Dyreborg et al. (2022) state that the number of well-conducted studies on safety interventions is not very high. Dyreborg et al. (2022) also mention that the research methods used in these studies are poor and it is not clear how the steps taken in the studies to improve safety climate connect with the final goal of reducing injuries. Obtaining scientifically rigid evidence on the relationship between safety culture and safety performance is

[1] *Work Health Technology, TNO, Sylviusweg 72, 2333 BE Leiden. Correspondentie: wouter.steijn@tno.nl*

[2] *Faculty of Technology, Policy and Management, TU Delft, Jaffalaan 5, 2628 BX Delft*

[3] *Sociale wetenschappen, Leiden universiteit, Wassenaarseweg 52, 2333 AK Leiden*

difficult to obtain (e.g. Hale et al. 2010; Hopkins, 2006) and sometimes a comparison even leads to conflicting results (Antonsen, 2009; Boskeljon-Horst et al., 2022). Those scientific endeavors that propose a framework to obtain this evidence (e.g. Ayob, Hassan & Hamid, 2022; Siuta et al., 2022), often lack a follow-up study in which these frameworks are actually operationalized and implemented for further validation. As a result, science is yet to provide a definitive answer on whether the relationship between culture and performance exists (Kilcullen et al., 2022; Morello et al., 2013; Dyreborg et al., 2022).

With regard to safety culture practice, there are organizations that make use of well-established culture assessment measures, such as the Hearts and Minds card sorting assessment method based on a maturity model (although initially not intended as an assessment tool (Lawrie, Parker & Hudson, 2006)) or the Dupont Bradley Curve to benchmark and develop a company's safety culture maturity (dss+, 2023). Yet they rarely provide insights into their effects with any scientific rigor (Fleming, 2001). Moreover, according to a thorough analysis of safety culture maturity models by Goncalves Filho & Waterson (2018), the models and the tools used to measure them lacked sufficient proof of their validity and reliability, and more studies are necessary. Generally, the use of interventions is based on other criteria (such as best practices shared by other organizations or based on legal requirements) rather than scientific evidence (van Kampen et al., 2023). Although there are some notable exceptions of scientific initiatives by organizations like the Safety Culture Maturity Model (Fleming, 2001), which has been applied and validated in various industries, such as construction, oil and gas, and nuclear. However, these initiatives are often more in the lines of safety climate than safety culture.

In 2016, Cooper stated that there was a mismatch between safety culture theory (i.e., science) and safety culture practice, and the same conclusion was reached in the analysis of Dutch (Swuste et al., 2019) and Belgian safety professionals (van Nunen, Reniers & Ponnet, 2018). Bridging the gap between these two fields in order to pool their resources more effectively could give the safety culture domain the necessary momentum toward scientifically validated safety culture interventions. With this goal in mind, we interviewed both safety culture consultants and scientists, as representatives of practice and science, to gain insights into their perspectives on safety culture and their approach to initiating change.

Method

In 2019, we approached 20 Dutch consultants and scientists working in the field of safety culture to participate in our interviews. In total, eleven experts took part, five of whom were scientists and six of whom were consultants. The consultants were experts whose core business is predominantly to assist organizations with their safety culture improvement programs. The scientists were experts whose core business is to conduct predominantly theoretical research on the topic of safety culture. The participants had varying backgrounds, ranging from organizational psychology to anthropology and social and behavioral sciences.

The interviews took 1.5 to 2 hours each and were semi-structured. The interviewers had a set of questions as guidance to address key topics during the interview, but, other than that, they followed the flow of the interview depending on the interviewees and their field of expertise. We also shared the questions with the interviewees beforehand as preparation. These questions can be found in Appendix A. Each interview started with a short introduction to our study and interests before asking the first question, namely, "Please tell us what organizational culture is for you", after which the focus was quickly moved towards the safety culture within organizations.

To allow the interviewers to focus on the interview, the entire interview was recorded with the consent of the interviewee. The recordings were sent to a transcription organization that used a speech recognition algorithm to transcribe the interviews and was able to distinguish the different speakers reliably. The resulting transcript was double-checked by a human, given the academic purpose of obtaining reliable data. This resulted in good-quality transcriptions in which only a few names were misspelled and there were some illegible sections. Next, a thematic analysis of the transcripts was carried out.

Thematic analysis

The thematic analysis was performed in seven phases: collecting, compiling, familiarizing, disassembling, reviewing, interpreting, and concluding. This approach was based on two sources. The first was a methodological review of thematic analysis that outlined the steps one should include to perform a rigorous thematic analysis (Castleberry & Nolen, 2018). The second was a practical step-by-step guide for performing thematic analysis (Maguire & Delahunt, 2017). The latter approach was based on Braun & Clarke's (2006) theoretical framework for thematic analysis and explained this framework by applying it to comparable qualitative data to those used for the current project: transcripts of structured interviews. We have already described the first two phases, collecting (i.e., semi-structured interviews) and compiling (i.e., transcriptions). In Appendix B, we provide a brief description of the remaining five phases.

Results

The thematic analysis resulted in the following two main themes that came across all the interviews: *different approaches of consultants and scientists and different aspects of safety culture.* Below, we will describe each of these aspects in further detail.

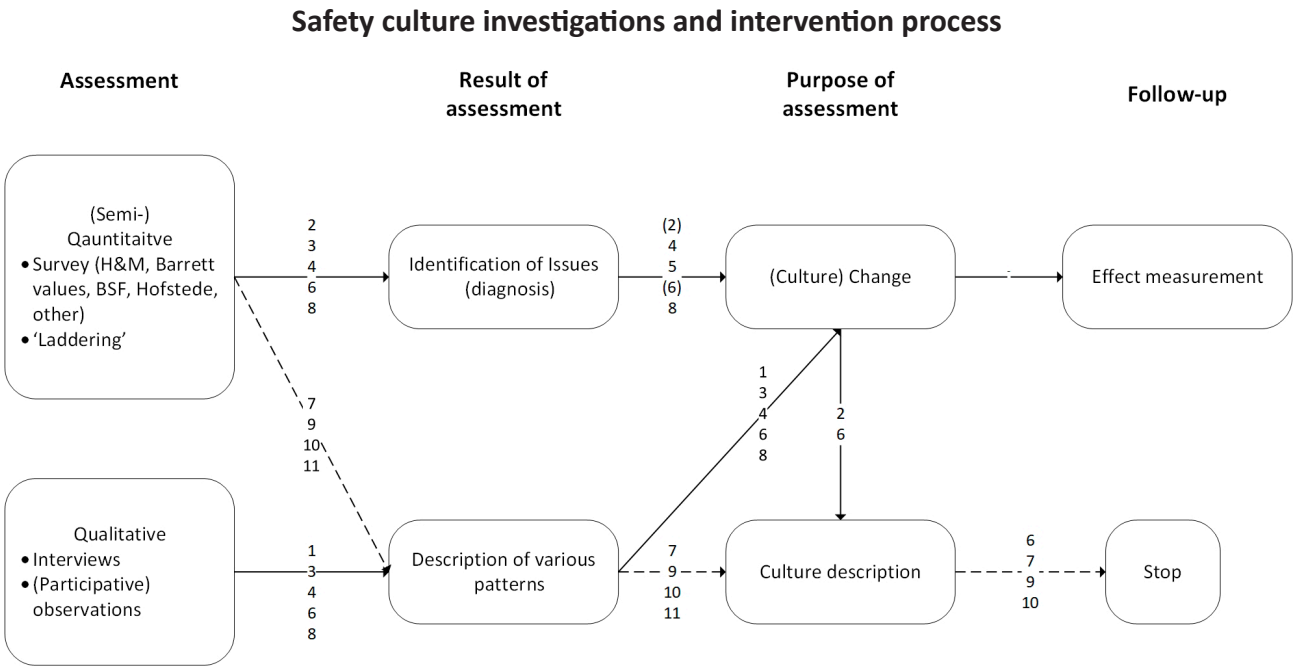

Figure 1. Interviewees (by number) mapped onto the safety culture investigation and intervention process. Interviewees 1-6 were more active in practice (consultant) and interviewees 7-11 were more active in the science field (scientist).

Differences between safety culture science and practice

Based on all of our interviews we were able to distinguish two separate paths (see Figure 1). In Figure 1, we mapped the processes that the interviewees described as part of their work concerning safety culture. It shows that all the interviewees generally follow the same steps. They start with an assessment, either quantitative or qualitative, with a specific result and purpose, and which can subsequently lead to a follow up.

The first identified path (full arrows in Figure 1) starts with a safety culture assessment to either diagnose the 'weak spots' within an organization or to describe various (behavioral) patterns inherent to the organization. This is with the purpose of initiating a change to remove undesired patterns or to alleviate weak spots. This path was described by experts who are predominantly active in safety culture practice (interviewees 1-6 in Figure 1). Noteworthy is that in some cases this may have led to a final description of the culture within an organization, but none of the interviewees mentioned any form of effect measurement as follow-up (indicated in Figure 1 by the lack of an arrow). In a few cases it was mentioned that clients do not ask for or want to pay for a follow-up measurement. The main focus lies on implementing the desired change within an organization. To implement this change, they make use of various models for assessing culture, such as the Hearts and Minds method (Hudson et al., 2004), Barrett Values model (Barrett, 2006), and the Laddering method (Rugg et al., 2002). The interviewed experts decide which model to use based on the situation and needs of the organization and which they have in their own toolbox. These experts are not so concerned with the exact definition of culture. They focus on changing processes, behaviors, and sometimes safety outcomes but almost always without gathering data concerning the effectivity of their methods. Most interviewees expect a culture to change when "something" changes in people's heads. This could be a mindset, increased awareness, or new insights or understandings. Often, this insight might result from a so-called "Aha-Erlebnis," which radically changes their view on aspects of reality shared with others and which might have a lasting impact on their behavior. In the case of safety culture, this might affect their safety behavior.

The second identified path (dashed arrows in Figure 1) is more characteristic of the scientific approach to safety culture. The assessment is primarily used to come to a description of the safety culture based on the assessed (behavioral) patterns. This description is the endpoint of this path; the intervention part is left to the client's discretion. Figure 1 shows that this path was primarily described by experts primarily active in the scientific field of safety culture (interviewees 7-11).

There is significant variety in the kinds of interventions used by the interviewees to change (safety) culture. These interventions are based on various mechanisms such as the creation of a shared understanding (e.g., storytelling, social contagion or (guided) dialogue), provoking personal insight (e.g., training or a confrontation with the description of culture) or by improving personal motivation (e.g., introducing rewards and punishment). A commonly mentioned approach is the use of so-called safety ambassadors. These are a small group of motivated employees who work with the consultants and function as their extension to enact change

within the organization. Once properly trained, these ambassadors can continue to enact the desired change once the consultants themselves have left. This approach also helps to reduce the gap between the management and the work floor and to facilitate social contagion of new behavior.

Different aspects of safety culture

Throughout the interviews, the interviewees mentioned various aspects that are related to different conceptualizations of culture in general. These aspects were related to, for example, behavioral practices, the structure in which the behavior takes place, and the underlying beliefs, values, and attitudes. No clear distinction was found in the use of these conceptualizations between experts from the field of practice or science. However, the variation in possible conceptualizations is relevant to take into account and should be made explicit when attempting to bring the two fields together.

The themes we identified in our interviews closely resembled the categories that Edwards, Davey & Armstrong (2013) proposed for a synthesized conceptualization of safety culture and that we will therefore use in the remainder of this paper. Here, we labeled these culture categories as normative, pragmatic, or ideational. Normative culture concerns changeable factors, such as the policies, procedures, and structure of an organization; pragmatic culture concerns behavioral practices and norms; and ideational culture involves beliefs, values, attitudes, assumptions, and expectations (Edwards, Davey & Armstrong, 2013). Below, we will explore each conceptualization further in turn.

- *Pragmatic culture:* Pragmatic culture concerns the behavioral practices and norms that are considered appropriate. In our interviews with the experts, we noticed that they regularly used terms related to the behavior of individuals and between groups. These were terms such as patterns, habits, and routines, that is, what people considered normal behavior, and can be seen to happen often. They included both physical behavior and communication patterns. In other words, it is not only what people do but also what they talk about that can be informative about the culture.
- *Normative culture:* Normative culture concerns changeable factors, such as the policies, procedures, and structure of an organization. The experts recognized and often referred to these factors as the context in which the previously described behaviors occur, and it is only natural that they influence (or limit) the kind of behavior that takes place. It was also noted, however, that the policies, procedures, and structure of an organization are the result of the culture, and, in this case, this is because of the (groups of) individuals who created this structure. This is indicative of the importance of power and leadership when thinking about culture (and changes thereof) and of circular processes (structure influences culture; culture influences structure) that show that culture is not a fixed condition.
- *Ideational culture:* Ideational culture concerns beliefs, values, attitudes, assumptions, and expectations. The interviewees often referred to these terms as well. An awareness of important values and that certain behavior is incompatible with those values can lead to people changing their behavior; however, people are not always aware of how values affect their behavior and find it difficult to name important values. Therefore, discovering which values are important can be a difficult process. The focus should be on identifying shared values and beliefs (as opposed to individual values) as they are likely to have the biggest impact on people's behavior in a certain context.

Discussion

In this paper, we set out to bridge the gap that exists between the practice and science of safety culture. In our interviews, we verified the existence of two separate paths that seem to be symptomatic for how experts treat safety culture in practice and scientifically. This result is in line with previous findings indicating a mismatch between the scientific work on safety culture and what happens in practice (Cooper et al., 2016; van Nunen, Reniers & Ponnet, 2018; Swuste et al., 2019). Neither approach has any form of follow up to verify the effect of enacted changes (path 1) or interventions that may have been undertaken as a result of the description (path 2). This makes it impossible for a feedback loop to occur between experts active in practice and science, which would be necessary to accumulate the evidence required to prove or disprove the effect of safety culture models on safety performance. A better understanding of what sets these fields apart is key to overcome these differences and work toward better synergy.

One factor that may be partly responsible for the observed gap between practice and science, is the fact that culture is multifaceted. Overall, we identified the normative, pragmatic, and ideational aspects (see also Edwards, Davey & Armstrong, 2013) in our interviews. Safety culture discussions do not always unambiguously label what aspect they are addressing when they mention culture. A holistic approach covering all aspects is most effective to address safety culture. However, each aspect of culture likely requires different methodologies to assess, describe and change (Cooper, 2016). If we want practice and science to line up and elevate the safety culture discussion, no confusion should exist on this level.

Another aspect is the fact that there is a clear discrepancy between the objectives in practice and in science. In practice, the focus lies on facilitating change within organizations rather than the necessity to gather (scientific) evidence of the effectiveness of their interventions. Their process is based on the toolbox they believe is effective. The experts we interviewed indicate they rely primarily on their own impressions or on what we could typify as "social proof." That is, the organization itself often viewed change as "improvement" and therefore considered

the intervention to be a "success". Science, on the other hand, rarely progresses beyond describing and measuring culture, its components, and the psychometric properties of its measurement tools. This is not always by choice; organizations that wish to improve their safety culture often leave little room in the project to follow the empirical cycle completely in a scientifically rigid manner.

In science, on the other hand, safety culture is predominantly described, modeled, and measured to progress the scientific discussion. Whereas in practice a much more pragmatic approach on safety culture is taken aimed at diagnosing the problem and implementing needed change. The latter approach is more likely to result in an intervention that is aligned with the wishes of the organization, even if scientists might argue that it is not the optimal approach to the identified problems. A recent study showed that relatively few safety professionals, including consultants, consider the scientific literature to be a viable source of safety interventions (van Kampen et al., 2023). A majority of scientific work is dedicated to determining how to measure culture (Ayob, Hassan & Hamid, 2022; Siuta et al., 2022) and to the psychometric properties of the tools used (e.g. Alruqi, Hallowell & Techera, 2018; Curran et al., 2018; Hartnell et al., 2019), rather than to methods to enact change in the culture and their effectiveness.

The scientific method consists of a continuous process of modeling, measuring, implementation, validation, and improvement (of the modeling). Here, we observed that currently most attention in science goes to modelling and measuring (see Figure 1: The description of various patterns), while in practice the focus lies at implementation (see Figure 1: Change). Little progress will be made in the validation or improvement of safety culture models and methods, as long as little effort will be put in follow-ups in the form of effect measurements. An important step that the scientific community could take is to align its research questions with the questions that are posed by organizations in practice. Rather than describing new models or verifying the psychometric properties of existing tools, more effort should be put into, for example, finding the active components that make interventions effective. A good example to this practice is the use of the Intervention Mapping (IM) method as developed by Eldredge and colleagues (2016). IM encourages a structured approach to the development of interventions through six iterative steps that cover (among others) a needs assessment, linking change objectives to theoretical methods and an evaluation plan. By following the steps, an evidence-based intervention is developed that is tailored to the specific needs of the target population (see van der Beek et al. (2023) for an implementation of the IM method in the safety domain).

Safety is not the only domain that struggles with optimizing the potential synergy between practice and science. Looking at other domains can give some potential interesting ways forward in search of practical solutions on how to bridge the gap between science and practice. Nyström et al. (2018), for example, sought to strengthen collaboration and partnership research for the improvement of health and social services between researchers and practitioners. Their article provides a detailed roadmap for research collaboration, based on five approaches to co-production of research: (1) co-creation; (2) co-design; (3) co-production; (4) co-implementation; and (5) co-evaluation. To support the practitioners involved at different levels in the healthcare system and build trust and understanding, time, interaction spaces and project management and communication skills are needed during the research collaboration.

In the adult education domain, Mohajerzad et al. (2021) explored how to promote evidence-informed education by bridging the gap between research and practice. They argued that scientific evidence is not the only factor in practitioners' decision-making, and that contextuality and quality of evidence should be considered. They also advocate for collaborative research that involves both academics and practitioners (i.e. teachers), and that focuses on the relevance and applicability of the findings. Their survey results show that practitioners tend to trust research findings regardless of the research team composition, but that cooperation between practitioners and scientists does not guarantee the use of research knowledge in the field.

These studies provide important lessons from which the safety domain can benefit. Namely, they highlight the importance of having a structured program when developing interventions, including steps for (amongst others) design, implementation and evaluation. In addition, they argue that experts from practice and science should cooperate starting with the development of the intervention rather than once implementation starts. Closer cooperation from the design phase can help align the research questions posed by science with the questions asked by organizations and result in more valid insights on the effects of the intervention that is implemented. This will in turn help progress the scientific continuous modeling, measuring, implementing, validating, and improving (of the intervention). This may shift the focus from describing new models or verifying the psychometric properties of existing tools into developing effective interventions and finding the active components that make interventions effective. This also implies that scientists will need to incorporate qualitative methods structurally in their safety culture related research.

There are several limitations that should be considered while interpreting our current results. Firstly, we did not include any views of organizations' managers or leaders regarding their culture improvement projects. They could have provided a different perspective on the meaning of culture for their organization. Secondly, our sample is relatively small (n=11) as we made use of a convenience

sample of Dutch safety professionals for our interviews. This could have implications for the generalization of the results.

Conclusion

Organizations are driven by a strong commitment to elevate their safety performance. A focus on enhancing safety culture stands out as one of the most promising avenues to achieve this goal. Ideally, this journey is a dynamic partnership, with both experienced consultants and innovative scientists engaging in a perpetual cycle of developing, testing, and refining the tools and strengthening the foundational theories that underlie it. However, in practice, this Plan-Do-Check-Act-like process does not work effectively. The scientific community is diligently crafting intricate theoretical frameworks and strategies, yet their efficacy in real-world scenarios often remains unproven, creating a roadblock in the scientific process. On the other side, consultants in practice, armed with these sometimes still underdeveloped scientific concepts, possess invaluable insights into the effectiveness of their practical implementations. Yet, regrettably, this essential information is not shared with the scientific community, hampering the progress of implementation and improvement. The solution lies in fostering an exchange between science and practice to close the feedback loop.

The scientific field can empower efforts in practice to better align their strategies with the specific needs of organizations, advancing the scientific foundation for safety culture as both a theory and a catalyst for transformative change. This is only possible if scientists are provided with the insights that can only be obtained with the application of safety culture theories and models in practice. In return, this collaboration equips consultants with a robust arsenal of scientifically validated tools, propelling them towards the realization of the shared goal of creating safer, more safe workplaces.

Acknowledgments

The authors would like to acknowledge the input provided by Frank Guldenmund from Delft TU during the execution of the study described in this paper. This research was supported by the Dutch Ministry of Social Affairs and Employment (SZW) [Occupational Safety Research Program 2019].

Literature

Alruqi WM, Hallowell MR, Techera U. (2018) Safety climate dimensions and their relationship to construction safety performance: A meta-analytic review. Safety Science; 109: 165-173.

Antonsen S. (2009) Safety culture assessment: A mission impossible? Journal of contingencies and crisis management; 17 (4): 242-254.

Arzahan ISN, Ismail Z, Yasin SM. (2022) Safety culture, safety climate, and safety performance in healthcare facilities: A systematic review. Safety Science; 147: 105624.

Ayob AN, Hassan CRC, Hamid MD. (2022). Safety culture maturity measurement methods: A systematic literature review. Journal of Loss Prevention in the Process Industries; 104910.

Barrett R. (2006) Building a values-driven organization – a whole systems approach to cultural transformation. Routledge.

Eldredge BLK, Markham CM, Ruiter RAC, Fernandez ME, Kok G, Parcel GS. (2016) Planning health programs: An intervention mapping approach (4th ed.). Jossey-Bass.

Braun V, Clarke V. (2006) Using thematic analysis in psychology. Qualitative Research in Psychology; 3 (2): 77-101.

Castleberry A, Nolen A. (2018) Thematic analysis of qualitative research data: Is it as easy as it sounds? Currents in Pharmacy Teaching and Learning; 10 (6): 807-815.

Cooper D. (2016). Navigating the safety culture construct: A review of the evidence. BSMS, Franklin. Retrieved from https://www.behavioral-safety.com/articles/safety_culture_review.pdf.

Curran C, Lydon S, Kelly M, Murphy A, Walsh C, O'Connor P. (2018) A systematic review of primary care safety climate survey instruments: Their origins, psychometric properties, quality, and usage. Journal of Patient Safety; 14 (2): e9-e18. https://doi.org/10.1097/PTS.0000000000000393.

DeJoy DM. (2005) Behavior change versus culture change: Divergent approaches to managing workplace safety. Safety Science; 43: 105-129.

dss+. (2023) The DuPont Bradley Curve. Retrieved from dss+ Bradley Curve™ | dss+ Consulting (consultdss.com).

Dyreborg J, Lipscomb HJ, Nielsen K, Törner M, Rasmussen K, Frydendall KB, Bay H, Gensby U, Bengtsen E, Guldenmund F, Kines P. (2022) Safety interventions for the prevention of accidents at work: A systematic review. Campbell Systematic Reviews; 18 (2): e1234.

Edwards J, Davey J, Armstrong K (2013) Returning to the roots of culture: A review and re-conceptualisation of safety culture. Safety Science; 55: 70-80.

Fleming M. (2001) Safety culture maturity model. Report prepared for the Health and Safety Executive, OTO-2000/049, United Kingdom. ISBN 0 7176 1919 2.

Goncalves Filho AP, Waterson P. (2018) Maturity models and safety culture: A critical review. Safety Science; 105: 192-211.

Groeneweg J. (1992) Controlling the controllable: The management of safety. Proefschrift Rijksuniversiteit Leiden, DSWO Press, Leiden, The Netherlands.

Guldenmund FW. (2000) The nature of safety culture: A review of theory and research. Safety Science; 34 (1–3): 215-257.

Hale AR, Guldenmund FW, van Loenhout PLCH, Oh JIH. (2010) Evaluating safety management and culture interventions to improve safety: Effective intervention strategies. Safety Science; 48: 1026-1035.

Hartnell CA, Ou AY, Kinicki AJ, Choi D, Karam EP. (2019) A meta-analytic test of organizational culture's association with elements of an organization's system and its relative predictive validity on organizational

outcomes. Journal of Applied Psychology; 104 (6): 832-850.
Hopkins A. (2006) Studying organizational cultures and their effects on safety. Safety Science; 44: 875-889.
Hudson P, Parker D, Lawrie M, van der Graaf G, Bryden R. (2004) How to win hearts and minds: The theory behind the program. In SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production. OnePetro, Society of Petroleum Engineers London, United Kingdom.
Keenan V, Kerr W, Sherman W. (1951) Psychological climate and accidents in an automotive plant. Journal of Applied Psychology; 35 (2): 108-111.
Kilcullen MP, Bisbey TM, Ottosen MJ, Tsao K, Salas E, Thomas EJ. (2022) The Safer Culture Framework: An application to healthcare based on a multi-industry review of safety culture literature. Human Factors; 64 (1): 207-227. https://doi.org/10.1177/00187208211060891.
Lawrie M, Parker D, Hudson P. (2006) Investigating employee perceptions of a framework of safety culture maturity. Safety Science; 44 (3): 259-276.
Maguire M, Delahunt B. (2017) Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. All Ireland Journal of Teaching and Learning in Higher Education; 8 (3): 3351-3364.
Mohajerzad H, Martin A, Christ J, Widany S. (2021) Bridging the gap between science and practice: Research collaboration and the perception of research findings. Frontiers in Psychology; 12: 790451. https://doi.org/10.3389/fpsyg.2021.790451.
Morello RT, Lowthian JA, Barker AL, McGinnes R, Dunt D, Brand C. (2013) Strategies for improving patient safety culture in hospitals: A systematic review. BMJ Quality & Safety; 22 (1): 11-18.
Nielsen KJ. (2014) Improving safety culture through the health and safety organization: A case study. Journal of Safety Research; 48: 7-17.
Nyström ME, Karltun J, Keller C, Andersson Gäre B. (2018) Collaborative and partnership research for improvement of health and social services: Researchers' experiences from 20 projects. Health Research Policy and Systems; 16: 46. https://doi.org/10.1186/s12961-018-0322-0.
Robson LS, Shannon HS, Goldenhar LM, Hale AR. (2001) Evaluating the effectiveness of strategies for preventing work injuries: How to show whether a safety intervention really works. National Institute for Occupational Safety and Health (NIOSH).
Rugg G, Eva M, Mahmood A, Rehman N, Andrews S, Davies S. (2002) Eliciting information about organizational culture via laddering. Information Systems Journal; 12 (3): 215-229.
Siuta D, Kukfisz B, Kuczyńska A, Mitkowski PT. (2022) Methodology for the determination of a process safety culture index and safety culture maturity level in industries. International Journal of Environmental Research and Public Health; 19 (5): 2668. https://doi.org/10.3390/ijerph19052668.
Swuste P, Zwaard W, Groeneweg J, Guldenmund F. (2019) Safety professionals in the Netherlands. Safety Science; 114: 79-88.
Turner BA., Pidgeon N, Blockley D, Toft B. (1989) Safety culture: Its importance in future risk management. In Position Paper for the Second World Bank Workshop on Safety Control and Risk Management, Karlstad, Sweden, 6-9.
van Kampen J, Lammers M, Steijn W, Guldenmund F, Groeneweg J. (2019) The effectiveness of 48 safety interventions according to safety professionals. Chemical Engineering Transactions; 77: 307-312.
van Kampen J, Lammers M, Steijn W, Guldenmund F, Groeneweg J. (2023) What works in safety: The use and perceived effectiveness of 48 safety interventions. Safety Science; 162: 106072.
van Nunen K, Reniers G. Ponnet K. (2018) Measuring and improving safety culture in organisations: An exploration of tools developed and used in Belgium. Journal of Risk Research: 21 (5): 622-644.

Appendix A: Semi-structured interview questions

Main questions

1. What is organizational culture to you? (Are you using a particular definition? A metaphor? Are you using a model? If so, what is this model based on and what is its nature?)
   a. Do you consider safety culture as a separate construct?
   b. What is the relationship between safety culture and organizational culture?
2. Can safety culture or organizational culture be investigated empirically? If so, how?
3. Can safety culture or organizational culture be diagnosed or measured? What do you mean by diagnosis/measurement? What is the purpose of diagnosing/measuring? Is that a feasible goal?
4. How do you diagnose/measure organizational and/or safety culture (qualitative, quantitative, or mixed methods)? Is such a diagnosis descriptive or normative?
5. Can the organizational culture or safety culture be changed? Is that an autonomous process or a controlled or manageable process with interventions? How would you define those terms?
6. Is there an empirical relationship between the change in the organizational culture or safety culture and the (change in) performance of that organization? How would you describe the change and the change process and define the (change in) performance?
7. Have we covered your thoughts on organizational and safety culture or are there any other matters that you consider relevant?

Additional questions

1. How do you see the relationship between the behavior of organization members and organizational culture? Is that a dynamic process? Is there development or a

process of maturing or relapse? How would you define the terms “dynamic,” “development,” “maturity,” and “relapse”?

2. How do you see the relationship between organizational culture, the behavior of organization members, and business processes? Is that a dynamic process? Is there a “development” or “dynamic process” of “maturing” or “relapse”? How would you define those terms?
3. Is there an empirical relationship between the organizational culture (with characteristics to be described), the behavior of organization members (with characteristics to be described), and the quality/reliability of the course of business processes (with characteristics to be described)? How would you define those terms? What does the method of data collection and analysis look like?
4. Is there an empirical relationship between the organizational culture (with characteristics to be described), the behavior of organization members (with characteristics to be described), the business processes (with characteristics to be described), and the results of those processes (with characteristics to be described)? How would you define those terms? What does the method of data collection and analysis look like?
5. Do you know a useful measure of safety? Is this a process measure or an outcome measure? How would you define those terms?
6. What was the reason for (the wish for) that change/change process? What was the purpose of that change? What was the starting point(s) of the change process?

Appendix B: Thematic analysis

We performed the thematic analysis in seven phases: collecting, compiling, familiarizing, disassembling, reviewing, interpreting, and concluding (Braun & Clarke, 2006). The last five steps are described below.

- *Familiarizing.* To familiarize themselves with the data, all the researchers read the transcripts. This allowed them to gain a sense of the entire content of all the interviews and to understand a phrase more accurately due to familiarization with its entire context (Castleberry and Nolen, 2018).
- *Disassembling.* After becoming familiar with the data, we created meaningful groupings in the transcripts, known as codes. These codes were divided into larger meaningful themes, a process called “coding” (Castleberry & Nolen, 2018). According to Braun & Clarke (2006), a theme captures an important concept relevant to the research question(s). In this phase, themes were only constructed in a semantic manner. This means that only what the interviewees had stated literally was coded. In this phase, researchers were not yet allowed to interpret these codes (Braun & Clarke, 2006). Codes were created by multiple researchers independently to ensure intercoder reliability in the next phase (Castleberry & Nolen, 2018).
- *Reviewing.* In this phase, we compared all the themes and codes found by the individual researchers. The researchers discussed their individual findings for each interview and selected the themes that overlapped between them. This discussion was supported by visualizations of the codes and themes that emerged per interview. Whenever a theme was not found by multiple researchers, a discussion was held to decide whether the theme should be included. This process took place for all the interviews. The final themes were merged into one document, with each theme being supported by relevant codes retrieved from all the interviews. The intercoder reliability was considered satisfactory as most themes had emerged with each researcher independently and most discussions were about including or excluding various codes.
- *Interpreting.* In this phase, the discussion extended beyond the semantic level and started identifying the underlying ideas, assumptions, conceptualizations, and ideologies. Additionally, we decided which themes were most valuable in relation to our research questions. To achieve a comprehensive perspective, three additional organizational safety culture experts took part in the discussions. The researchers who performed the semantic level of analysis also participated to elucidate their specific findings and their context.
- *Concluding.* The final phase involved drawing conclusions based on the interpretation and related discussion concerning the data. In the results section, we will provide our main conclusion based on this process.