'Ladder'-based safety culture assessments inversely predict safety outcomes

Boskeljon-Horst, Leonie; Sillem, Simone; Dekker, Sydney W.A.

DOI
10.1111/1468-5973.12445

Publication date
2022

Document Version
Final published version

Published in
Journal of Contingencies and Crisis Management

Received: 27 January 2022 | Revised: 12 December 2022 | Accepted: 14 December 2022

DOI: 10.1111/1468-5973.12445

ORIGINAL ARTICLE

WILEY

'Ladder'-based safety culture assessments inversely predict safety outcomes

Leonie Boskeljon-Horst[1] | **Simone Sillem**[2] | **Sydney W. A. Dekker**[3]

[1]Royal Netherlands Air Force Headquarters, Breda, The Netherlands

[2]Department of Values, Technology and Innovation, Faculty of Technology, Policy and Management, Delft University of Technology, Delft, The Netherlands

[3]Safety Science Innovation Lab, School of Humanities, Languages and Social Science, Griffith University, Nathan Campus, Australia

Correspondence
Leonie Boskeljon-Horst, Royal Netherlands Air Force Headquarters, Luchtmachtplein 1, 4822 ZB Breda, The Netherlands.
Email: L.Boskeljon.Horst@mindef.nl and L.Boskeljon-Horst@tudelf.nl

Funding information
None

Abstract

There is little empirical evidence on the predictive value of safety culture assessments (SCAs) in relation to how accident-prone an organisation might be. Recently, Antonsen not just demonstrated how a quantitative SCA mispredicted future safety outcomes, but actually showed an inverse relationship between the assessment and subsequent critical incident investigation findings. To add to our understanding, this article presents research on whether a SCA has a predictive capacity for safety outcomes. Like in Antonsen's research, an opportunity emerged when a helicopter taxiing accident, resulting in a rotor strike occurred for a helicopter squadron that had just undergone a SCA. The assessment used 'culture ladder' rubrics for its findings, which allowed us to look for specific features in the subsequent independent accident investigation (in which the researchers were not involved). As with Antonsen's findings, our research shows that a 'ladder'-based assessment has little predictive value. Any predictive value it has is in the inverse of the assessment findings. For instance, where the SCA showed that the safety culture was very mature regarding finding a balance between safety and the mission at hand or the breaking of rules, the accident investigation pointed these out as the causes of the accident.

KEYWORDS

predictive value, safety, safety culture assessment, safety culture ladder, safety outcome

1 | INTRODUCTION

Despite the popularity of safety culture assessments (SCAs) since the 2000s, there is little evidence for their predictive capacity for actual safety outcomes. This study asks whether a SCA offers actionable insight into how accident- or incident-prone an organisation might be. A SCA might provide a static description of the safety culture of an organisation. The same safety culture as measured with an assessment is present when an accident happens. A comparison between a proactive SCA and the reactive description of the safety culture, as found in an incident or accident investigation, should, therefore, reveal similarities. To test this assumption, and building on previous research by Antonsen (2009), a comparative study was conducted. The study investigated the results of a SCA within an operational squadron of the Royal Netherlands Air Force (RNLAF) and compared these results to the results of an investigation of an accident that occurred within the same squadron just a few weeks after the SCA.

Antonsen has concluded that the predictive value of SCAs is limited (2009). In his study, a SCA using a 5-point Likert scale was compared to the results of a qualitative investigation of a major oil and gas incident. The results of a SCA using detailed workplace descriptions are compared to an accident investigation in this research article. More detail or nuance in the SCA may lead to more predictive value concerning how accident-prone a company is—whereas an assessment using a 5-point Likert scale, which lacks detail, may not reveal that information. Therefore, the objective is to determine the predictive value of this specific assessment regarding the accident and how accident-prone this squadron was. The study reviews whether the findings of the accident investigation match the findings of the SCA. It considers whether the safety culture weaknesses found in the SCA were the causes of the incident.

In the following paragraphs, a theoretical framework is provided, and the findings are presented. First, the results of the SCA and the findings of the accident investigation are given, followed by a comparison of these results. This study concludes with a discussion of the findings and their practical implications.

2 | THEORETICAL FRAMEWORK

2.1 | Safety culture

Since the late 1980s and early 1990s, the cause of accidents has been found in the fact that operators perform their duties or interact with technology in a particular culture (Wiegmann et al., 2004), specifically a safety culture. It is commonly agreed that interest in safety culture dates from the Chernobyl accident in 1986 (Health & Safety Executive, 2005; Obadia et al., 2007; Patankar et al., 2012; Thaden & Mitchell-Gibbons, 2008; Wiegmann et al., 2002). The term 'safety culture' was first introduced by the International Nuclear Safety Advisory Group (INSAG) (1991) and defined as 'that assembly of characteristics and attitudes in organisations and individuals which establishes that, as an overriding priority, nuclear plant safety issues receive the attention warranted by their significance' (Obadia et al., 2007). Since the 1990s safety culture then been seen as the key to improving safety (Antonsen et al., 2017) and an enormous amount of research has been conducted to try to precisely define the concept and create a fitting model for it (Guldenmund, 2010).

2.2 | The influence of safety culture on safety

Many researchers have found that safety culture has a positive impact on the safety of an organisation. For instance, Van Vuuren (2000) has found that the causes of incidents and accidents were largely due to organisational factors and that the safety culture counted for one-third of those organisational factors. The majority of these safety culture failures were related to the safety attitudes of people involved in the incidents and accidents. Sexton and Klinect (2017) have found that crews with positive perceptions of safety culture could catch more errors, perform better, make fewer violations and errors and have fewer undesired aircraft states than crews with negative perceptions of safety culture. Stolzer et al. (2011) have considered safety culture, when reinforced throughout the organisation, to be one of the most effective and systemic ways to reduce the levels of accident and incident within an organisation. Berglund (2020) has found that safety culture prevented the Fukushima incident from worsening. Rubin et al. (2020) have seen safety culture as a potential cause of risk-taking behaviour, where poor safety norms led to more risk-taking behaviour.

With the Swiss cheese model (Reason, 1997) in mind, Obadia et al. (2007) have considered safety culture to be an additional defence-in-depth in the organisation, contributing to the reduction of risks at each layer of defence and, therefore, reducing the chance of an accident in the system. Zohar (2000) has found an empirical relationship between safety climate measures and the occurrence of minor work injuries. In their study, (Cooper & Phillips, 2004) have found a link between safety climate and safety behaviour, which is probably based on perception of the importance of safety training, a factor in safety climate. Johnson (2007) has found that improving safety climate is likely to improve safe behaviour, although the results reflect a correlation and not causation.

Critiques of these findings have also been made. Recently, Antonsen (2009) has showed, based on an empirical analysis of the influence of safety culture, that it is hard to prove the relationship between safety culture and safety, but also that a proactive quantitative SCA has little predictive value regarding the occurrence of incidents and accidents. Overall, the relationship between safety culture and safety has met with considerable debate (see Henriqson et al., 2014). This research is focused on whether this concept has any predictive value in terms of safety outcomes.

2.3 | Modelling safety culture

Considering safety culture as a cause of accidents requires the ability to describe and model the concept. Safety culture can be seen as something that can become better or worse as opposed to being a static concept. This view results in the maturity model concept (Corrigan et al., 2020). Westrum (2004) has distinguished the pathological, bureaucratic and generative cultures, each with a characteristic response to problems, which can be predicted based on the way information flows through the organisation (Flannery, 2001; Westrum, 2004). To depict evolutionary progression in the development of safety culture, with a true safety culture at the end of that progression (Salas et al., 2001), Hudson has elaborated on Westrum's model (1991, cited in Salas et al., 2001) by adding two stages. Hudson also changed the name bureaucratic to calculative, creating a safety culture model consisting of five stages: pathological, reactive, calculative, proactive and generative (1999). Foster and Hoult (2013) in (Corrigan et al., 2020), using slightly more neutral terms, have identified five stages as vulnerable, reactive, compliant, proactive and resilient. This maturity model depicted shown in Figure 1.

2.4 | Safety culture maturity and safety

Safety culture maturity models centre on the notion that an organisation needs to develop a mature safety culture if it takes safety seriously (Hudson, 2001a). The characteristics of a mature or generative safety culture resemble those of high reliability organisations, which are known for their good safety performance records (Parker et al. 2006). To become a generative or true safety culture, an organisation must progress through the other, less mature stages

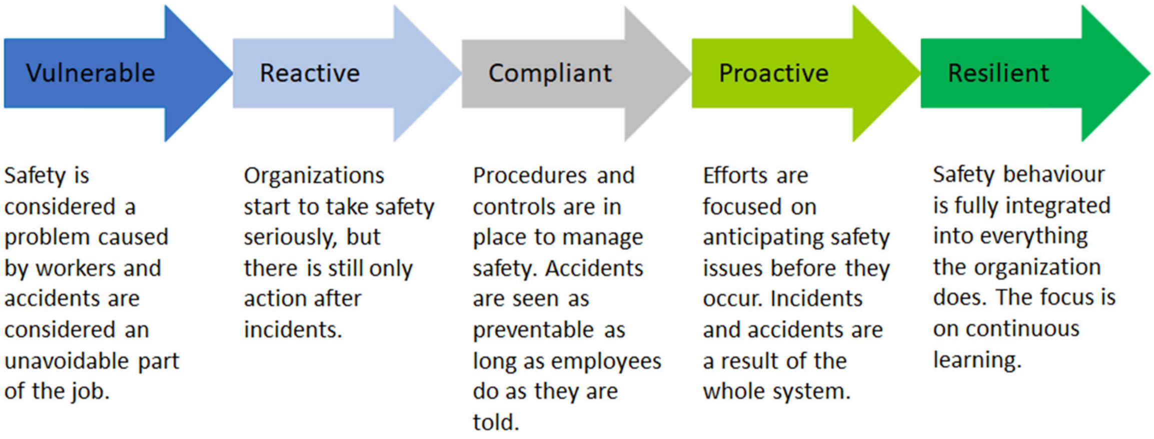

FIGURE 1 The maturity model of safety culture, based on (Salas et al., 2001), (Hudson, 2001b), (Hudson, 2001c), (Filho et al., 2010), (Johnsen et al., 2003) (Parker et al. 2006) (Yeoman, 2011)

(Hudson, 2001a). An organisation with an advanced, generative safety culture has fewer accidents compared to organisations with less mature safety cultures (Hudson, 2001c). A mature and effective generative safety culture can lead to substantial diminishment of the accident rate in aviation (Hudson, 2001a), whereas in organisations characterised by a pathological safety culture, the workplaces are dangerous (Hudson, 1999).

The literature on safety culture maturity claims a positive relationship between safety culture and safety, but that this relationship can be specified based on the maturity level of a given safety culture. The most mature level, generative, results in the safest performance and the least number of accidents in an organisation. However, Goncalves Filho and Waterson (2018) have discussed the strengths and weaknesses of safety culture maturity models, concluding that these models assume steady progress from one level to the next, therefore, seeing values and behaviours regarding safety as static, when in fact this may not be the case. If, however, these values and behaviours are dynamic and variable, the predictive value, that is the possibility to describe the safety state of an organisation, is lost.

In this study, the relationship between the safety culture maturity concept and safety is questioned. A safety culture maturity assessment should provide insight into the protection from accidents in an organisation. If an accident happens, the investigation of this occurrence should reveal aspects of safety culture that are still underdeveloped. These should be the same aspects as revealed by the SCA, if the latter is to be considered a stable description with predictive value regarding safety and accidents. This study focuses on the predictive value of such a SCA.

3 | METHOD

This research focused on a helicopter squadron of the RNLAF. In 2013, at the end of a training mission, a helicopter struck a light post with its rotor blades only weeks after this squadron had taken part in a proactive SCA. This accident was investigated by an accident investigation committee. The researchers were not part of this committee. The researchers had collected SCA data using a survey a few weeks before the accident and were able to compare this data to the results of the accident investigation.

For the SCA a survey was developed, based on the survey used in the Hearts and Minds programme and comparable surveys by Hudson (Parker et al. 2006; Salas et al., 2001). From the literature (Zohar, 1980) (Cox & Cox, 1991) (Williamson et al., 1997) (Cox & Flin, 1998) (Cox & Tomás, Cheyne, et al., 1998) (Lee, 1998) (Cheyne & Cox, 2000) (Ferraro, 2000) (Flin et al., 2000) (Glendon & Stanton, 2000) (Griffin & Neal, 2000) (Grote & Künzler, 2000) (Guldenmund, 2000) (Vuuren, 2000) (Gadd & Collins, 2002) (Sorensen, 2002) (Wiegmann et al., 2002) (Patankar, 2003) (Taylor & Thomas, 2003) (Dolfini-Reed & Streicher, 2004) (GAIN Working Group E, 2004) (Gordon & Kirwan, 2004) (Hamaideh, 2004) (Cai, 2005) (Fogarty, 2005) (Health & Safety Executive, 2005) (Falconer, 2006) (Parker et al. 2006) (Thaden & Mitchell-Gibbons, 2008) (Piers et al., 2009) (Fogarty & Shaw, 2010) 129 indicators were collected. A deduction exercise, using a resemblance strategy, resulted in the reduction of these 129 indicators into nine indicators. Each indicator contains at least two, but not more than 4 items to represent this indicator, resulting in a total of 28 items. For each item, the behaviour indicative of each maturity level was described resulting in five separate situational descriptions with the same subject but different behavioural wordings (Filho et al., 2010; Hudson, 2001c; Johnsen et al., 2003). Respondents chose the description that best resembled their current working environment. Within each set, the five descriptions were in a random order, so the respondents did not know which description fit which maturity level. The statements were constructed as neutrally as possible, so it would not be clear from the wordings which statement was 'best' and which 'worst'.

To confirm the five descriptions per item represent accurately the five maturity stages, four other researchers received 28 envelopes, each envelope containing five cards depicting one description per card. The researchers were asked to put the descriptions in order ranging from the least developed safety culture to the highest developed safety culture. The results were discussed and a few wordings in the SCA adapted.

Using descriptions instead of a Likert scale provides a better understanding of the meaning of an answer. With a questionnaire with statements and corresponding 5 or 7-point Likert scales there is a higher risk of interpretation flaws due to misunderstanding the statement and less detail in the description.

For the assessment, specifically the mode (the answer most frequently given by the respondents), was used. When describing a concept such as safety culture, the frequency of perceptions or opinions is the starting point, not a calculated mean. This concurs with the observation that culture emerges from group processes (Kramer, 2019) and the general agreement that culture is subsequently determined by the most shared ideas, values and perceptions found in the majority of the group. This is the case, unless, of course, there is a minority with substantial formal or informal power. In that case, it is the minority that has the largest influence on culture. However, this is not the case within the squadron studied, since it is a helicopter squadron in which the principles of crew resource management are vital for both operational effectiveness and physical safety (Salas et al., 2001).

The accident investigation reported on the rotor strike. This accident was chosen for this study because of the coincidence that it happened just a few weeks after the proactive SCA. There are, luckily, few accidents a year, which makes it more difficult to find appropriate data.

A content analysis of the accident report was performed and the findings determined by the accident investigation committee were compared and related to the indicators used in the survey to determine the similarities and the possible predictive value of a SCA. Each finding in the report was related separately to a maturity description of one of the items. Relating the findings to the indicators was carried out with a 'best fit' strategy based on the description (both wording and meaning) in the report and the description of the indicators. This process allowed each finding to be related to only one indicator. For instance, the accident investigation found that the crew deviates from standard radio calls and different acknowledgements are provided instead of the prescribed ones in the training manual. This relates to the rule-making indicator at the calculative level, which states: Noncompliance to rules and regulations is unacceptable, even if the relevance or appropriateness of the rules/regulations is doubted.

The investigation report showed no findings relating to the 8th and 9th indicators of the survey, role model behaviour of staff, and reinforcement or punishment. The focus of the accident investigation was on the crew and the movement of the helicopter at the time of the accident, not on commanding officers or the way safe behaviour was being reinforced. Therefore, for these indicators, a comparison cannot be made and the SCA results for these two indicators were omitted from this study.

The results of the SCA are presented, followed by the results of the occurrence investigation and, finally, the comparison between the two assessments is given.

4 | RESULTS OF THE SCA

In 2013, this squadron had 76 employees (pilots and crewmembers). Thirty-one employees returned a distributed survey, resulting in a response percentage of 41%. Although the response rate is not that high, if the intention of research is to gain in-depth knowledge about safety culture, a smaller number of respondents can still be relied on (Bergersen, 2003).

Of the 31 respondents, 11 are pilots-in-command, 9 are co-pilots and 11 are loadmasters. Regarding their position within the squadron, of the 31 respondents 2 are senior managers within the squadron, 9 middle management and the remaining 20 general operators. To guarantee the anonymity of the respondents, no other information regarding the respondents were used in this study.

Feedback provided by 10 respondents after returning the survey showed that filling out the survey took on average 40 min due to the fact that respondents had to read a total of 240 statements (28 times 5) and had to make 28 decisions. The returned surveys showed no pattern in the answers provided, as has been seen with 5-point Likert scale questionnaires where only the third/middle category is marked. Because no pattern was detected it is assumed the respondents filled out the survey in a serious manner.

To determine if the SCA was predictive of the squadron being incident- or accident-prone, the results of the survey were compared to the findings of the accident investigation. The results of the SCA are depicted in Table 1–7. The tables show the percentage of the respondents choosing the statement that most fitted their work environment, which was the squadron under investigation. These results show how the safety culture for the operational squadron was described. For each table, the following legend applies (Figure 2).

The results for the first indicator showed that crews were trusted when breaking the rules for safety reasons (respectively 63% and 45%), but that it was perceived that rules were also broken for operational advantage, indicating a difference between paper and practice (32%).

The results for the second indicator indicated respondents perceived empowerment by employees relating to safety (respectively 87% and 53%). Regarding risky training missions, 42% thought they could determine themselves whether to execute them. However, 26% thought that, although possible, this was a decision never made.

The results for the third indicator showed mixed results. Safety was seen as an inseparable part of every mission and, therefore, a priority (respectively 52%, 45% and 38%). The operational goals were considered more important and safety was usually triggered by unwanted outcomes, which diminished over time (26%, 23% and 19% respectively). Because of the rules and regulations, no trade off was necessary (19%).

A significant proportion of respondents perceived the sharing of mistakes and safety issues as part of their job (46%, 50%, 29% and 42% respectively). They felt they were encouraged in this behaviour and saw the same behaviour in their colleagues (27%, 40%, 48% and 32% respectively).

TABLE 1 Indicator 1 rule breaking

1a	Crews are trusted to recognise the situations in which compliance to regulations needs to be doubted. It is exected that crew initiatives are executed safely and that risks are mitigated by crew decision making.	A certain amount of incompliance with rules and regulations is accepted, as long as this is discussed up front.	Noncompliance to rules and regulations is unacceptable, even if the relevance or appropriateness of the rules/regulations is doubted.	It is possible to break the rules without anyone noticing, especially when more time has passed since the last incident.	Regarding following the rules the saying is: 'who cares, as long as were are not getting caught!'.
	63% generative	19% proactive	19% calculative	0 reactive	0 pathological
1b	Breaking the rules is only allowed for safety reasons. Crews are trusted to make the right judgment.	On paper rules are not meant to be broken, but in practice the case is different. Management does not always live up to its word when an operational advantage is to be gained.	There is no such thing as breaking the rules. Each decision is about risk control. The crew is trusted to act safely.	The most important thing is to conduct the mission, even if that means breaking rules or regulations. It is up to the crew to decide if they are capable or not.	It is never allowed to break the rules. The safety management system in place ensures safety.
	45% proactive	32% reactive	16% generative	6% pathological	0 calculative

TABLE 2 Indicator 2 empowerment of employees

2a	Every employee has the possibility and authority to make his/her own estimate of safety and act accordingly.	Personnel responsible for safety has the authority to make safety decisions.	Only very experienced crewmembers have limited possibilities to make decisions that affect the safety of normal flight missions.	Pilots and crewmembers have no authority to make decisions that affect the safety of normal flight missions.	The authority to make safety decisions is related to someones formal positions and part of regulations.
	87% generative	6% proactive	3% reactive	3% pathological	0 calculative
2b	Crews can decide for themselves whether or not to execute risky training missions and determine if the risks are mitigated to an acceptable level.	It is possible for employees to avoid risky training missions, however nobody does this. A mission that in hindsight should not have been executed does not result in consequences for the crew.	According to regulations it is not expected of employees to execute risky training missions.	According to regulations it is not expected of employees to execute risky training missions.	Even risky training missions are to be executed at all times. Flight performance is that high, this is not an issue. Refusal or avoidance is a sign of a lack of flying skills.
	42% generative	26% reactive	19% proactive	13% calculative	0 pathological
2c	Crewmembers are actively reminded of the possibility to intervene if they think safety is at risk.	When it comes to safety, crewmembers are equal. The authority gradient is hardly present. Copilots are actively encouraged to intervene if they deem this necessary.	Rules, such as the two-challenge-rule, offer copilots and crewmembers, including the inexperienced ones, the possibility to contribute to the safety of the mission.	Crewmembers should not question actions of the captain except when these actions threaten the safety of the flight.	Except for total incapacitation of the captain, the co-pilot should never assume command of the aircraft.
	53% proactive	22% generative	19% calculative	6% reactive	0 pathological

TABLE 3 Indicator 3 balance between mission and flight safety

3a	Safety is inseparably linked to every aspect of my job. I always make sure that my aircraft and crew are as safe as possible.	Right after an incident has occurred I pay a little more attention to safety when doing my job.	Where it is possible, safety is equally important as the result when doing my job.	Following the rules is the most important part of my job. Safety is an expected result of compliance.	Personally I feel that safety issues are not the most important part of my job.
	52% generative	26% reactive	16% proactive	3% calculative	3% pathological
3b	The organization tries to make safety a top priority by acknowledging that safety can contribute to the organizational goal (operational success). The organization is pretty good in combining safety with the organizational goal, although attaining operational success is still decisive.	Organizational goals are and remain the most important priorities, but effort is made to comply with legal norms and regulations.	Most of the available time is spend on achieving operational success. Safety is guaranteed through rules and regulations and requires less attention.	Safety is a core activity in this organization. Safety and operational success are valued equally. Safety is a prerequisite to achieve operational success.	Operational success is the number one priority. If this can be achieved safely it's a plus.
	45% proactive	23% reactive	19% calculative	13% generative	0 pathological
3c	Safety is part of every mission, which means that without an acceptable level of safety the mission cannot be succesful.	Following the rules results in safety. This means no trade off between safety and performing flight missions is necessary.	When more time has passed since the last incident, more risks are taken during training missions and safety is taken less into account.	If it all comes done to it, people in this organization prefer to take safety risks to cancelling training missions.	When executing training missions safety receives an ever greater priority.
	38% generative	19% calculative	19% reactive	16% pathological	9% proactive

In estimating their own capabilities, the respondents saw stress, fatigue, emergencies and personal problems as possible safety threats (23%, 16% and 17%) that needed to be addressed before a flight mission (55%, 62% and 74% respectively). Preparing for emergencies and monitoring one another were seen as ways to mitigate the risk (19% and 73%).

Concerning lessons learned, 39% indicated that nothing changes, no lessons are learned, and that only employees directly involved received feedback (22%). At the same time, trend analysis was conducted (29%) and incident and accident reports were spread to share lessons learned (44%).

As can be seen in Table 7, learning and possible improvement are the focus of incident and accident investigations (41% and 54%) and the entire organisation is investigated and benefits (25% and 20%). Accountability is placed upon involved operators and their supervisors (26%) and on the systems that failed (35%). While 91% of the respondents perceived a working environment in which incidents and accidents are seen as unpreventable, 65% of the respondents thought being wary helps and 26% thought acceptance is all that is left.

5 | ACCIDENT INVESTIGATION FINDINGS

The accident involved a helicopter that struck a light post while taxiing. The intention was to park the helicopter after a training mission. The designated parking spot, as indicated by a marshaller, did not offer enough room to manoeuvre the helicopter. The crew feared the rotor blades would come in contact with another helicopter already parked. It was decided to park next to the already parked helicopter, but on the other side. While performing a 270-degree turn, the light post was struck. The accident investigation used a systems approach, but no specific research method was applied. The investigation report provides the findings presented in Table 8 as causally linked to the accident, which can subsequently be related to safety culture weaknesses. Table 8 represents all the factual findings written down in the report. No selection was made nor findings left out of this study. Although hindsight bias is always a risk and it is not possible to determine potential hindsight bias with the investigating committee, by using the entire investigation report the risk of hindsight bias in this study is reduced to a minimum.

6 | COMPARING THE SCA TO THE ACCIDENT INVESTIGATION

Based on a content analysis and best fit strategy using resemblance, the findings in the investigation report were related to the indicators used in the survey and the maturity level according to models by Hudson (Salas et al., 2001) and Foster and Hoult (Foster & Hoult, 2013) were determined. The results are shown in Tables 9–15. Eleven findings in the investigation report could not be related to the indicators because they either displayed facts unrelated to the crew,

TABLE 4 Indicator 4 openness

4a	I am encouraged by my leadership and co-workers to report any unsafe conditions I may observe.	I am being encouraged to communicate my safety concerns to those who can do something about it.	We have a system in which I can report safety issues anonymously. All pilots and crewmembers are encouraged to use this system.	There is a way for me to address safety issues, however this is not encouraged.	Inexperienced pilots and crewmembers have no way to report safety issues.
	46% generative	27% proactive	13% calculative	13% reactive	0 pathological
4b	Sharing mistakes is SOP as far as I'm concerned. There is no shame or admiration for doing this openly.	I am actively encouraged to openly share my mistakes, for instance in articles in our safety magazine.	I'd prefer my other Flight crew/air crew/flight members not to know when I made a mistake. I talk about it as little as possible.	I am ashamed when I make a mistake in front of my other crewmembers. I will therefore not share my mistakes with others.	Mistakes can be shared by reporting them anonymously in a database, which I do.
	50% generative	40% proactive	7% reactive	3% pathological	0 calculative
4c	I have no problems with writing my name on a safety report after I was unintentionally involved in an incident or accident.	I always report any issues that I feel are a threat to safety.	I am willing to report below standard performance or unsafe acts of colleagues.	I am reluctant to report safety issues unless this can be done anonymously.	I am not prepared to report safety issues as I fear a negative response or personnal consequences.
	48% proactive	29% generative	10% calculative	6% reactive	6% pathological
4d	Every morning briefing starts with reporting safety issues in front of the group by the people that came across them.	My colleagues always tell me about any safety issues they have encountered.	Safety issues are only discussed by the people that came across them with the people who were a witness at that time.	Safety issues are shared through the monthly newsletter.	Safety issues are not being shared among colleagues.
	42% generative	32% proactive	23% reactive	3% calculative	0 pathological

TABLE 5 Indicator 5 perception of limitations

5a	When I am tired I prefer not to fly. If I have to I will use my colleagues as a safety net to compensate for the dimishing safety.	I am less effective when stressed or fatigued.	When I am tired I don't fly. The risk of fatigue is to big. There are no sufficient measures to be able to fly safely.	When I am tired, coffe or a powernap compensates for the lack of energy so that it doesn't influences my performance.	Even when fatigued, I perform effectively during critical times in a flight.
	55% proactive	23% calculative	10% generative	10% reactive	3% pathological
5b	During the briefing possible emergencies and solutions are adressed. What if's, alternates, etc. are discussed to prepare the crew in case an emergency happens.	The crew is always prepared for emergencies with checklists, what if's, alternates, etc. It is a given that under those circumstances the quality of decision making diminishes and possible options are revised continuously during the flight.	I am more likely to make judgment errors in an emergency.	My decision making ability is as good in emergencies as in routine flying conditions.	The decision making proces is suboptimal during emergencies, for which I can compensate with my own capacities.
	62% proactive	19% generative	9% calculative	6% pathological	3% reactive
5c	Since personnal problems can negatively affect performance, these deserve attention. It is key to discuss them up front so the rest of the crew can take this into account.	Personal problems can negatively affect my performance. But there are not many options to deal with this.	A truly professional crewmember can leave personal problems behind when flying.	Personnal problems are nonexistent to a professional crewmember. His or her private life has nothing to do with the job.	Having personnal problems is a risk that is hard to mitigate and incompatible with flying.
	74% proactive	16% calculative	6% reactive	3% pathological	0 generative
5d	Monitoring eachother for signs of stress or fatigue is a structural element of a pilot's or crewmember's task.	Crewmembers should monitor eachother for signs of stress or fatigue.	I should actively monitor my crewmembers for signs of stress or fatigue.	It is not necessary to monitor one another, since every pilot/ crewmember is capable of doing his/her job, including when under stress or experiencing fatigue.	When a pilot/crewmember seems stressed or fatigued during flight, it might be wise to monitor him/her.
	73% generative	17% calculative	7% proactive	3% pathological	0 reactive

TABLE 6 Indicator 6 lessons learned

6a	Nothing changes after a safety report has been written. There are no recommendations, feedback or actions taken.	The database containing the reported incidents and accidents is used for trend analysis. The results of this analysis are used to enhance safety.	Safety reports are stored in a database. On predetermined moments an analysis is made and spread throughout the organization as a form of feedback.	All safety reports are published. With each new report the possibility of a trend is researched. All levels of the organization actively use the the information from the safety reports in their daily work.	Possible recommendations in a safety report only refer to the specific incident investigated.
	39% pathological	29% proactive	13% calculative	10% generative	10% reactive
6b	Incident or accident reports are spread at the supervision level after completion to share the lessons learned with alle employees.	Only the employees involved in an incident/accident receive feedback after an investigation. Sometimes the occurrence is mentioned in the annual report.	After an investigation, no lessons learned are shared.	Alle employees are informed about the lessons learned of an investigation, which takes the form of new regulations and an article in a safety magazine.	All employees are personally informed of an investigation and the lessons learned as a result of that investigation.
	44% proactive	22% reactive	16% pathological	13% calculative	6% generative

such as the host nation, or content not covered by the indicators. These findings were, therefore, not compared.

Regarding the first indicator, rule breaking, the SCA results showed that the first item was perceived at the generative level (63%), the second item at the proactive (45%) and the reactive level (32%). The accident investigation, however, showed findings that all related to the calculative level, indicating that the results of the SCA do not match the results of the accident investigation. The SCA describes a safety culture that acknowledges that to be and stay safe, rules sometimes need to be broken. It describes a safety culture in which the pilots and crewmembers know the difference between situations in which rules can be broken for safety reasons and situations in which compliance with the rules is necessary. The accident investigation, however, showed that, in this case, the crew could not see the need to comply. The crew deviated a few times from rules and procedures, which, according to the investigation report, led to the accident.

Regarding the second indicator, empowerment of employees, SCA results showed the respondents perceived the first item at the generative level (87%), the second item at both a generative (42%) and reactive (26%) level (pertaining to the avoidance or risky missions, which was not the case in this investigated situation) and the third item at the proactive level (53%). The accident investigation showed all generative results except for the fourth result, which corresponds to the proactive level. It can be seen from the investigation results that the investigated behaviour showed empowerment in practice, as predicted in theory by respondents. This means that the behaviour shown matched the behaviour described in the SCA.

Regarding the third indicator, balance between mission and flight safety, the SCA results showed the respondents perceived the first item at the generative level (52%), the second item at the proactive (45%) and reactive (23%) level and the third item at the generative (38%) and calculative (19%) level. The accident investigation, however, showed results that corresponded with either the calculative or reactive level, indicating that the SCA showed a more positive perception of the way safety was incorporated in the daily work than the accident investigation. Although, in the SCA, safety was seen as inextricably linked to the mission by respectively 52%, 45% and 23%, the accident investigation showed different outcomes. The SCA and the accident investigation described different behaviour and, therefore, different safety cultures.

Regarding the fourth indicator, openness, the SCA results showed the respondents perceived the first item at both the generative and proactive level (46% and 27% respectively), the second item at the generative level (50%), the third item at the proactive (48%) and generative (29%) level and fourth item again at the generative and proactive level (42% and 32%, respectively). As concluded from the SCA, openness was part of their daily existence, according to the respondents. The results in the accident investigation corresponded with the pathological, calculative and proactive level. Although a positive climate resulting from good crew cooperation was noticed, the lack of communication was seen as a

TABLE 7 Indicator 7 cause of an accident/incident

7a	Accountability for an incident/accident is placed with the involved operators and their supervisors as well as the systems that failed.	Involved operators and their supervisors are seen as the cause of an incident/accident and are held accountable.	All levels in the organization are seen as possibly accountable. Management admits she should accept part of the accountability in case of an incident/accident.	I feel that no-one should ever be blamed for the outcome of their professional actions. Incidents/accidents are caused by the interactions in the entire system, both past and present. Management and employees share accountability.	When an incident or accident happens it is often exclusively due to the error or violation of a single person. This individual is held accountable.
	35% calculative	26% reactive	23% proactive	10% generative	6% pathological
7b	Incidents/accidents cannot be prevented. The organization is always wary. The chance of an incident/accident can be diminshed by making safety a part of every task executed.	Incidents/accidents are ‘part of the job' and cannot be prevented. They are a fact of life.	Incidents/accidents are the result of system failure and lack of (compliance with) regulation. They can be prevented by focussing on implementation of systems and enough regulation.	Incidents/accidents are often bad luck. They can be prevented by removing the involved operators.	All incidents/accidents can be prevented by incorporating proactively safety in management decision making and developing proactive measures.
	65% generative	26% pathological	6% calculative	3% reactive	0 proactive
7c	In case of an incident/accident, the investigation focusses on understanding and learning with the aim of prevention. The Why and How are more important than the Who or What.	When investigating an incident/accident, management looks at the entire system, including processes and procedures.	Investigation of an incident/accident focusses on who or what caused the occurrence (direct cause). The investigation is simple and factual.	In case of an incident/accident the investigation focusses on the failing system and rules that were broken	In case of an incident/accident the investigation focusses on finding the accountable employees.
	41% generative	25% proactive	19% reactive	13% calculative	3% pathological
7d	I see any errors that I or a colleague make as a way for the organisation to improve. The ones making a mistake, speak up about it voluntarily so everybody can learn from their experience.	When an incident or accident has occurred this eventually always benefits safety within my organisation.	incidents/accidents show the organization where regulations can be improved.	Incidents/accidents are an unwanted disturbances of daily operations that doe not benefit the organization whatsoever.	Incidents and accidents show which employees are incapable of their job.
	54% generative	20% proactive	20% calculative	6% pathological	0 reactive

Safety culture maturity levels				
pathological	reactive	calculative	proactive	generative

FIGURE 2 Legend accompanying Table 1–7

lack of openness and causally related to the accident. However, based on the investigation report, it is unclear whether the lack of communication was due to a lack of openness or simply the result of not seeing the need to communicate. Based on the description in the investigation report, the safety culture outlined in the investigation report does not match the safety culture detailed in the SCA.

Regarding the fifth indicator, perception of one's own limitations, the SCA results showed the respondents perceived the first three items at the proactive level (55%, 62% and 74%, respectively) and the fourth item at the generative level (73%), indicating that personal limitations needed to be addressed, mitigated and monitored. Of course, occurrences, such as being annoyed or distracted, are relatively common. The results of the SCA did not indicate that these emotions would no longer occur, but that they were acknowledged and mitigated to ensure a safe flight. The accident investigation showed nine findings that corresponded to the pathological level and two findings corresponding with the generative level. Before the flight, the crew had taken care of their physical well-being (rest and food). However, during the flight, the taxiing phase, the crew realised they were irritated and distracted, but did not counter these feelings, discuss them or otherwise mitigate them to ensure the safety of this phase. The investigation report showed they failed to understand the effects of their emotional well-being. The safety culture described by the accident investigation did not match the safety culture described by the SCA.

Regarding the sixth indicator, lessons learned, the SCA results showed the respondents perceived the first item at the pathological (39%) and proactive (29%) level and the second item at the proactive (44%) and reactive (22%) level. The accident investigation revealed only one finding corresponding to this indicator, which was at the reactive level. Although the SCA seemed to describe an organisation that learned from previous incidents and accidents, it also showed that it did not. The results of the SCA were ambiguous. The accident investigation concluded that the organisation did not learn from previous occurrences, given the similarities with a previous accident. The safety culture described by the accident investigation, therefore, only partly matched the safety culture described by the SCA.

Regarding the seventh indicator, cause of an accident or incident, the SCA results showed respondents perceived the first item to be at the calculative (35%) and reactive (26%) level, the second and fourth item at the generative level (65% and 54% respectively) and the third item at the generative (41%) and proactive (25%) level. The accident investigation showed findings corresponding to the reactive and calculative level. This indicates that the assessment showed a safety culture in which accidents and incidents were seen as emerging from the entire organisation and something the organisation could learn from, but that responsibility and accountability were often placed on the involved employees. The investigation report partly reflected the same safety culture. Regarding responsibility and accountability, the accident was seen as preventable had the involved operators been more careful.

Although this comparison seems to show differences and similarities between the two assessments (SCA and accident investigation), it is not clear how significant the differences are or how well the results match. The results of both assessments are, therefore, depicted in graphs, to study the differences further. For each graph, the horizontal axis depicts the number of items in the survey and the number of findings in the accident investigation. The vertical axis depicts the maturity level ranging from 1 to 5, with 1 indicating the pathological level and 5 the generative level. The orange bars represent the results from the SCA and the blue bars the results from the accident investigation. The individual survey items are not compared to the individual accident investigation findings, only the pattern found in the SCA is compared to the pattern found in the accident investigation.

When comparing the results of the SCA and the accident investigation, the same pattern is found in Indicators 3—balance between mission and flight safety, 4—openness, 6—lessons learned and 7—cause of an accident or incident. The following four graphs show this pattern.

Based on these graphs, the SCA shows a more positive, more mature safety culture for these four indicators than the accident investigation. These indicators relate to the role and behaviour of the operators. According to the SCA the operators working at the involved helicopter squadron maintained a good balance between flight mission and safety. The operators were open, learned from previous incidents or accidents and considered an incident or accident to be a result of the entire system, although they recognised that the operators involved were marked as responsible. The accident investigation, however, revealed the operators as the cause of the helicopter accident, supported by a lack of openness, not enough priority placed on safety and the fact that a similar accident happened 3 years previously. For these four indicators, the operators had a more positive perception of their safety culture than the organisational investigation committee.

The results for Indicator 1—rule breaking show the same pattern when looking at the content of the indicator.

Although the maturity levels suggest a linear development for the subject of rule breaking, it is more of a parabola development. The underlying theme regarding the breaking of rules is the amount of freedom one experiences. According to the maturity ladder for both the pathological and the generative level, operators can break the rules if they think they need to. In a pathological safety culture, this is possible because succeeding in the mission is the most

TABLE 8 investigation findings

Findings from the accident investigation
No technical shortcomings found, so the investigation focuses on preparation and execution of the flight mission.
Deviation in the flight plan; supposed to be a two-ship formation, with the other PIC as section lead.
During the preparation a three hour delay occurred due to technical issues with the helicopter. The crew brought food for a proper lunch; the delay did not result in fatigue.
The flight mission lasted for 2 h and 10 min and was executed without any remarkable events.
The flight mission was considered a routine flight. Crew had flown together multiple times and experienced good crew cooperation.
During the reconnaissance the crew has inspected the platform. The light posts were not mentioned by the host nation.
The reconnaissance checklist is very general, it doesn't prompt to inspect any specifics.
No written report was made of the reconnaissaince or the identified risks.
The light posts were never seen as a risk or danger to the flight mission.
The briefing by the host nation did not mention the light posts.
Only one marshaller was to assist the parking of the helicopter. During prior parking there were two marshallers.
Crew estimates that the parking spot directed by the marshaller is to small and poses to many risks. It is considered unsafe.
Based on the directions and estimates of the loadmaster, the crew decides to park south of the already parked other helicopter.
Before the training mission, the helicopter was also parked south of the already parked other helicopter.
Copilot's attention is drawn to the marshaller who reacts very irritated upon the decision to park south and leaves the platform.
Copilot gets irritated by the irritation of the marshaller.
Physical contact between crew (loadmaster) and marshaller could have resolved the mutual misunderstanding about the parking intentions and would therefore have prevented the accident.
Crew decides to continue to taxi and park without support of the marshaller.
Since a safe parking is the responsibility of the PIC, parking without the support of a marshaller is permitted.
Crew deviates right from taxi line to provide separation with the already parked helicopter.
Subsequentely, the crew deviates left from the taxi line to provide separation with the preflight tent of which the sides started to flap.
The number one loadmaster keeps his attention at preflight tent and a group of bystanders filming the helicopter at a distance of approximately 150 m.
When the number one loadmaster can no longer see the preflight tent he directs his attention to a fire extinguishers on the platform.
It is concluded that the focus of the crew was more on the consequences of the downwash during taxiing than on the dimensions of the rotor blades relating to the MOD.
The lesser focus on the dimensions of the rotor blades was a result of the fact that the week before taxiing on the platform was safe and uneventful.
At some point in the 270 degree turn the copilot notices the light post but does not consider it a risk.
Passing the light post on the right, the copilot decides not to mention it on the intercom to the rest of the crew, because it seems to fall outside the minimum obstacle distance (MOD)
The PIC notices the light post and decides to taxi straight on to create more distance with the edge of the platform.
There is no communication within the crew regarding the decision to taxi straight on.
At the edge of the platform the PIC decides to deviate left of the taxi line to create space for the tail of the helicopter. The focus is on the edge of the platform, not on the light post.
The number one loadmaster has not seen the light post.
The PIC is relatively unexperienced, less than 60 flight hours as a PIC.
The number 2 loadmaster is relatively unexperienced, just over 90 flight hours of experience.
All crewmembers were fit to fly. They had had enough rest and no personal problems.
The RNLAF has to be compliant with the militaire aviation requirements.
RNLAF rules and regulations ensure this compliance.
Every employee is expected to know the relevant rules and regulations, so that they will be compliant.
During training of pilots and loadmasters enough attention is given to the subject of distance between rotor blades and obstacles.
The crew was distracted due to the vehicle parked behind the already parked helicopter, the behaviour of the marshaller, the tent on the platform and the group of bystanders.
As a result of the cognitive distraction the situation awareness of the crew dimished. Human attention span is limited.
Based on the lookout sections of the PIC, COP and number one loadmaster the light post could have been identified as a risk to safety had they realised that the light post was within the MOD.
According to the training manual a distance of a minimum of 5 m. has to be guaranteed.
Based on the CVR it is concluded that the crew deviates from standard radio calls. Different acknowledgements are provided instead of the prescribed ones in the training manual.
A turn to the left is executed without clearance from the loadmaster.
In flight training issues such as platforms that deviate from minimum separation distances are not addressed, which is considerd a training deficiency.
The crew experienced a false sense of safety while taxiing on a platform that does not entail the minimum separation distances for helicopters.
A similar accident, in which a helicopter collided with a container while taxiing, happened three years before this accident.

TABLE 9 comparision occurrence results to SCS indicator 1 rule breaking

Findings accident investigation	Indicator	Level
The RNLAF has to be compliant with the militaire aviation requirements.	Rule breaking	Calculative
RNLAF rules and regulations ensure this compliance.	Rule breaking	Calculative
Every employee is expected to know the relevant rules and regulations, so that they will be compliant.	Rule breaking	Calculative
According to the training manual a distance of a minimum of 5 m. has to be guaranteed.	Rule breaking	Calculative
Based on the CVR it is concluded that the crew deviates from standard radio calls. Different acknowledgements are provided instead of the prescribed ones in the training manual.	Rule breaking	Calculative
A turn to the left is executed without clearance from the loadmaster.	Rule breaking	Calculative

TABLE 10 comparision occurrence results to SCS indicator 2 empowerment of employees

Findings accident investigation	Indicator	Level
Crew estimates that the parking spot directed by the marshaller is to small and poses to many risks. It is considered unsafe.	Empowerment of employees	Generative
Based on the directions and estimates of the loadmaster, the crew decides to park south of the already parked other helicopter.	Empowerment of employees	Generative
Crew decides to continue to taxi and park without support of the marshaller.	Empowerment of employees	Generative
Since a safe parking is the responsibility of the PIC, parking without the support of a marshaller is permitted.	Empowerment of employees	Proactive
Crew deviates right from taxi line to provide separation with the already parked helicopter.	Empowerment of employees	Generative
Subsequentely, the crew deviates left from the taxi line to provide separation with the preflight tent of which the sides started to flap.	Empowerment of employees	Generative
At some point in the 270 degree turn the copilot notices the light post but does not consider it a risk.	Empowerment of employees	Generative
The PIC notices the light post and decides to taxi straight on to create more distance with the edge of the platform.	Empowerment of employees	Generative
At the edge of the platform the PIC decides to deviate left of the taxi line to create space for the tail of the helicopter. The focus is on the edge of the platform, not on the light post.	Empowerment of employees	Generative

TABLE 11 comparision occurrence results to SCS indicator 3 balance between mission and flight safety

Findings accident investigation	Indicator	Level
During the reconnaissance the crew has inspected the platform. The light posts were not mentioned by the host nation.	Balance mision and flight safety	Calculative
No written report was made of the reconnaissance or the identified risks.	Balance mision and flight safety	Calculative
The light posts were never seen as a risk or danger to the flight mission.	Balance mision and flight safety	Reactive

TABLE 12 comparision occurrence results to SCS indicator 4 openness

Findings accident investigation	Indicator	Level
The flight mission was considered a routine flight. Crew had flown together multiple times and experienced good crew cooperation.	Openness	Proactive
Passing the light post on the right, the copilot decides not to mention it on the intercom to the rest of the crew, because it seems to fall outside the minimum obstacle distance (MOD)	Openness	Calculative
There is no communication within the crew regarding the decision to taxi straight on.	Openness	Pathological

TABLE 13 comparision occurrence results to SCS indicator 5 perception of one's own limitations

Findings accident investigation	Indicator	Level
During the preparation a 3 h delay occurred due to technical issues with the helicopter. The crew brought food for a proper lunch; the delay did not result in fatigue.	Perception of one's own limitations	Generative
Copilot's attention is drawn to the marshaller who reacts very irritated upon the decision to park south and leaves the platform.	Perception of one's own limitations	Pathological
Copilot gets irritated by the irritation of the marshaller.	Perception of one's own limitations	Pathological
The number one loadmaster keeps his attention at preflight tent and a group of bystanders filming the helicopter at a distance of approximately 150 m.	Perception of one's own limitations	Pathological
When the number one loadmaster can no longer see the preflight tent he directs his attention to a fire extinguishers on the platform.	Perception of one's own limitations	Pathological
It is concluded that the focus of the crew was more on the consequences of the downwash during taxiing than on the dimensions of the rotor blades relating to the MOD.	Perception of one's own limitations	Pathological
The lesser focus on the dimensions of the rotor blades was a result of the fact that the week before taxiing on the platform was safe and uneventful.	Perception of one's own limitations	Pathological
All crewmembers were fit to fly. They had had enough rest and no personal problems.	Perception of one's own limitations	Generative
The crew was distracted due to the vehicle parked behind the already parked helicopter, the behaviour of the marshaller, the tent on the platform and the group of bystanders.	Perception of one's own limitations	Pathological
As a result of the cognitive distraction the situation awareness of the crew dimished. Human attention span is limited.	Perception of one's own limitations	Pathological
The crew experienced a false sense of safety while taxiing on a platform that does not entail the minimum separation distances for helicopters.	Perception of one's own limitations	Pathological

TABLE 14 comparision occurrence results to SCS indicator 6 lessons learned

Findings accident investigation	Indicator	Level
A similar accident, in which a helicopter collided with a container while taxiing, happened three years before this accident.	Lessons learned	Reactive

TABLE 15 comparision occurrence results to SCS indicator 7 cause of an accident/incident

Findings accident investigation	Indicator	Level
No technical shortcomings found, so the investigation focuses on preparation and execution of the flight mission.	Cause of an accident/ incident	Reactive
Physical contact between crew (loadmaster) and marshaller could have resolved the mutual misunderstanding about the parking intentions and would therefore have prevented the accident.	Cause of an accident/ incident	Calculative
Based on the lookout sections of the PIC, COP and number one loadmaster the light post could have been identified as a risk to safety had they realised that the light post was within the MOD.	Cause of an accident/ incident	Calculative

important goal and breaking the rules is allowed to achieve this goal. In a generative safety culture, this is possible because operators are trusted to recognise when they need to comply and when they need to deviate regarding safety and risk control. In both the pathological and generative description, operators have considerable freedom regarding breaking the rules. This leads to a different scoring, with the pathological and generative levels receiving similar codes, as do the reactive and proactive levels, based on the amount of freedom. This line of thinking results in the pathological and generative levels being represented by a 5, the reactive and proactive levels by a 3 and the calculative levels by a 1. The numbers 5, 3 and 1 are arbitrary, chosen to show the difference more clearly than in the case of the numbers 3, 2 and 1. In this parabola graph, the numbers 1–5 on the horizontal axis represent the five maturity levels, which are pathological (1), reactive (2), calculative (3), proactive (4) and generative (5). The vertical axis represents the scoring of the levels. This shows the parabola effect of this particular indicator.

Translating the parabola effect to the SCA results and investigation findings results in a different depiction of the original graph of indicator 1.

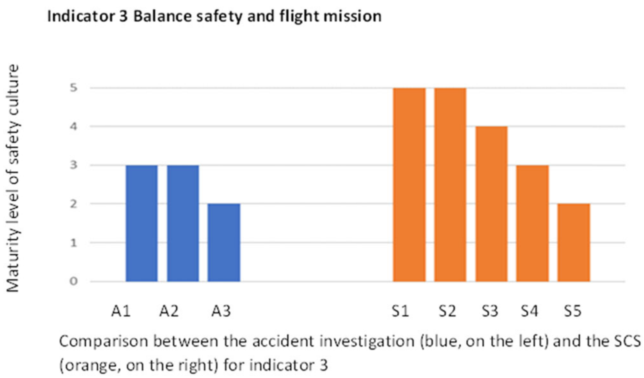

FIGURE 3 Indicator 3 balance safety and flight mission

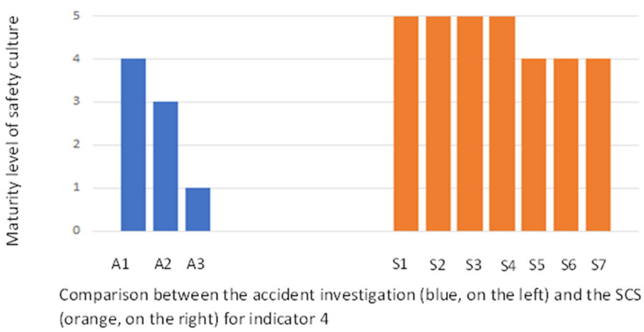

FIGURE 4 Indicator 4 openness

FIGURE 5 Indicator 6 lessons learned

Taking into account the parabola effect, it shows, as with Indicators 3, 4, 6 and 7, that for Indicator 1 the operators perceived their safety culture as more positive and more mature than in the investigation committee findings. According to the investigation committee, based on the military flight regulations, which are also part of the safety culture, operators may not break the rules. According to the operators, however, they work in a safety culture that does allow them to break the rules when necessary.

Indicator 2—empowerment of employees seems to show a different pattern, in which the accident investigation reveals a more positive and more mature safety culture than the SCA.

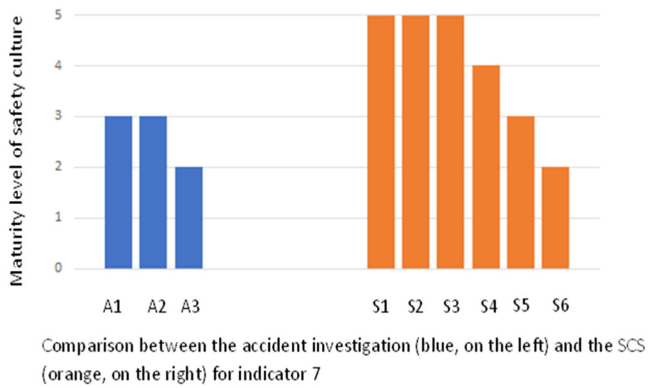

FIGURE 6 Indicator 7 cause of incident/accident

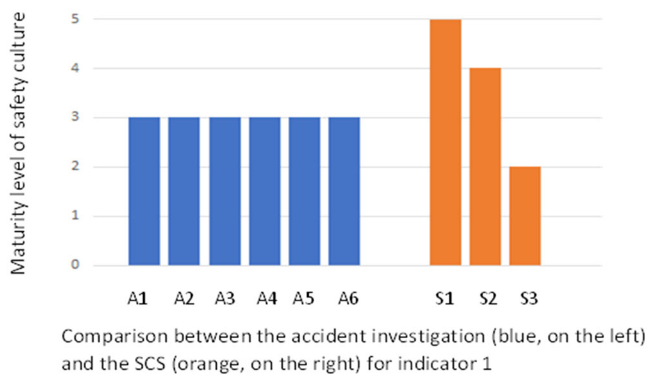

FIGURE 7 Indicator 1 rule breaking

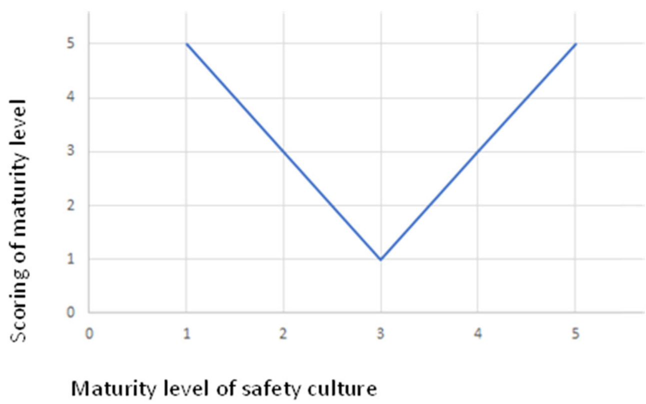

FIGURE 8 Scoring of indicator 1 rule breaking

What leads to this pattern is the perception of operators, seen in the SCA, that risky training missions can be avoided but would not be, indicating a focus more on mission and goal completion than on safety. In practice, however, as seen by the accident investigation, parking at the parking spot indicated by the marshaller, the final part of the mission, was refused in favour of a parking spot deemed to be safer. The operators did not predict their own behaviour regarding the use of their judgement to make a decision. For the investigation results, a relationship can be seen with Indicator 7—cause of an incident or accident. According to the investigation, the cause of this accident was the involved crew, considering the decisions they made, which they should not have. In the investigation too much empowerment (freedom) was found, resulting in the accident.

The fifth indicator, perception of limitations, also seems to convey a mixed message.

When looking closer at the content of both the SCA and the accident investigation, it becomes clear that the operators perceive their safety culture as more positive and mature than was found during the accident investigation. The difference between the two was significant. According to the investigation, operators showed a mature safety culture when preparing for a flight, as evidenced in their buying lunch and being well rested, but that, during the flight, the crew displayed behaviour resembling a rather immature safety culture.

Looking at the seven indicators together, the operators considered their safety culture rather mature and indicated that they played an active role in achieving a safe flight. The investigation showed a more immature safety culture in which the behaviour of the operators led to a lack of safety rather than safety. The described safety cultures do not match and, therefore, the SCA seems to have limited predictive value regarding the situation found in the accident investigation.

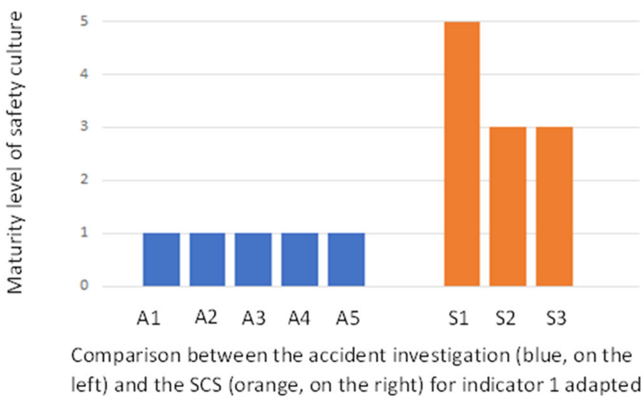

FIGURE 9 Indicator 1 rule breaking adapted

7 | DISCUSSION

The findings of this study show, as did the research by Antonsen (2009), that the results of the accident investigation do not match the results of the SCA. The comparison of the SCA with the accident investigation shows that the two different methods describe two different safety cultures, adding to the literature by demonstrating the limitative predictive value of a SCA.

Based on the SCA, it seemed very unlikely an accident would happen that was caused by factors, such as empowerment or

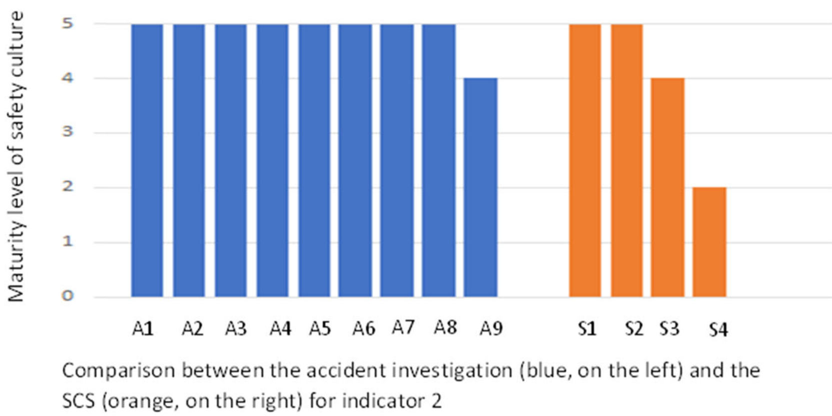

FIGURE 10 Indicator 2 empowerment of employees

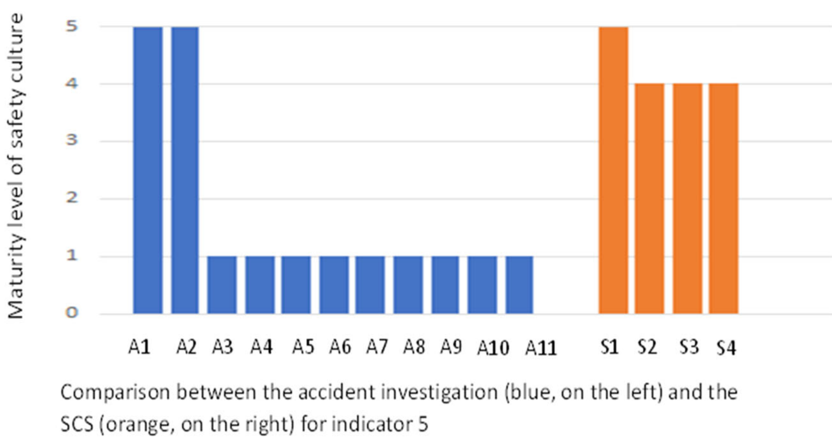

FIGURE 11 Indicator 5 perception of limitations

perception of limitations. Yet this is exactly what was determined by the accident investigation committee. The occurrence of an accident is not remarkable. There is always a chance an accident happens. However, according to the theoretical framework, the chance of an accident is diminished in the mature safety culture which the SCA described (Hudson, 1999, 2001c). What is remarkable, however, is that the accident investigation determined factors reflecting a mature safety culture, such as empowerment, as the cause of this specific accident. This means that the SCA has only limited predictive value. A safety culture maturity description of a workplace or an organisation does not provide protection from accidents. It might in fact provide an illusion, even leading to the idea of a 'check in the box' once a SCA shows an organisation has reached the proactive or even the generative maturity level.

This finding is supported by the fact that the SCA and the accident investigation were only a few weeks apart. The SCA was conducted in September. The accident happened in October of the same year. No significant changes occurred within the squadron organisation during that time. Safety culture is seen as a derivative of culture (Guldenmund, 2010), which is seen as a stable concept (Schein, 2004). It is, therefore, highly unlikely that the safety culture of the squadron changed between the assessment and the accident.

Perhaps the difference in results when comparing the two measurements was caused by the instruments used. However, the SCA is based on a validated and well-used instrument (Energy Institute, 2015, 2019). The difference between the original instrument and the SCA is in the wordings used, matching the RNLAF organisation. The underlying theoretical framework is the same. The accident investigation was the official instrument to conduct an accident investigation, based on and driven by international norms, agreements and standards applicable to air safety investigations (International Civil Aviation Organization, 1993). The investigation committee followed these and internal guidelines and procedures regarding the process of accident investigation.

A limitation in this study was the response rate of 41% of the SCA, although previous research (Bergersen, 2003) shows that the number of respondents is adequate for scientific analysis. Furthermore, the balance between the number of pilots-in-command and co-pilots within the response group resembles the balance in the squadron. Only the loadmasters are a little under represented. The balance regarding position (senior management, middle management and operators) again resembles the squadron. These representations add to the generalizability of the findings of the SCA. Another limitation could be the content of the SCA. The SCA is based on a literature study of literature on safety culture indicators, which can never be complete, inevitably resulting in questions not asked. This research has been set as much as possible in the context of available scholarship, as described in the method section.

The question remains of what to do with two conflicting descriptions of the supposedly same safety culture. According to Antonsen (2009), it makes sense to consider the description based on the accident investigation as most accurate, since investigations can be seen as more thorough and are considered more important in the organisation. The reason for this preference is that an organisation is highly motivated to find the cause of an incident or accident to prevent future ones. That motivation is lacking when 'simply' assessing the safety culture when nothing is out of the ordinary. When conducting an accident investigation, the investigators describe the behaviour of others. In principle, they have nothing to gain by describing this behaviour more positively or negatively than observed. The SCA asks respondents to describe their workplace and their own behaviour in it, as do most SCAs. It is possible that the respondents have a more positive perception of this than is supported by reality. Unfortunately, the choice to have respondents describe their safety culture instead of having an external observer make the evaluation falls outside the scope of this study.

The choice to consider the squadron safety culture description based on the accident investigation as the most accurate leads to the conclusion that the predictive value of the SCA is limited but that it inversely predicted the safety outcomes. Aspects of the safety culture that, according to the SCA, are very mature (proactive and generative) were identified as the causes of this accident.

The instruments used, the SCA and the accident investigation, have a different focus. The SCA focuses, as do all safety culture maturity assessment instruments, on the way safety is created or enhanced (Parker et al. 2006). The results show the perceived safety culture in the absence of an accident. The results of the accident investigation show the accident investigators' perception of that same safety culture in the presence of an accident, that is, in the absence of safety, according to the accident investigators.

Perhaps it is this difference in research focus or aim that makes it impossible to compare the content and even find an inverse prediction. If prediction is what an organisation aspires to, it might be more fruitful to design an instrument that measures safety and to subsequently test this instrument for predictive value not only in situations characterised by a lack of safety but also in situations where safety was actually achieved. A safety culture maturity assessment apparently cannot accomplish that.

This is supported by the literature on safety culture as a complex and dynamic concept which assessments using surveys are unable to convey (Guldenmund, 2007) and a quantitative approach is limited when the aim is prevention (Guldenmund, 2000). Safety culture is continually changing as a result of the dynamic social reality surrounding it (Richter & Koch, 2004) (Gephart et al., 2009) (Silbey, 2009). Furthermore, accident and incidents cannot be seen as a simple indicator of safety culture (Cooper, 2000) (Richter & Koch, 2004).

Based on this, one could ask the question if there is an instrument that does adequately provide insight into the safety culture of an organization, be it an assessment or an accident investigation tool. As the literature shows, as mentioned above, questionnaires and assessments do not do the complex phenomenon justice. Neither do most accident investigation methods as they focus on what went wrong, which is usually the exception (Hollnagel, 2014). What could be done however, is to study the way people do their work. Within a complex setting, operators are faced with multiple goals that are often not compatible with one another (Woods et al., 2010). Operators find ways to deal with these goal conflicts by employing what is called local ingenuity (Boskeljon-Horst et al., 2022): routines that help resolve goal conflicts and become part of the regular repertoire of operators, not necessarily contrary to the literal wording or intent of rules and procedures, but were not originally intended and not included in current documentation. How local ingenuity is employed to achieve safety, one of goals to be achieved, adds to the understanding of the workings of a complex system operators are part of and provides more insight than a SCA or an accident investigation.

8 | CONCLUSION

The research aim in this article was to determine whether the SCA possesses predictive value regarding how accident-prone an organisation is. In the case under study, the organisation was a helicopter squadron. However, the official instrument to conduct an accident investigation had different results and a different safety culture description than the safety culture maturity assessment. The results of the comparison show that the SCA has limited predictive value for the safety culture during an accident but that it inversely predicts the safety outcome. Factors indicative of a mature, proactive and generative, safety culture were found to be the causes of this particular accident. The findings by Antonsen (2009), based on the comparison of a quantitative SCA to an accident report, were replicated, this time using a SCA with detailed workplace descriptions. This research reinforces the conclusion Antonsen draws regarding the limited predictive value of safety culture maturity assessments. Development of a means to predict safety outcomes is planned using micro experiments within a helicopter squadron. These results are discussed elsewhere Figures 3–11.

AUTHOR CONTRIBUTIONS

Conceptualization: Leonie Boskeljon-Horst and Sydney W. A. Dekker. *Methodology*: Leonie Boskeljon-Horst and Sydney W. A. Dekker. *Formal analysis*: Leonie Boskeljon-Horst. *Investigation*: Leonie Boskeljon-Horst. *Writing—original draft preparation*: Leonie Boskeljon-Horst and Sydney W. A. Dekker. *Writing—review and editing*: Leonie Boskeljon-Horst, Sydney W. A. Dekker, and Simone Sillem. *Visualization*: Leonie Boskeljon-Horst. *Supervision*: Sydney W. A. Dekker and Simone Sillem. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGEMENTS

The researchers would like to gratefully acknowledge the contributions of various subject-matter experts at the helicopter squadron, particularly the relevant pilot in command, as well as LtCol. Michel de Rivecourt, LtCol. Arthur Rademaker and Dr. Eva van Baarle.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data supporting the findings of the safety culture assessment are in encompassed in the article. Availability of data that support the findings of the accident investigation requires a request to the Dutch Department of Defense.

ORCID

Leonie Boskeljon-Horst http://orcid.org/0000-0001-5760-6191
Simone Sillem http://orcid.org/0000-0001-5599-8603

REFERENCES

Antonsen, S. (2009). Safety culture assessment: A mission impossible. *Journal of Contingencies and Crisis Management*, *17*(4), 242–254. https://doi.org/10.1111/j.1468-5973.2009.00585.x

Antonsen, S., Nilsen, M., & Almklov, P. G. (2017). Regulating the intangible. Searching for safety culture in the Norwegian petroleum industry. *Safety Science*, *92*, 232–240. https://doi.org/10.1016/j.ssci.2016.10.013

Bergersen, C. (2003). *Tool to be used to survey and improve safety culture in the European railway industry*. Unpublished Master's Thesis, Norwegian University of …. http://scholar.google.com/scholar?hl=en%26btnG=Search%26q=intitle:Tool%2BTo%2BBe%2BUsed%2BTo%2BSurvey%2Band%2BImprove%2BSafety%2BCulture%2BIn%2BThe%2BEuropean%2BRailway%2BIndustry#0

Berglund, J. (2020). After fukushima: safety culture and fostering critical thinking. *Safety Science*, *124*(November 2019), 104613. https://doi.org/10.1016/j.ssci.2020.104613

Boskeljon-Horst, L., De Boer, R. J., Sillem, S., & Dekker, S. W. A. (2022). Goal conflicts, classical management and constructivism: How operators get things done. *Safety*, *8*(37), 37.

Cai, W. (2005). *The impact of safety culture on performance: A case study of a construction company*. Indiana University.

Cooper, M. D. & Phillips, R. A. (2004). Exploratory analysis of the safety climate and safety behavior relationship. *Journal of Safety Research*, *35*, 497–512. https://doi.org/10.1016/j.jsr.2004.08.004

Cooper, M. D. (2000). Towards a model of safety culture. *Safety Science*, *36*(2), 111–136. https://doi.org/10.1016/S0925-7535(00)00035-7

Corrigan, S., Kay, A., Ryan, M., Brazil, B., & Ward, M. E. (2020). Human factors & safety culture: Challenges & opportunities for the port environment. *Safety Science*, *125*, 103854. https://doi.org/10.1016/j.ssci.2018.02.030

Cox, S. & Cox, T. (1991). The structure of employee attitudes to safety: A European example. *Work & Stress*, *5*(2), 93–106.

Cox, S. & Flin, R. (1998). Safety culture: Philosopher's stone or man of straw. *Work & Stress*, *12*(3), 189–201. https://doi.org/10.1080/02678379808256861

Cox, S. J. & Cheyne, A. J. T. (2000). Assessing safety culture in offshore environments. *Safety Science*, *34*, 111–129.

Cox, S. J., Tomás, J. M., Cheyne, A., & Oliver, A. (1998). Safety culture: The prediction of commitment to safety in the manufacturing industry. *British Journal of Management*, *9*(SPEC. ISS.), 3–11. https://doi.org/10.1111/1467-8551.9.s1.2

Dolfini-Reed, M. A. & Streicher, B. L. (2004). Creating a safety culture. *CRM D001056.A2/1Rev*, 1–57. https://doi.org/10.21236/ada605037

Energy Institute. (2015). *Improving safety culture through Hearts and Minds. September*. https://minerva.jrc.ec.europa.eu/EN/content/minerva/7592dbf6-056b-46e6-96c0-2e580b9c3b69/201509mjv_safety_culture_nlkinglawrie

Energy Institute. (2019). *Understanding your HSE culture. Hearts & Minds*. https://publishing.energyinst.org/heartsandminds/toolkit/UYC

Falconer, B. T. (2006). Attitudes to safety and organisational culture in Australian military aviation. *Aviation*, 337–350. http://scholar.google.com.au/scholar?start=60&hl=en&num=30&cites=2395082783304815858#12

Ferraro, L. (2000). *Measuring safety climate: The implications for safety performance* (p. 1217).

Filho, A. P. G., Andrade, J. C. S., & Marinho, M. M. O. (2010). A safety culture maturity model for petrochemical companies in Brazil. *Safety Science*, *48*(5), 615–624. https://doi.org/10.1016/j.ssci.2010.01.012

Flannery, J. A. (2001). Safety culture and its measurement in aviation. *Transport*, *36*(25), 1–29. http://www.mendeley.com/research/safety-culture-measurement-aviation-1/

Flin, R., Mearns, K., O'Connor, P., & Bryden, R. (2000). Measuring safety climate: Identifying the common features. *Safety Science*, *34*(1–3), 177–192. https://doi.org/10.1016/S0925-7535(00)00012-6

Fogarty, G. J. (2005). Psychological strain mediates the impact of safety climate on maintenance errors. *International Journal of Applied Aviation Studies*, *5*(1), 53–63.

Fogarty, G. J. & Shaw, A. (2010). Safety climate and the theory of planned behavior: Towards the prediction of unsafe behavior. *Accident Analysis & Prevention*, *42*(5), 1455–1459. https://doi.org/10.1016/j.aap.2009.08.008

Foster, P. & Hoult, S. (2013). The safety journey: Using a safety maturity model for safety planning and assurance in the UK coal mining industry. *Minerals*, *3*, 59–72. https://doi.org/10.3390/min3010059

Gadd, S. & Collins, A. M. (2002). *Safety culture: A review of the literature, Report No. HSL/2002/25*.

GAIN Working Group E. (2004). *A roadmap to a just culture: Enhancing the safety environment*.

Gephart, R. P., Van Maanen, J., & Oberlechner, T. (2009). Organizations and risk in late modernity. *Organization Studies*, *30*(2–3), 141–155. https://doi.org/10.1177/0170840608101474

Glendon, A. I. & Stanton, N. A. (2000). Perspectives on safety culture. *Safety Science*, *34*, 193–214.

Goncalves Filho, A. P. & Waterson, P. (2018). Maturity models and safety culture: A critical review. *Safety Science*, *105*(June), 192–211. https://doi.org/10.1016/j.ssci.2018.02.017

Gordon, R. & Kirwan, B. (2004). Developing a safety culture in a research and development environment: Air Traffic Management domain. *Europe Chapter of the Human Factor and Ergonomic Society Conference*, *129*(August 2014), 2865. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.79.9628&rep=rep1&type=pdf

Griffin, M. A. & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology*, *5*(3), 347–358. https://doi.org/10.1037/1076-8998.5.3.347

Grote, G. & Künzler, C. (2000). Diagnosis of safety culture in safety management audits. *Safety Science*, *34*(1–3), 131–150. https://doi.org/10.1016/S0925-7535(00)00010-2

Guldenmund, F. W. (2000). The nature of safety culture: A review of theory and research. *Safety Science*, *34*(1–3), 215–257. www.elsevier.com/locate/ssci

Guldenmund, F. W. (2007). The use of questionnaires in safety culture research—an evaluation. *Safety Science*, *45*(6), 723–743. https://doi.org/10.1016/j.ssci.2007.04.006

Guldenmund, F. W. (2010). *Understanding and exploring safety culture*. BOXPress.

Hamaideh, S. (2004). *Safety culture instrument: A psychometric evaluation*. Jordan University of Science and Technology.

Health & Safety Executive. (2005). A review of safety culture and safety climate literature for the development of the safety culture inspection toolkit. Research Report 367. Health (San Francisco).

Henriqson, É., Schuler, B., van Winsen, R., & Dekker, S. W. A. (2014). The constitution and effects of safety culture as an object in the discourse of accident prevention: A Foucauldian approach. *Safety Science*, *70*, 465–476. https://doi.org/10.1016/j.ssci.2014.07.004

Hollnagel, E. (2014). *Safety-I and Safety-II The Past and Future of Safety Management. Taylor & Francis Group*.

Hudson, P. (1999). *Safety culture—Theory and practice*. 1–2.

Hudson, P. (2001a). Aviation safety culture. *Safeskies*, 1–23.

Hudson, P. (2001b). Safety culture: The ultimate goal. *Flight Safety Australia*, 29–31. http://82.94.179.196/bookshelf/books/1091.pdf

Hudson, P. (2001c). Safety management and safety culture the long, hard and winding road. *Occupational Health & Safety Management Systems Proceedings of the First National Conference*, 3. http://www.ohs.com.au/ohsms-publication.pdf#page=11

INSAG. (1991). *Safety series no.no. 75-Insag-4 International atomic energy agency, Vienna, 1991 categories in the iaea safety series. In Safety Series, (Issue 75).*

International Civil Aviation Organization. (1993). *Investigation of human factors in accidents and incidents. Human Factors Digest #7*.

Johnsen, S. O., Vatn, J., Jersin, E., Veiseth, M., Rosness, R., Lamvik, G., Steiro, T., Hagen, O., & Herrera, I. (2003). *Review of existing knowledge applicable to safety culture at interfaces in European Railway undertakings*. SINTEF Report.

Johnson, S. E. (2007). The predictive validity of safety climate. *Journal of Safety Research*, *38*(5), 511–521. https://doi.org/10.1016/j.jsr.2007.07.001

Kramer, J. (2019). *Deep Democracy. Jitske Kramer & Boom uitgevers Amsterdam.*

Lee, T. (1998). Assessment of safety culture at a nuclear reprocessing plant. *Work & Stress*, *12*(3), 217–237.

Obadia, I. J., Vidal, M. C. R., & e Melo, P. F. F. F. (2007). An adaptive management system for hazardous technology organizations. *Safety Science*, *45*(3), 373–396. https://doi.org/10.1016/j.ssci.2006.07.002

Parker, D., Lawrie, M., & Hudson, P. (2006). A framework for understanding the development of organisational safety culture. *Safety Science*, *44*(6), 551–562. https://doi.org/10.1016/j.ssci.2005.10.004

Patankar, M. S. (2003). A study of safety culture at an aviation organisation. *International Journal of Applied Aviation Studies*, *3*(1), 243–258.

Patankar, M. S., Brown, J. P., Sabin, E. J., & Bigda-Peyton, T. G. (2012). *Safety Culture; Building and Sustaining a Cultural Change in Aviation and Healthcare*. Ashgate Publishing Limited.

Piers, M., Montijn, C., & Balk, A. (2009). Safety management system and safety culture working group (SMS WG) Safety culture framework for the ecast SMS-WG. *ECAST: European Strategic Safety Initiatives, March*, 1–14.

Reason, J. (1997). *Managing the Risks of Organizational Accidents*. Ashgate.

Richter, A. & Koch, C. (2004). Integration, differentiation and ambiguity in safety cultures. *Safety Science*, *42*(8), 703–722. https://doi.org/10.1016/j.ssci.2003.12.003

Rubin, M., Giacomini, A., Allen, R., Turner, R., & Kelly, B. (2020). Identifying safety culture and safety climate variables that predict reported risk-taking among Australian coal miners: An exploratory longitudinal study. *Safety Science*, *123*(December 2019), 104564. https://doi.org/10.1016/j.ssci.2019.104564

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does crew resource management (CRM) training work? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *43*(4), 641–674.

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

Sexton, J. B., & Klinect, J. R. (2017). The link between safety attitudes and observed performance in flight operations. *Human Error in Aviation*, *1996*, 513–518. https://doi.org/10.4324/9781315092898-25

Silbey, S. S. (2009). Taming prometheus: Talk about safety and culture. *Annual Review of Sociology*, *35*, 341–369. https://doi.org/10.1146/annurev.soc.34.040507.134707

Sorensen, J. N. (2002). Safety culture: A survey of the state-of-the-art, *76*, 189–204.

Stolzer, A. J., Halford, C. D., & Goglia, J. J. (2011). *Implementing safety management systems in aviation*. Ashgate.

Taylor, J. C. & Thomas III, R. L. (2003). Toward measuring safety culture in aviation maintenance: The structure of trust and professionalism. *The International Journal of Aviation Psychology*, *13*(4), 321–343. https://doi.org/10.1207/S15327108IJAP1304_01

Thaden, T. L. V. & Mitchell-Gibbons, A. (2008, July). *The safety culture indicator scale measurement system (SCISMS). Human Factors Division Institute of Aviation. Technical Report HFD-08-03/FAA-08-2, 121*.

van Vuuren, W. (2000). Cultural influences on risks and risk management: Six case studies. *Safety Science*, *34*, 31–45. https://docs.google.com/a/wmu.se/file/d/0B2lfx_f4ofuXNzlSNlQ2UXV4MGs/edit

Westrum, R. (2004). A typology of organisational cultures. *Quality and Safety in Health Care*, *13*(Suppl 2), 22–27. https://doi.org/10.1136/qshc.2003.009522

Wiegmann, D. A., Zhang, H., Thaden, T., Von Sharma, G., & Mitchell Gibbons, A. (2002, June). *A synthesis of safety culture and safety climate research. Aviation Research Lab, Institute of Aviation. Technical Report ARL-02-3/FAA-02-2*. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Modality+Differences+in+Advanced+Cockpit+Displays:+Comparing+Auditory+Vision+and+Redundancy+for+Navigational+Communications+and+Traffic+Awareness#0

Wiegmann, D. A., Zhang, H., Von Thaden, T. L., Sharma, G., & Gibbons, A. M. (2004). Safety culture: An integrative review. *The International Journal of Aviation Psychology*, *14*(2), 117–134. https://doi.org/10.1207/s15327108ijap1402_1

Williamson, A. M., Feyer, A. M., Cairns, D., & Biancotti, D. (1997). The development of a measure of safety climate: The role of safety perceptions and attitudes. *Safety Science*, *25*(1), 15–27.

Woods, D. D., Dekker, S. W. A., Cook, R., Johannesen, L., & Sarter, N. (2010). *Behind human error*. Ashgate Publishing Limited.

Yeoman, C. (2011). *What is safety culture?* https://Sites.Google.Com/Site/Highleyyeoman/Papers-Articles/What-Is-Safety-Culture.

Zohar, D. (1980). Safety climate in industrial organizations: Theoretical and applied implications. *Journal of Applied Psychology*, *65*(1), 96–102. https://doi.org/10.1037/0021-9010.65.1.96

Zohar, D. (2000). A group level-model of safety climate: Testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology*, *85*(4), 587–596.

How to cite this article: Boskeljon-Horst, L., Sillem, S., & Dekker, S. W. A. (2022). 'Ladder'-based safety culture assessments inversely predict safety outcomes. *Journal of Contingencies and Crisis Management*, 1–20. https://doi.org/10.1111/1468-5973.12445