ARTICLE IN PRESS

Journal of Safety Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Safety Research

journal homepage: www.elsevier.com/locate/jsr

The path toward successful safety performance measurement

Aki Jääskeläinen *, Sari Tappura, Julius Pirhonen

Industrial Engineering and Management, Tampere University, Korkeakoulunkatu 8, P.O. Box 541, 33101 Tampere, Finland

ARTICLE INFO

Article history:
Received 3 September 2021
Received in revised form 21 October 2021
Accepted 23 August 2022
Available online xxxx

Keywords:
Occupational Health and Safety (OHS)
Performance measurement
Performance management
Safety management
Survey study

ABSTRACT

Introduction: Safety management is widely seen as a key contributor to occupational health and safety (OHS) performance. Performance measurement is an important tool for management in reaching its goals. Safety performance measurement has gained increasing attention in the literature. However, little is known so far of the path towards successful safety performance measurement resulting in better OHS performance. *Methods:* This study analyzes the maturity of safety performance measurement in relation to OHS performance and the role of employee commitment and practices of using performance information in facilitating the performance benefits. The empirical data were gathered with a survey that received 270 responses from five industrial organizations. Partial least squares structural equation modeling (PLS-SEM) was used to analyze the data obtained. *Results:* It is found that commitment to performance measurement is the strongest explaining factor of both supervisor and employee OHS performance, while the maturity of performance measurement has a direct effect on supervisor safety performance only. *Practical Applications:* The results show how safety performance measurement can be implemented to derive the potential benefits. While managers may benefit directly from advanced performance measures, the wider performance benefits among employees materialize only by using performance measurement properly and committing employees to it.

1. Introduction

It is widely recognized that capable safety management can lead to great financial benefits by reducing the costs related to accidents (Tappura, Sievänen, Heikkilä, Jussila, & Nenonen, 2015). Safety management can also improve organizations' productivity and economic results (O'Toole, 2002), and even lead to a competitive advantage (Fernández-Muñiz, Montes-Peón, & Vázquez-Ordás, 2009). This study sheds light on how safety performance measurement and management activities relate to safety performance. The term safety performance is used to refer specifically to occupational health and safety (OHS) performance, defined as the measurable results concerning an organization's management of its OHS risks (BS 4:1800, 2008). OHS performance is measured by evaluating supervisor and employee performance.

Performance measurement may be crucial for several reasons: (a) organizations can evaluate their current performance and monitor their development from the past; (2) organizations can control their development to reach set objectives and targets; (c) organizations can budget for their plans; (d) organizations can motivate staff, managers, and other stakeholders and use the information in rewarding; (e) organizations can use their performance information to promote themselves to customers, investors, and other stakeholders and to benchmark themselves against competitors; (f) organizations can improve continuously by using the information and learning from the past (Behn, 2003; Lebas, 1995). However, the benefits of performance measurement are multifaceted and a subject of academic controversy (Bourne, Kennerley, & Franco-Santos, 2005; Franco-Santos, Lucianetti, & Bourne, 2012).

Much of the literature on safety performance measurement still focuses on technical issues in measurement, definitions for measures, and, most notably. the types of measures such as the difference between leading and lagging measures (Koivupalo & Reiman, 2017; Podgórski, 2015; Sheehan, Donohue, Shea, Cooper, & Cieri, 2016). Although the general performance measurement literature has already shifted its focus from technical issues in measurement into the use of measurement information (Franco-Santos et al., 2012; Henri, 2006; Nudurupati, Bititci, Kumar, & Chan, 2011), the implementation and use of performance measurement has gained only limited attention in safety research. It is evident that only the actual use of information can lead to improved organizational performance (Bititci et al., 2011; Franco-Santos et al., 2012).

OHS practitioners seem to agree that their organizations' success in health, safety, and environmental (HSE) concerns are linked

* Corresponding author.
E-mail addresses: aki.jaaskelainen@tuni.fi (A. Jääskeläinen), sari.tappura@tuni.fi (S. Tappura), julius.pirhonen@tuni.fi (J. Pirhonen).

https://doi.org/10.1016/j.jsr.2022.08.014
0022-4375/

Please cite this article as: A. Jääskeläinen, S. Tappura and J. Pirhonen, The path toward successful safety performance measurement, Journal of Safety Research, https://doi.org/10.1016/j.jsr.2022.08.014

to how well they track, manage, and use the information provided by leading metrics (Sinelnikov, Inouye, & Kerper, 2015). Firms often have a lot of safety performance information. However, it is widely known that transforming data into action is a real problem for many organizations and their leading metrics for safety may be data-collecting machines rather than continuous improvement metrics created based on specific OHS-related action to measure (Sinelnikov et al., 2015).

It appears that safety performance measurement literature is ahead of practice. Organizations are still mostly focusing more on "lagging" indicators (Jääskeläinen, Tappura, & Pirhonen, 2020; Reiman & Pietikäinen, 2012) measuring events from the past, like reports of accidents. In this case, organizations are measuring the absence of safety rather than its presence, which for purposes of achieving improvements is not ideal. A low accident rate does not assure that risks are under control, especially in companies where the probability for accidents is low but major hazards are present (Arezes & Sérgio Miguel, 2003). When used correctly, "leading" indicators, like unsafe behavior reporting, enable organizations to identify and correct deficiencies before they cause any injuries or damage and promote a culture of prevention (Sinelnikov et al., 2015; Zwetsloot, Leka, Kines, & Jain, 2020). Hence it could be assumed that performance measurement may have safety performance implications if used appropriately.

Safety climate and culture are important viewpoints to safety performance measurement. Safety culture can be defined as the attitudes, beliefs, perceptions, and values that employees and managers share in relation to safety within an organization or workplace (Cox & Cox, 1991). Safety culture is expressed through the organizational climate, while the climate can be taken to mean the manifestation of culture within the organization (Guldenmund, 2000). Safety climate could be defined as the surface features of an organization's underlying safety culture. This is discerned from the employees' attitudes and perceptions at a given point in time, that is, a snapshot of the state of safety (Cox & Flin, 1998; Flin, 2003). Safety climate scores can be regarded as essential safety performance indicators as such (Hoffmeister et al., 2014; Zohar, 2010). There is a plethora of instruments available for measuring safety climate and culture, but they have rarely been validated in different contexts (Glendon, 2008). In addition, they have not been linked to the maturity of using performance information in safety management.

Studies show that safety culture is linked to safety performance (e.g., Lee, 1998; Carder & Ragan, 2003; Vinodkumar & Bhasi, 2009; Stemn, Bofinger, Cliff, & Hassall, 2019). Also, positive safety climate perceptions and organizational attitudes have been associated with better self-reported physical and mental health (Haslam, O'Hara, Kazi, Twumasi, & Haslam, 2016). However, Koivupalo and Reiman (2017) state that the relationship between safety culture and safety performance is heavily dependent on how and when both culture and safety performance are measured. Due to the supposed dependence, models for measuring safety culture have emerged steadily in the last two decades. There are some indications that performance measurement relates to safety performance (Stemn et al., 2019) but the relationships between safety performance measurement, supportive practices, safety climate, and safety performance are still mostly unclear.

This study specifies safety culture as the commitment of management and employees to safety performance measurement and management. The use of performance information in safety management is still inefficient and commitment to performance measurement insufficient (Jääskeläinen et al., 2020). Hence, it appears that performance measurement is not enough if the practices of using performance and supportive commitment are not in place. The link between performance measurement and safety performance is supposedly indirect. The existing research does not fully capture this complex connection between safety performance measurement and safety performance. Furthermore, the emphasis in the safety management literature has focused on detecting different components affecting safety and these links are not widely demonstrated (Zohar, 2010) This paper addresses the gap by researching how performance measurement and management practices affect safety performance among both supervisors and employees.

The empirical material was gathered with a survey from five industrial organizations. The survey was sent to 725 respondents and it gathered 270 responses. Partial least squares structural equation modeling (PLS-SEM) was used to determine the links between performance measurement and safety performance. The results show that commitment to performance measurement is the strongest explaining factor of both supervisor and employee safety performance, while the status of performance measurement only has a direct effect on supervisor safety performance. The study also reveals that the use of performance measurement fully mediates the link between performance measurement status and employee safety performance.

2. Literature review

2.1. Key concepts

2.1.1. Safety and safety performance

Safety is considered to be freedom from unacceptable risk or harm (ISO/IEC, 2004). In this paper, safety is seen as a synonym for occupational health and safety (OHS). OHS can be defined as conditions and factors that affect, or could affect, the health and safety of employees and any other person in the workplace (BS 4:1800, 2008; ISO 45001:2018). OHS is usually managed with an occupational health and safety management system (OHSMS), which is a part of an organization's management system used to develop and implement its OHS policy and manage its OHS risks. It is a set of interrelated elements and includes organizational structure, planning activities, responsibilities, practices, procedures, and resources. (ISO 45001:2018).

OHSMS is not a well-defined management system and there are no clear boundaries between OHS activities, OHS management, and OHSMS (Nielsen, 2000). OHSMS can be seen as a management system to improve the organization's OHS performance, focusing not only on operations and general physical work but also on employees' health and safety (ISO 45001:2018).

Performance is a measurable result of the management of activities, processes, products (including services), systems, or the organization. It may relate to either quantitative or qualitative findings. (IWA 26:2017). Tangen (2005) defines performance as the umbrella term of excellence, which includes profitability and productivity as well as other non-cost factors such as quality, speed, delivery, and flexibility.

Safety performance can be seen as a subsystem of organizational performance (Wu, Chen, & Li, 2008). Safety performance relates to the effectiveness of the prevention of injury and ill health and the provision of safe and healthy workplaces (ISO 45001:2018). Safety performance refers to the measurable results of an organization's management of its OHS risks, which includes measuring the effectiveness of its controls and the provision of safe and healthy workplaces (BS 4:1800, 2008; ISO 45001:2018). It could be also seen as OHS-related actions and behaviors that employees exhibit to endorse the health and safety in the workplace, such as safety compliance and safety participation (Burke & Signal, 2010; Neal, Griffin, & Hart, 2000). Good safety performance impacts, for example, on efficiency through reduced accident costs or improved productivity (Sievänen, Nenonen, &

Hämäläinen, 2013; Tappura et al., 2015) Safety performance is often measured by OHS objectives and organizational indicators such as safety climate, absenteeism, ill health, and injury rate (Hale, Guldenmund, Van Loenhout, & Oh, 2010).

2.1.2. Performance measurement

According to Neely (1998, p. 5) performance measurement is a process of quantifying the efficiency or effectiveness of a past action. Hannula and Lönnqvist (2002, p. 47) state that performance measurement is a process used to determine the status of an attribute relevant to the performance of the measurement object. Lemieux-Charles et al. (2003) define performance measurement as monitoring, which shows where change is required, and which will, in turn, produce the desired behavior that will produce improved performance. The different definitions highlight different things, which shows that performance measurement can be used in many ways.

Since performance measurement is a broad concept, there is no single correct way to measure performance. There are some commonly used models measuring performance, such as the Balanced Scorecard (Kaplan & Norton, 1996), the Performance Pyramid (Lynch, 1995), and the Performance Prism (Neely, Adams, & Kennerley, 2002). What these have in common is that performance measurement is related to the organization's strategy and can thus be used as a tool to implement the organization's strategy and to communicate strategic information from management to employees. Another unifying factor is the "balanced" approach of all of these, which means measuring performance from many different aspects, considering economic, employees', and customers' views.

2.2. Safety performance and its antecedents

Safety climate and leadership are significant facilitators of safety performance (Clarke, 2013; Wu et al., 2008). Safety leadership has received a lot of attention in the field of safety and studies have identified it as a key aspect in safety climate (e.g., Barling, Loughlin, & Kelloway, 2002; Zohar, 2010). Safety leadership can be described as "the process of interaction between leaders and followers, through which leaders could exert their influence on followers to achieve organizational safety goals under the circumstances of organizational and individual factors" (Wu, 2005). Management commitment is highly important in gathering occupational health and safety data, role model type of behavior, and promoting occupational health and safety throughout the organization (Lingard, Wakefield, & Cashin, 2011; Zohar, 2010). A study by Aksorn and Hadikusumo (2008) revealed that "management support" was the most significant factor in safety program implementation.

For its part, safety climate, referring to employees', managers', and directors' commitment and perceptions regarding safety, has attracted even more attention (Wu et al., 2008). Safety culture is another commonly used term in relation to safety climate. These terms are sometimes confused and often used synonymously, but there are slight differences in the meanings (see e.g. Cooper, 2000; Flin, Mearns, O'Connor, & Bryden, 2000).

It is clear that safety culture and climate are correlated with safety performance (e.g., Lee, 1998; Carder & Ragan, 2003; Vinodkumar & Bhasi, 2009; Stemn et al., 2019), similarly to safety leadership (Clarke, 2013; Wu et al., 2008). A longitudinal study by McCabe, Alderman, Chen, Hyatt, and Shahi (2017) concluded that safety climate accounted for up to 20% of the variance in safety outcomes. Another study by Lingard, Cooke, and Blismas (2012) revealed that supervisors' safety perceptions mediate the relationship between safety climate and injury rates. There are also more specific studies showing that constant promotion by top management is a key to the success of occupational safety interventions (Hale et al., 2010; Saksvik, Nytrø, Dahl-Jørgensen, & Mikkelsen, 2002). There is still a lack of research on identifying the specific mechanisms with which leaders promote a better safety climate in risky operations (Zohar, 2010) and thus better overall safety performance.

There are several factors other than safety leadership and climate that have been positively related to safety performance. For example, appropriate safety education and training, teamwork, clear and realistic goals, positive group norms, personal competency, safety equipment, and communication of safety have been described as critical factors affecting safety performance (Aksorn & Hadikusumo, 2008; Sawacha, Naoum, & Fong, 1999). Nonetheless, many of these factors have been included in safety culture/climate assessments.

Safety performance measurement also has a role in safety performance, especially in improving it. Leading indicators have been proven to predict safety performance. For example, Salas and Hallowell (2016) measured the frequency of supplier internal safety audits and the results predicted injury rates later in the same project. Studies suggest that the best leading indicators to predict future safety performance are safety observation and review and pre-task safety meetings (Rajendran, 2013; Salas & Hallowell, 2016).

2.3. Safety performance measurement

Maturity of performance measurement consists of factors, such as scope, sophistication and relevance of measurement (Jääskeläinen & Roitto, 2015). Scope can be used to refer to the balance of measures between financial and non-financial or leading and lagging measures (Cocca & Alberti, 2010; Van Aken, Letens, Coleman, Farris, & Van Goubergen, 2005). Sophistication may include, for example, the ability of measurement to explain relationships between measurement objects and information system sophistication (Cocca & Alberti, 2010; Marx, Wortmann, & Mayer, 2012). Relevance may be related to up-to-date measures and reliability of information provided by measures (Wettstein & Kueng, 2002). The prior discussion on safety performance measurement has especially paid attention to scope (e.g., balance between leading and lagging indicators) and relevance (e.g., high amount of performance information) and is briefly reviewed next.

Safety performance measurement can offer knowledge and help in internal analysis and decision-making (Arezes & Sérgio Miguel, 2003). Most of the measurements in organizations focus on achieving positive safety outcomes. For example, when addressing economic performance, positive aspects such as liquid profit, ROI, or market share are evaluated, but when safety performance is measured, the measurable variables are often negative, like total injuries or lost workday rates (Arezes & Sérgio Miguel, 2003). In this case, organizations are measuring the absence of safety rather than its presence, which gives a unique character to safety performance measurement and illustrates its complexity.

Organizations typically have a lot of performance information related to safety. Safety indicators are usually divided into leading (also known as proactive or predictive) and lagging (also known as reactive, trailing, or outcome) indicators. Lagging indicators are the most commonly used indicators, which have a retrospective focus on the reduction of workplace injuries (Kaassis & Badri, 2018). Leading indicators are less used, but they can provide advanced warning of potential problems and therefore the possibility of implementing preventive measures before mishaps occur (Kaassis & Badri, 2018). Zwetsloot et al. (2020) recently presented a further classification, proactive leading indicators, to better support the development of a prevention culture.

Blair and O'Toole (2010) suggest that the ratio between leading and lagging indicators should be about 80:20 or place even greater emphasis on leading indicators. Regarding putting the weight on leading indicators, it is argued that with a heavy focus on these the results will reflect on the outcome indicators as lower injury rates and workers' compensation costs. The leading indicators appear to be in their potential to predict and prevent adverse outcomes, but they are also intended to help the transformation of an organization's culture from passive and problem-focused to a more proactive and solution-driven (Sinelnikov et al., 2015). Lagging indicators still cannot be omitted from measurement practices since they yield valuable information about the development of safety performance and about the relationships between leading and lagging indicators. Moreover, many researchers do not distinguish between leading and lagging indicators, but use more general terms such as key indicator, safety performance indicator, or key performance indicators (Swuste, Theunissen, Schmitz, Reniers, & Blokland, 2016).

Another common way to measure organizational safety performance involves assessing workers' and managers' attitudes to and perceptions of safety. These attitudes and perceptions are usually measured with a safety climate survey. As stated earlier, safety climate/culture has been proved to correlate with safety performance, which has led to the point that safety climate/culture scores are commonly used as leading indicators. In a study by Hoffmeister et al. (2014), safety climate scores were considered the most important safety performance indicators. The state of safety culture has focused heavily on the results of safety culture maturity models. The most common aim of these models has been general safety management assessment, followed by an evaluation of the communication of safety and managerial commitment to safety (Goncalves Filho & Waterson, 2018).

Measuring safety culture or climate is a potential approach to improve safety performance. In a Nordic study, Kines et al. (2011) developed and empirically tested a questionnaire for measuring safety climate status in different countries and industries. It included seven dimensions related to both management (e.g., commitment, empowerment) and workers (e.g., commitment, trust). Stemn et al. (2019) found a strong negative correlation between the incident rate and most elements of the safety culture maturity framework. Carder and Ragan (2003) successfully used their survey tool for safety measurement to improve safety performance in several companies. Safety climate measures can also provide in-depth information about the root causes of OHS problems and can be used as a useful diagnostic tool (Lingard et al., 2011). In a study by Lingard et al. (2011) both safety climate survey and leading indicators were used successfully to identify weaknesses and to develop practical solutions to issues that would not even have been identified if the measurement had relied solely on the use of traditional lagging measures.

Safety culture as a concept remains a problem. Its definition entails many difficulties, likewise the decision as to what should be measured as the consequence of its presence or absence; either incidents or other intermediate measures of safety (Waterson, Jenkins, Salmon, & Underwood, 2017). Swuste et al. (2016) state that organizations with good scores on occupational safety use more complex indicators, but they hardly use the information to improve the organization. Safety indicators seem to be based mainly on company experience or common sense (Swuste et al., 2016). Using sources to pick indicators might be difficult since some references present more than 400 safety indicators (Amir-Heidari, Maknoon, Taheri, & Bazyari, 2017). This leads to a situation where safety indicators are more data-collecting machines than continuous improvement metrics created on the basis of specific OHS-related actions (Sinelnikov et al., 2015).

Increasing information processed by managers and decision-makers may also negatively influence the quality of their decisions (Hwang & Lin, 1999; Iselin, 1988). Decision-making will improve with a certain amount of information, but decision-makers will face an 'information overload' and information processing will decrease resulting in worse decisions when the amount of information exceeds a certain point. Since managers like to have as much information as possible, it is important to aggregate and summarize the information to keep the number of information dimensions to a minimum (Hwang & Lin, 1999).

It is generally known that performance management systems significantly affect people's behavior, organizational capabilities, and performance. There is evidence that performance measurement systems play a key role in strategy, communication, and management processes, generating capabilities that enable organizations to excel (Franco-Santos et al., 2012). Performance measurement systems alone are still not enough, and the information must be used, and used correctly.

A study by Henri (2006) strongly suggests that only interactive use of performance measurement fosters capabilities of market orientation, entrepreneurship, innovativeness, and organizational learning, while diagnostic use of performance measurement seems to contribute negatively to the deployment of these capabilities. Diagnostic use represents tight control of operations and strategies, and highly structured channels of communication and restricted flow of information. Interactive use, on the other hand, represents loose and informal control, and open channels of communication and free flow of information throughout the organization. Another study shows that, contrary to expectations, the higher performing organizations usually place less emphasis on performance management than the lower performing cluster (Bititci et al., 2011). These high performing organizations associate managerial activities more with culture and communication processes (Bititci et al., 2011).

Performance measurement can be also divided into feedback and feed-forward measures. The purpose of feedback controlling is to exploit the firm's existing capabilities and the purpose of feed-forward controlling is to identify new capabilities. The use of these different measures is typically tied to how these measures are used in the evaluation of managers (Grafton et al., 2010). To conclude, according to these sources the most sophisticated way to use performance measurement for organizational learning could be using feed-forward measures in an interactive manner. In safety management, the equivalent to interactive use of feedforward measures could be using leading indicators in an interactive way.

There is no consensus in the literature as to which are the most important metrics for safety measurement and how to utilize the measurement data effectively. Safety culture maturity models are still the most common way to evaluate the state of safety, but often these models do not accurately capture the practices of performance measurement. Hence, the understanding on the status of performance measurement in relation to safety performance is limited. A recent study by Alruqi and Hallowell (2019) identified nine common leading indicators that correlated significantly with worksite injuries: safety record, safety resource, staffing for safety, owner involvement, safety training/orientation, personal protective equipment, safety incentives program, safety inspections and observation, and pre-task safety meeting. Stemn et al. (2019) found that one of the strongest negative correlations between evaluated maturity of practices and incident rate was in the item named performance measurement with a correlation coefficient of almost −1. The topic of safety performance measurement has clearly come in for more attention lately, but the understanding of its desirable characteristics, supportive factors and relationship to safety performance measurement continues to be limited.

3. Hypotheses

Performance measurement offers no benefits if the information gathered is not utilized. To derive all the benefits of performance measurement and to use it efficiently, the organization must be committed to it. Tung, Baird, and Schoch (2011) state that top management's continued involvement and support are crucial for performance measurement to be effective. Therefore, top management need to personally commit to performance measurement and guarantee that sufficient resources are provided on a continuous basis to properly developing and managing the existing performance measurement systems. A study by Cavalluzzo and Ittner (2004) shows that management's commitment to performance measurement has a significant positive influence on the use of performance measurement. These findings support earlier research by Shields (1995), who argues that top management's support is crucial to new measurement system implementation because managers can allocate resources to initiatives they deem worthwhile and withhold resources from ideas they do not support. A study by Jääskeläinen and Sillanpää (2013) demonstrated the importance of operative level commitment in a performance measurement system implementation. Committed employees successfully implemented and used the new performance measures even though using it required more resource investments than the previous system. All in all, commitment to performance measurement by both operative employees and management is recognized to be a critical part performance measurement development (e.g., Cocca & Alberti, 2010). Thus, the first hypothesis is proposed.

H1 Commitment to performance measurement relates positively to the use of safety performance measurement

Maturity of performance measurement may explain performance differences between organizations (Franco-Santos et al., 2012). Many studies highlight the practices of performance measurement, but as Kaplan and Norton (2001) state, "It's not just what is measured, but how the measurements are used that determines the organizational success." According to Amaratunga and Baldry (2002), measurement is important because it provides the basis for an organization to assess how well it is progressing towards its objectives, helps to identify strengths and weaknesses, and decides future initiatives, but it does not tell why something happened or what to do about it. Organizations must make the transition from measurement to management to use the performance measurement effectively to reach the goal of improving organizational performance (Amaratunga & Baldry, 2002). There is evidence that organizations that do not integrate ongoing performance measurement and feedback into their management development programs tend to experience lower than expected performance improvements and greater dissatisfaction of their management (Longenecker & Fink, 2001). A study by Henri (2006) supports these results and strongly suggests that interactive use of performance measurement systems positively influences organizational capabilities, such as organizational learning. A study by Grafton et al. (2010) corroborates the findings of Henri (2006) and suggests that the use of performance measures impacts on the strategic capabilities of the organization and subsequently its performance.

H2 The use of performance measurement mediates the relationship between the status of performance measurement practices and safety performance: a) supervisor performance, b) employee performance

There is ample evidence that being committed to something relates positively to the results to which the commitment is made. A study by Wu et al. (2008) shows that safety climate, which includes employee and management commitment to safety, is positively related to safety performance. Psychology research, on the other hand, shows that goals are significantly more likely to lead to performance gains if individuals are committed to achieving them (Klein, Wesson, Hollenbeck, & Alge, 1999). Accounting research, consistent with the results from the psychology literature, also shows that commitment to goals has significant direct and indirect effects on managerial performance (Chong & Chong, 2002; Wentzel, 2002). Nasomboon (2014) reported that leadership commitment also directly affects employee engagement and organizational performance. Performance improvements require willingness to change (Chan, 2004), which may be difficult with a top-down authoritarian management style (Coate, 1993). Participatory performance measurement development may enhance the performance benefits of performance measurement (Groen, Wouters, & Wilderom, 2012). Thus, the third hypothesis is proposed.

H3 Commitment to performance measurement relates positively to safety performance: a) supervisor performance, b) employee performance

Fig. 1 provides a summary of the model which will be tested in the empirical part.

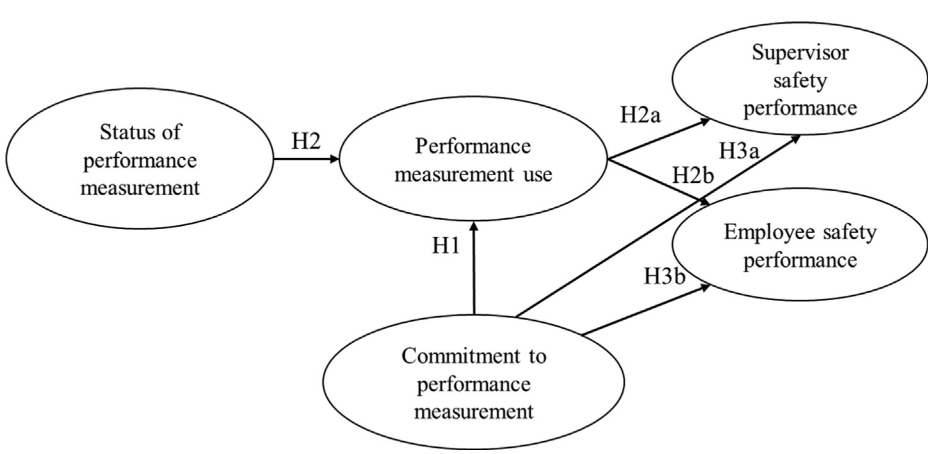

Fig. 1. Research model.

4. Methodology

4.1. Empirical data

The empirical data were gathered with a survey addressed to five companies operating in the service, food, infrastructure construction, manufacturing, and mining industries. Most of the respondents were from the Nordic countries. The mining company operates in Eastern Europe, while all the other companies have their main operations in the Nordic countries. The manufacturing company has a global spread of operations and respondents represented several nationalities. Respondents of the food industry company were mostly from the various Nordic countries. Service and infrastructure industry respondents represented a single Nordic country and mining industry respondents were from a single location. The respondents of this study were selected by focusing on those groups in the company with sufficient knowledge of performance measurement either as users (managers, supervisors) or developers (experts). The respondent lists were prepared together with the participating companies to ensure responses from the different levels and functions of the companies and to improve the reliability of responses.

Electronic survey was used. The respondents gained access to the survey for three weeks. Two reminders were sent during this time. Non-response bias was tested on three groups of respondents: initial invitation, first reminder, and second reminder. A t-test between these groups was done for all research constructs and no statistical differences between the respondent groups were found. Hence, non-response bias should not be a concern in the study.

The survey was sent to 725 respondents and 270 responses were received, yielding a response rate of 37%. The response rate in the individual companies varied from 27% (infrastructure construction) to 69% (mining). Time-to-complete of the survey was 16 minutes on average and the respondents had the opportunity to change their response during this process of filling the questionnaire. Screening of the data and exclusion of missing values (pairwise) meant that the data used in the analysis consisted of 249 responses.

The largest respondent group was middle management, followed by supervisors. These groups play an essential role in facilitating the use of performance measurement (Jääskeläinen & Luukkanen, 2017). Safety experts, managers, and senior managers formed the other notable respondent groups. The majority of the respondents had worked over five years for their present employers. A food industry company provided the highest and a mining industry company the lowest number of responses. The share of other industries was relatively equal (15–20%). Table 1 presents the characteristics of respondents.

Table 1
Demographic information on the respondents.

Total sample	249
Work experience with current employer	less than 1 year 8.1%
	1–3 years 17%
	3–5 years 8.5%
	5–10 years 24.8%
	more than 10 years 41.5%
Position	Senior management 10%
	Middle management 40%
	Supervisor 20.4%
	Safety expert or manager 15.9%
	HR expert or manager 1.5%
	Other expert duties 12.2%
Respondents per company	Company 1 (Infrastructure construction) 58 (21.5%)
	Company 2 (Manufacturing) 53 (19.6%)
	Company 3 (Service) 41 (15.2%)
	Company 4 (Food industry) 93 (34.4%)
	Company 5 (Mining industry) 25 (9.2%)

4.2. Measurement of research variables

Measurement of the research variables in this study utilized a recently developed safety performance measurement maturity model (Jääskeläinen, Tappura, & Pirhonen, 2019). The companies participating in this study were already involved in the testing of the developed measurement instrument and confirmed its applicability in their firm context. The model has three main perspectives: safety performance measurement practices, commitment to safety performance measurement, and use of safety performance measurement. These three themes cover the three lifecycle viewpoints of performance measurement including design, implementation, and use (Bourne, Mills, Wilcox, Neely, & Platts, 2000). In addition, supervisor safety performance and employee safety performance is measured to test the relationship between performance measurement and safety performance. The measurement scale is described in Appendix 1 and more detailed evaluation criteria and survey items in Appendix 2. While safety performance is examined on a Likert scale, the status of performance measurement is assessed by written evaluation criteria presenting four maturity levels (1 = lowest level and 4 = highest level). Written descriptions were used since they can improve objectivity and justify the choice of response options (Cocca & Alberti, 2010; Jääskeläinen & Roitto, 2015).

Safety performance measurement practices represent performance measurement design and include the most established content of the model. Performance measurement practices included four items combining performance measurement and information systems, as suggested by Nudurupati et al. (2011). Commitment to performance measurement commonly regarded as an important success factor in implementing performance measurement (Jääskeläinen & Sillanpää, 2013; Kennerley & Neely, 2002) and it is closely related to safety culture (Fernández-Muñiz, Montes-Peón, & Vázquez-Ordás, 2007). Commitment of both managers (2 items) and employees (2 items) was examined.

There is no widely accepted definition for the use of performance measurement (Tangen, 2005), which is also dependent on the area of management (e.g., safety management). In the use of safety performance measurement, the first evaluation viewpoint was the use of information in planning, that is, the ex-ante perspective (two items). Management in turn represented the ex-post perspective and was measured by two items related to supply chain management, competencies, and rewarding (Cocca & Alberti, 2010; Tung et al., 2011). The extent of performance measurement use at different organizational levels (Van Aken et al., 2005) was also included with one item.

The section measuring the state of occupational safety performance enables links between performance measurement and safety performance. Safety performance was measured by how supervisors dealt with occupational safety in the organizations (5 items) and how employees dealt with safety in the organization (4 items), which were adapted from Nordic Safety Climate Questionnaire (Kines et al., 2011). Both these dimensions are widely regarded as important indicators for safety performance (e.g., Brondino, Silva, & Pasini, 2012; Givehchi et al., 2017).

4.3. Analysis methods

Analysis was conducted with statistical software IBM SPSS Statistics 24 and SmartPLS 3.0. PLS-SEM is a component-based analysis method that does not make assumptions on data distributions. It supports the purpose of this study in predicting and explaining the variance of safety performance (cf. Reinartz, Haenlein, & Henseler, 2009). A bootstrapping technique was used with 5,000 rounds. Table 2 summarizes the data characteristics (correlations, means, standard deviations) in the main research constructs. It should be noted that safety performance had a scale ranging from 1 to 5 whereas the scale of other constructs ranged from 1 to 4.

Common method bias was tested with Harman's single factor test, which revealed that none of the factors covered more than 50% of the data variance. Variance inflation factor (VIF) was used for testing multicollinearity (O'Brien, 2007). VIF of more than 3.3 can be interpreted as an indication of collinearity and common method bias in the model examined (Kock, 2015). As the VIFs of the model are clearly below the threshold (highest VIF is 1.80), it can be concluded that common method bias and collinearity are not likely to create challenges for the analysis (see Tables 3–5).

Reflective constructs were used in the study. Composite reliability (CR) (Fornell & Larcker, 1981), average variance extracted (AVE), and factor loadings (Hair, Anderson, Tatham, & Black, 2014) were used to evaluate the reliability and internal consistency of the constructs. Appendix 2 reports the results. All CRs were above the threshold of 0.7 (Nunnally, 1978) and all AVEs exceeded 0.5 (Fornell & Larcker, 1981). All the items of the survey had loadings above 0.6 (Hulland, 1999).

5. Results

As hypothesized, commitment to performance measurement is conducive to its effective use ($\beta = 0.360$). There is an even stronger relationship ($\beta = 0.399$) between the status of performance measurement and effective use meaning that use of performance information requires appropriate measurement systems. The presence of this relationship means that the first path for H2 is supported. F^2 effect size for these relationships is medium (cf. Hair et al., 2014). The model explains 44.5% of the variation in the use of performance measurement. The Q^2 statistics applying cross-validate redundancy approach demonstrate that the model has medium level of predictive relevance for the use of performance measurement (cf. Hair et al., 2014). The results explaining the use of performance measurement are summarized in Table 3.

Table 4 overviews the results on the antecedents of supervisor safety performance. Status of performance measurement has a significant effect on supervisor safety performance. The use of performance measurement is likewise positively related to supervisor safety performance, although with slightly less significance. This result demonstrates that H2a is supported, but the mediation observed is partial since there is also a direct link between the status of performance measurement and supervisor safety performance. Commitment to performance measurement seems to be the strongest explaining factor for the supervisor safety performance. The result also demonstrates that H3a is supported. Effect sizes (F^2) for all these relationships are small. The model explains 24.8% of the variation in supervisor safety performance and its predictive relevance is small.

The results on the antecedents of employee safety performance are presented in Table 5. There is no direct relationship between the status of performance measurement and employee safety performance. However, the use of performance measurement is positively related to employee safety performance. This result demonstrates that H2b is supported and that the use of performance measurement fully mediates the link between performance measurement status and employee performance. Commitment to performance measurement is again the strongest explaining factor of employee safety performance thereby verifying H3b. Effect sizes for both significant paths are small. The model explains 24.9% of the variation in employee safety performance and its predictive relevance is medium.

A permutation test was used (Hair, Hult, Ringle, & Sarstedt, 2017) to find out whether there were significant differences between industries in the path model studied. All the possible pairwise group comparisons between industries were conducted. Most of the comparisons led to insignificant results, thereby lending support to the external validity of the overall results with the whole data set. Table 6 summarizes the results by including only those pairwise comparison results that were found to be significant. These results should be taken only as indicative since the number of responses per industry was low, especially in the case of the mining industry.

It can be seen that the service industry most often differs from the other industries studied, that is, has different results in some of the paths from the food, infrastructure construction, and manufacturing industries. The most notable difference is that good performance measurement status can as such be beneficial to employee safety performance in the service industry. In addition, advanced performance measurement use is not similarly important for employee safety performance in service industry. The difference between infrastructure construction and mining industries is not meaningful since the significantly different path was not significant in either of the two industries. In the comparison between infrastructure construction and service industries, the service industry results seem to be aligned with the whole data set results, while the infrastructure construction industry differs in the relationship between status of performance measurement and supervisor safety performance. More specifically, it appears that in the infrastructure construction industry there is a need for appropriate performance measurement use to achieve both employee and supervisor safety performance. Mere good status of performance measurement does not suffice.

6. Discussion

Some studies have reported a link between performance measurement and overall organizational performance (Franco-Santos et al., 2012) but the understanding of its consequences for safety performance is limited. Several studies have addressed how safety performance is affected by factors such as organizational, economic, communicational, psychological, procedural, and technical factors (e.g., Kines et al., 2011; Sawacha, 1999), but the understanding of more specific antecedents of safety performance is

Table 2
Characteristics of the data.

Construct	Mean (std. dev.)	1.	2.	3.	4.	5.
1. Performance measurement status	2.85 (0.61)	1				
2. Performance measurement use	2.24 (0.62)	0.585	1			
3. Commitment to performance measurement	2.65 (0.59)	0.497	0.557	1		
4. Safety performance (supervisors)	4.03 (0.70)	0.366	0.410	0.418	1	
5. Safety performance (employees)	3.61 (0.76)	0.341	0.411	0.454	0.632	1

All Pearson's correlations are significant at the 0.001 level.

Table 3
Antecedents of performance measurement use.

Antecedent	PLS SEM analysis results					Conclusion
	β	VIF	t-value	Significance (p-value)	F^2	
Status of performance measurement	0.399	1.341	8.444	0.000	0.214	1. part of H2 supported
Commitment to performance measurement	0.360	1.341	7.154	0.000	0.184	H1 supported
R^2	0.445					
R^2 adjusted	0.441					
Q^2	0.201					

Table 4
Antecedents of supervisor safety performance.

Antecedent	PLS SEM analysis results					Conclusion
	β	VIF	t-value	Significance (p-value)	F^2	
Status of performance measurement	0.183	1.629	2.678	0.007	0.026	Partial mediation for H2a
Performance measurement use	0.174	1.803	2.190	0.029	0.021	2. part of H2 supported → H2a supported
Commitment to performance measurement	0.246	1.587	3.517	0.000	0.050	H3a supported
R^2	0.248					
R^2 adjusted	0.240					
Q^2	0.139					

Table 5
Antecedents of employee safety performance.

Antecedent	PLS SEM analysis results					Conclusion
	β	VIF	t-value	Significance (p-value)	F^2	
Status of performance measurement	0.092	1.629	1.421	0.155	0.007	Full mediation for H2b
Performance measurement use	0.201	1.803	2.736	0.006	0.028	2. part of H2 supported → H2b supported
Commitment to performance measurement	0.297	1.587	3.864	0.000	0.074	H3b supported
R^2	0.249					
R^2 adjusted	0.240					
Q^2	0.153					

Table 6
Industry comparisons.

Industry comparison	Statistically significant difference in path (permutation test p value)	Result in path (industry 1)	Result in path (industry 2)
Food and service	Status of performance measurement → Employee safety performance (0.043)	−0.016, N.S. (food)	0.524, p = 0.013 (service)
	Performance measurement use → Employee safety performance (0.023)	0.414, p = 0.005 (food)	−0.211, N.S. (service)
Infrastructure construction and mining	Status of performance measurement → Performance measurement use (0.028)	0.206, N.S. (infrastructure construction)	0.687, N.S. (mining)
Infrastructure construction and service	Status of performance measurement → Supervisor safety performance (0.031)	−0.052, N.S. (infrastructure construction)	0.472, p = 0.019 (service)
Manufacturing and service	Status of performance measurement → Employee safety performance (0.026)	−0.087, N.S. (manufacturing)	0.524, p = 0.008 (service)
	Performance measurement use → Employee safety performance (0.007)	0.593, p = 0.001 (manufacturing)	−0.211, N.S. (service)

far from clear. Earlier studies have examined performance measurement, for example, as a part of a safety culture survey (e.g., Stemn et al., 2019) or measured the impact of leading indicators on overall safety performance (e.g., Alruqi & Hallowell, 2019). However, to the best of our knowledge, this is the first study to address the links between the maturity of performance measurement and management and safety performance. More specifically, this study contributes to the literature by shedding light on how performance measurement is related to safety performance and how commitment to performance measurement and the use of performance information help to achieve the benefits. This study also provides new understanding by distinguishing between the impacts of performance measurement on employee and supervisor safety performance. This division also makes the results interesting from the viewpoint of performance measurement, which is often seen primarily as a tool for management.

The results support our hypothesis and the earlier performance measurement research (e.g., Kaplan & Norton, 2001; Amaratunga & Baldry, 2002) that performance measurement must be supported by practices of using the information to gain the desired benefits. This is a specific aspect of management practices (Vredenburgh, 2002) and policies (Geldart, Smith, Shannon, & Lohfeld, 2010) reported in earlier studies as antecedents of safety performance. An interesting result was that performance measurement use was more important for employee safety performance than for supervisor safety performance. One explanation could be that rewarding employees was covered in the construct of performance measurement use. But it also makes sense that, specifically, employees benefit from the actual managerial work supported by performance measurement. Another interesting finding was that the status of performance measurement had a significant effect on supervisor safety performance, but it had no relationship with employee safety performance. Hence, it appears that the measures as such create benefits for the supervisors while the actual use of performance information is essential for all employees to gain benefits.

The results show that commitment to performance measurement is even more important than the status of performance measurement and its usage for both supervisor and employee performance. Management and employee commitment to safety have been proven to be significant factors affecting safety performance (e.g., Chen, Wu, Chuang, & Ma, 2009; Wu et al., 2008) and are often the key attributes captured in measuring safety culture. Management commitment can influence the use of performance measurement, but also has a direct impact on safety performance (Wu et al., 2008). Hence, the finding strongly supports earlier research and adds to it by giving some indication of the importance between cultural factors and more technical performance measurement factors in improving safety performance. It is notable that commitment also improves the successful use of performance measurement and hence has many positive effects in the model studied. Goal commitment in general relates positively to commitment to the results (Klein et al., 1999), and safety performance measurement is no exception. From a performance measurement point of view, commitment to performance measurement is especially crucial in the implementation phase of a new or updated performance measurement system (Jääskeläinen & Sillanpää, 2013). Commitment is needed to overcome the resistance typically encountered when implementing new systems.

The model accounts for around 25% of the variation in supervisor and employee safety performance, which means numerous other factors also affect safety performance. Nevertheless, it explains a significant part of the variance given that performance measurement has received only limited attention in the field of safety. For a reference point, McCabe et al. (2017) in their longitudinal study concluded that safety climate represented up to 20% of the variance in safety outcomes. These results are in alignment with the study by Stemn et al. (2019), which reported that performance measurement had the strongest correlation with incident rate of the whole safety culture maturity model.

Earlier research gives a good overall understanding of the factors other than performance measurement and management contributing to safety performance, such as safety leadership and specific leadership style (Barling et al., 2002; Tappura & Nenonen, 2016; Wu et al., 2008; Zohar, 2010), safety training (Tappura, Jääskeläinen, & Pirhonen, 2021; Vinodkumar & Bhasi, 2010), safety communication and dialogue (Griffin & Neal, 2000; Hale et al., 2010; Yorio & Wachter, 2014) as well as collaboration among company personnel (Chen et al., 2009; Geldart et al., 2010). Yorio and Wachter (2014) suggest that the complementary use of different practices has a greater effect on safety performance than individual practices. This also implies that the safety benefits of performance measurement and management may be greater when successfully connected to the above-mentioned factors reflecting the idea of management controls as a package presented in the literature on performance measurement and management control (Malmi & Brown, 2008).

Most of the findings of this study were similar even when comparing the five industries included with each other. However, some interesting differences were also found. Most notably, the service industry seemed to have different results regarding the role of the status of performance measurement in relation to safety performance. It appears that performance measurement as such is supportive of employee safety performance, which indicates that performance measurement provides benefits in reporting and formal control mode, possibly well supported by IT systems. This also means that additional benefits from the various ways of using performance measurement are not similarly important as in the other industries studied. It should be noted that the service company involved in this study operates in the facilities management business reflected by rather standard and continuous logic of operations, and these observations cannot be extended to other types of service industries. By contrast, the infrastructure construction industry seems to be heavily dependent on the mature use of performance measurement. The existence of good performance measurement and reporting systems does not create benefits even for supervisor safety performance. This raises the bar for deriving safety benefits from performance measurement.

7. Conclusions

This study contributes to the literature on safety management by demonstrating a path toward safety performance supported by mature performance measurement and management. The maturity of performance measurement and management is a much more complex issue than the much-discussed shift from lagging to leading indicators. This study further contributed to the literature through the use of a new evaluation instrument that elucidated the status practices of performance measurement and management in the context of safety management. It also advanced the current knowledge by shifting the attention from technical measurement design into implementation of measurement and actual use of performance information.

The results show how safety performance measurement can be implemented to gain its potential benefits. While managers may benefit directly from advanced performance measures, the wider performance benefits among employees materialize only by through using performance measurement properly and committing employees to it.

This study also has limitations that should be acknowledged. Since the data were gathered in a single period of time, it cannot fully verify causalities between the factors studied. This study concentrated on the role of safety performance measurement and management in creating conditions for the process toward better safety in an organization. It should be acknowledged that this process includes many physical and non-physical factors, which were not in the focus of this study. Qualitative studies are also needed to better understand and elaborate on the complex interconnections between the multifaceted factors contributing to successful safety performance measurement and management. The respondent population of this study did not include operative employees since the survey was not designed for that employee group. A further qualitative study would be an appropriate way of elucidating employee perceptions of the benefits of performance management and the maturity of performance measurement. Qualitative studies also have potential in identifying means to develop the use of safety performance measurement to satisfy the needs of different employee groups.

This study identified some differences in the results between industries, which may be explained by specific industrial contingencies. Due to the limited number of observations per industry, these context-specific findings should be regarded as indicative. For example, the service industry observations represent only a specific type of services (i.e., facilities maintenance) and not a wider spectrum of services. Further studies should test further and elaborate the differences between various contexts in gaining safety performance benefits. For example, the balance between performance measurement (as a formal control mechanism) and more informal supportive practices may vary between different contexts, as was seen, for example, in the case of service industry in this study. Further studies could also ascertain the relationships between safety culture and performance measurement in the pursuit of safety performance.

Acknowledgements

This research has received funding from the Finnish Work Environment Fund. The authors are grateful for the support.

Appendix 1. Measurement scale for the maturity of safety performance measurement and management

Performance measurement status

Links between occupational safety performance measurement objects (MS1).

- Linkages between measurement objects have not been considered.
- Linkages between measurement objects are discussed in the organization.
- Factors explaining the main measurement results are partially identified.
- Linkages between measurement objects are analyzed and modeled (e.g. with a strategy map). There is a common understanding in the organization regarding the factors that should be improved to affect the main measurement results.

Reliability of occupational safety-related performance information (MS2).

- Top managers do not trust the performance information.
- There are several interpretations of the performance information. Employees do not trust the performance information.
- There are differing interpretations of some parts of the performance information. Top managers trust the performance information.
- Indicators provide mainly unambiguous information. Employees trust the performance information.

Process for reviewing and updating occupational safety performance indicators (MS3).

- New indicators are not taken into use.
- New indicators are taken into use in a random manner.
- New indicators are taken into use when needed but the usefulness of the old indicators is not evaluated.
- There is a regular evaluation and development of indicators. Old indicators are discarded when necessary.

Information systems in gathering occupational safety-related performance information (MS4).

- Performance information is gathered manually when needed.
- Performance information is gathered manually to a large extent. Only information on a few key indicators is gathered automatically.
- Most of the performance information is gathered with information systems. Information systems enable the provision of information in real-time.
- Performance information is gathered automatically. The most important information systems communicate which each other and include consistent data.

Commitment to performance measurement

The role of employees in gathering occupational safety-related performance information (CC1).

- Employees do not report incidents affecting safety.
- Employees report only incidents seriously affecting safety.
- Employees also report incidents with a minor effect on safety.
- Employees are active in taking initiatives to improve safety performance.

Employee commitment to occupational safety performance measurement (CC2).

- Personnel regard measurement as an extra burden.
- There is no major criticism of measurement among employees.
- Measurement is regarded as useful in the work community. The views of employees are taken into account when developing measurement.
- Employees feel that measurement improves fairness (e.g. in rewarding). Employees initiate efforts to improve measurement.

Managerial support for occupational safety performance measurement (CC3).

- Performance measurement has no managerial support at any level.
- Top management supports performance measurement.
- Supervisors regard performance measurement as important and employees are encouraged to gather and report performance information.
- Sufficient resources and training are provided to implement performance measurement.

Resources for occupational safety performance measurement (CC4).

- There are no resources for sustaining safety performance measurement practices.
- There are sufficient resources for reporting the current indicators.
- There are sufficient resources for systematic analysis of our current indicators
- There are sufficient resources for systematic development of new indicators and evaluation of the old indicators.

Performance measurement use

Use of performance information in planning occupational safety issues (USE1).

- Performance information is utilized in analyzing only past incidents.
- Performance information is utilized to identify and analyze risks.
- Performance information is systematically utilized to prevent occupational safety problems and to improve work practices.
- A wide range of experts collaborates in the prevention of incidents affecting occupational safety and development of work practices. This work is supported by a wide range of performance information.

Defining action plans related to occupational safety (USE2).

- Indicators are not used in identifying aspects for development.
- Indicators are used in the identification of aspects in need of development (e.g. identifying a part in the process which causes many safety hazards).

- Indicators are used to support the preparation of action plans (e.g. prioritizing of procedures).
- Definition and implementation of action plans are done systematically and mainly based on performance information (e.g. action plans are prioritized and controlled with the support of performance information).

Development of occupational safety competencies (USE3).

- Indicators are not linked to occupational safety competencies.
- Indicators are used to identify occupational safety competencies (e.g. results of appraisal interviews, training costs/employees per year).
- Occupational safety competencies are constantly monitored in the organization (e.g. self-evaluations, employees fulfilling qualifications) to identify development targets.
- Individual competence development plans are created for the employees on the basis of performance information.

Use of occupational safety performance measurement at different levels (USE4).

- Indicators are utilized only at the level of the whole organization.
- Indicators are utilized at the supervisor level.
- Indicators are utilized at the employee level (e.g. in appraisal interviews).
- Indicators are utilized at the level of suppliers and subcontractors operating on our premises.

Use of performance information in occupational safety management of supply chains (USE5).

- Supplier/contractor safety performance is not monitored.
- Supplier/contractor safety is evaluated in contract preparation and a target level for safety is set.
- Performance information regarding safety of suppliers/contractors is regularly monitored.
- Performance information supports communication and collaboration development with suppliers/contractors.

Safety performance

Likert scale (1–5), strongly agree, agree, neither agree nor disagree, disagree, strongly disagree.

Supervisor safety performance (SPV1-5, see Appendix 2).

Employee safety performance (EP1-4, see Appendix 2).

Appendix 2. Measurement items, loadings, construct reliability and validity scores

Construct	Item	Loading	CR	AVE
Performance measurement status	MS1 Links between occupational safety performance measurement objects	0.743	0.807	0.512
	MS2 Reliability of occupational safety-related performance information	0.658		
	MS3 Process for reviewing and updating occupational safety performance indicators	0.708		
	MS4 Information systems in gathering occupational safety-related performance information	0.666		
Commitment to performance measurement	CC1 The role of employees in gathering occupational safety-related performance information	0.758	0.830	0.551
	CC2 Employee commitment to occupational safety performance measurement	0.784		
	CC3 Managerial support for occupational safety performance measurement	0.670		
	CC4 Resources for occupational safety performance measurement	0.702		
Performance measurement use	USE1 Use of performance information in planning occupational safety issues	0.725	0.837	0.508
	USE2 Defining action plans related to occupational safety	0.700		
	USE3 Development of occupational safety competencies	0.681		
	USE4 Use of occupational safety performance measurement at different levels	0.642		
	USE5 Use of performance information in occupational safety management of supply chains	0.690		
Supervisor safety performance	SVP1 Supervisors encourage employees here to work in accordance with safety rules - even when the work schedule is tight.	0.806	0.894	0.630
	SVP2 Supervisors ensure that every-one receives the necessary information on safety	0.822		
	SPV3 Supervisors ensure that safety problems discovered during safety rounds/evaluations are corrected immediately.	0.785		
	SPV4 Supervisors make sure that every-one can influence safety in their work environment.	0.807		
	SPV5 Supervisors look for causes, not guilty persons, when an accident occurs.	0.709		

(*continued on next page*)

Appendix 2 (*continued*)

Construct	Item	Loading	CR	AVE
Employee safety performance	EP1 Employees try hard together to achieve a high level of safety.	0.833	0.901	0.696
	EP2 Employees help each other to work safely.	0.854		
	EP3 Employees never accept risk-taking even if the work schedule is tight.	0.751		
	EP4 Employees take each other's opinions and suggestions concerning safety seriously.	0.860		

References

Aksorn, T., & Hadikusumo, B. H. (2008). Critical success factors influencing safety program performance in Thai construction projects. *Safety Science, 46*(4), 709–727.

Alruqi, W. M., & Hallowell, M. R. (2019). Critical success factors for construction safety: Review and meta-analysis of safety leading indicators. *Journal of Construction Engineering and Management, 145*(3).

Amaratunga, D., & Baldry, D. (2002). Moving from performance measurement to performance management. *Facilities, 20*(5/6), 217–223.

Amir-Heidari, P., Maknoon, R., Taheri, B., & Bazyari, M. (2017). A new framework for HSE performance measurement and monitoring. *Safety Science, 100*, 157–167.

Arezes, P. M., & Sérgio Miguel, A. (2003). The role of safety culture in safety performance measurement. *Measuring Business Excellence, 7*(4), 20–28.

Barling, J., Loughlin, C., & Kelloway, E. K. (2002). Development and test of a model linking safety-specific transformational leadership and occupational safety. *Journal of Applied Psychology, 87*(3), 488.

Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review, 63*(5), 586–606.

Bititci, U. S., Ackermann, F., Ates, A., Davies, J., Garengo, P., Gibb, S., et al. (2011). Managerial processes: Business process that sustain performance. *International Journal of Operations & Production Management, 31*(8), 851–891.

Blair, E., & O'Toole, M. (2010). Leading measures: Enhancing safety climate and driving safety performance. *Professional Safety, 55*(08), 29–34.

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *International Journal of Operations & Production Management, 20*(7), 754–771.

Bourne, M., Kennerley, M., & Franco-Santos, M. (2005). Managing Through Measures: A Study of Impact on Performance. *Journal of Manufacturing Technology Management, 16*(4), 373–395.

Brondino, M., Silva, S. A., & Pasini, M. (2012). Multilevel approach to organizational and group safety climate and safety performance: Co-workers as the missing link. *Safety Science, 50*(9), 1847–1856.

BS 18004:2008. (2008) Guide to achieving effective health and safety performance. British Standards Institution, p. 4–5.

Burke, M. J., & Signal, S. M. (2010). Workplace safety: A multilevel, interdisciplinary perspective. *Research in Personnel and Human Resources Management, 29*, 1–47.

Carder, B., & Ragan, P. W. (2003). A survey-based system for safety measurement and improvement. *Journal of Safety Research, 34*(2), 157–165.

Cavalluzzo, K. S., & Ittner, C. D. (2004). Implementing performance measurement innovations: Evidence from government. *Accounting, Organizations and Society, 29*(3–4), 243–267.

Chan, Y. (2004). Performance measurement and adoption of balanced scorecards – a survey of municipal governments in the USA and Canada. *International Journal of Public Sector Management, 17*(3), 204–221.

Chen, C.-Y., Wu, G.-S., Chuang, K.-J., & Ma, C.-M. (2009). A comparative analysis of the factors affecting the implementation of occupational health and safety management systems in the printed circuit board industry in Taiwan. *Journal of Loss Prevention in the Process Industries, 22*(2), 210–215.

Chong, V. K., & Chong, K. M. (2002). Budget goal commitment and informational effects of budget participation on performance: A structural equation modeling approach. *Behavioral Research in Accounting, 14*(1), 65–86.

Clarke, S. (2013). Safety leadership: A meta-analytic review of transformational and transactional leadership styles as antecedents of safety behaviours. *Journal of Occupational and Organizational Psychology, 86*(1), 22–49.

Coate, E. (1993). The introduction of total quality management at Oregon State University. *Higher Education, 25*(3), 303–320.

Cocca, P., & Alberti, M. (2010). A framework to assess performance measurement systems in SMEs. *International Journal of Productivity and Performance Management, 59*(2), 186–200.

Cooper, M. D. (2000). Towards a model of safety culture. *Safety Science., 36*(2), 111–136.

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

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

Fernández-Muñiz, B., Montes-Peón, J. M., & Vázquez-Ordás, C. J. (2009). Relation between occupational safety management and firm performance. *Safety Science, 47*(7), 980–991.

Fernández-Muñiz, B., Montes-Peón, J. M., & Vázquez-Ordás, C. J. (2007). Safety culture: Analysis of the causal relationships between its key dimensions. *Journal of Safety Research, 38*(6), 627–641.

Flin, R. (2003). "Danger—men at work": Management influence on safety. *Human Factors and Ergonomics in Manufacturing and Service Industries, 13*(4), 261–268.

Flin, R., Mearns, K., O'Connor, P., & Bryden, R. (2000). Measuring climate: Identifying the common features. *Safety Science, 34*, 177–192.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

Franco-Santos, M., Lucianetti, L., & Bourne, M. (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research. *Management Accounting Research, 23*(2), 79–119.

Geldart, S., Smith, C. A., Shannon, H. S., & Lohfeld, L. (2010). Organisational practices and workplace health and safety: A cross-sectional study in manufacturing companies. *Safety Science, 48*, 562–569.

Givehchi, S., Hemmativaghef, E., & Hoveidi, H. (2017). Association between safety leading indicators and safety climate levels. *Journal of Safety Research, 62*, 23–32.

Glendon, I. (2008). Safety culture: snapshot of a developing concept. *Journal of Occupational Health and Safety, 24*(3), 179–189.

Goncalves Filho, A. P., & Waterson, P. (2018). Maturity models and safety culture: A critical review. *Safety Science, 105*, 192–211.

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*, 347–358.

Groen, B. A., Wouters, M. J., & Wilderom, C. P. (2012). Why do employees take more initiatives to improve their performance after co-developing performance measures? A field study. *Management Accounting Research, 23*(2), 120–141.

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

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2014). *Multivariate Data Analysis* (7th ed.). Upper Saddle River, NJ: Pearson.

Hair, J. F., Jr, Hult, G. T. M., Ringle, C., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Sage Publications.

Hale, A. R., Guldenmund, F. W., Van Loenhout, P. L. C. H., & Oh, J. I. H. (2010). Evaluating safety management and culture interventions to improve safety: Effective intervention strategies. *Safety Science, 48*(8), 1026–1035.

Hannula, M., & Lönnqvist, A. (2002). *Suorituskyvyn mittauksen käsitteet (engl. Concepts of performance measurement)*. Helsinki: Metalliteollisuuden kustannus.

Haslam, C., O'Hara, J., Kazi, A., Twumasi, R., & Haslam, R. (2016). Proactive occupational safety and health management: Promoting good health and good business. *Safety Science, 81*, 99–108.

Henri, J. F. (2006). Management control systems and strategy: A resource-based perspective. *Accounting, Organizations and Society, 31*(6), 529–558.

Hoffmeister, K., Gibbons, A. M., Johnson, S. K., Cigularov, K. P., Chen, P. Y., & Rosecrance, J. C. (2014). The differential effects of transformational leadership facets on employee safety. *Safety Science, 62*, 68–78.

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal, 20*, 195–204.

Hwang, M. I., & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science, 25*(3), 213–218.

Iselin, E. R. (1988). The effects of information load and information diversity on decision quality in a structured decision task. *Accounting, Organizations and Society, 13*(2), 147–164.

ISO 45001:2018 Occupational health and safety management systems. Requirements with guidance for use. *International Organization for Standardization*, Geneva, Switzerland, 47 p.

ISO/IEC (2004). *Guide 2: Standardization and related activities – General vocabulary* (p. 60). Geneva, Switzerland: International Organization for Standardization.

Jääskeläinen, A., & Luukkanen, N. (2017). The use of performance measurement information in the work of middle managers. *International Journal of Productivity and Performance Management, 66*(4), 479–499.

Jääskeläinen, A., & Sillanpää, V. (2013). Overcoming challenges in the implementation of performance measurement: Case studies in public welfare services. *International Journal of Public Sector Management, 26*(6), 440–454.

Jääskeläinen, A., & Roitto, J. M. (2015). Designing a model for profiling organizational performance management. *International Journal of Productivity and Performance Management, 64*(1), 5–27.

Jääskeläinen A., Tappura S., Pirhonen J. (2019). Maturity Analysis of Safety Performance Measurement. In: Ahram T. et al. (Eds.), *Human Systems Engineering and Design II. IHSED 2019.* Advances in Intelligent Systems and Computing, Vol. 1026 (pp. 529–535). Springer, Cham.

Jääskeläinen A., Tappura S., Pirhonen J. (2020). Safety Performance Measurement Maturity in Finnish Industrial Companies. In: Arezes P. et al. (Eds.), *Occupational and Environmental Safety and Health II*. Studies in Systems, Decision and Control, vol 277 (pp. 41–49). Springer, Cham.

Kaassis, B., & Badri, A. (2018). Development of a preliminary model for evaluating occupational health and safety risk management maturity in small and medium-sized enterprises. *Safety, 4*(1), 5.

Kaplan, R. S., & Norton, D. P. (2001). Transforming the balanced scorecard from performance measurement to strategic management: Part II. *Accounting Horizons, 15*(2), 147–160.

Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review., 74*(1), 75–85.

Kennerley, M., & Neely, A. (2002). A framework of the factors affecting the evolution of performance measurement systems. *International Journal of Operations & Production Management, 22*(11), 1222–1245.

Kines, P., Lappalainen, J., Mikkelsen, K. L., Olsen, E., Pousette, A., Tharaldsen, J., et al. (2011). Nordic Safety Climate Questionnaire (NOSACQ-50): A new tool for diagnosing occupational safety climate. *International Journal of Industrial Ergonomics, 41*, 634–646.

Klein, H. J., Wesson, M. J., Hollenbeck, J. R., & Alge, B. J. (1999). Goal commitment and the goal-setting process: Conceptual clarification and empirical synthesis. *Journal of Applied Psychology, 84*(6), 885–896.

Koivupalo, M., & Reiman, A. (2017). Health and safety performance indicators in global steel company. *Safety Science Monitor, 20*(2).

Lebas, M. J. (1995). Performance measurement and performance management. *International Journal of Production Economics, 41*(1–3), 23–35.

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

Lemieux-Charles, L., McGuire, W., Champagne, F., Barnsley, J., Cole, D., & Sicotte, C. (2003). The use of multilevel performance indicators in managing performance in health care organizations. *Management Decision, 41*(8), 760–770.

Lingard, H., Wakefield, R., & Cashin, P. (2011). The development and testing of a hierarchical measure of project OHS performance. *Engineering, Construction and Architectural Management, 18*(1), 30–49.

Lingard, H., Cooke, T., & Blismas, N. (2012). Do perceptions of supervisors' safety responses mediate the relationship between perceptions of the organizational safety climate and incident rates in the construction supply chain? *Journal of Construction Engineering and Management, 138*(2), 234–241.

Longenecker, C. O., & Fink, L. S. (2001). Improving Management Performance in Rapidly Changing Organizations. *Journal of Management Development, 20*(1), 7–18.

Lynch, R. L. (1995). *Measure up! Yardsticks for continuous improvement*. Blackwell Publishing.

Malmi, T., & Brown, D. A. (2008). Management control systems as a package—Opportunities, challenges and research directions. *Management Accounting Research, 19*(4), 287–300.

Marx, F., Wortmann, F., & Mayer, J. H. (2012). A maturity model for management control systems. *Business & Information Systems Engineering, 4*(4), 193–207.

McCabe, B. Y., Alderman, E., Chen, Y., Hyatt, D. E., & Shahi, A. (2017). Safety performance in the construction industry: Quasi-longitudinal study. *Journal of Construction Engineering and Management, 143*(4).

Nasomboon, B. (2014). The relationship among leadership commitment, organizational performance, and employee engagement. *International Business Research, 7*(9), 77–90.

Neal, A., Griffin, M. A., & Hart, P. M. (2000). The impact of organizational climate on safety climate and individual behavior. *Safety Science, 34*, 99–109.

Neely, A. (1998). *Measuring Business Performance. Why, What and How?* London: Profile Books Ltd.

Neely, A. D., Adams, C., & Kennerley, M. (2002). *The performance prism: The scorecard for measuring and managing business success*. London: Prentice Hall Financial Times.

Nielsen, K. T. (2000). Organization theories implicit in various approaches to OHS management. In *Systematic Occupational Health and Safety Management* (pp. 99–123). Pergamon Press.

Nudurupati, S. S., Bititci, U. S., Kumar, V., & Chan, F. T. (2011). State of the art literature review on performance measurement. *Computers & Industrial Engineering, 60*(2), 279–290.

Nunnally, J. C. (1978). *Psychometric theory* (Vol. 2) New York: McGraw-Hill.

O'Brien, R. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity, 41*(5), 673–690.

O'Toole, M. (2002). The relationship between employees' perceptions of safety and organizational culture. *Journal of Safety Research, 33*, 231–243.

Podgórski, D. (2015). Measuring operational performance of OSH management system – A demonstration of AHP-based selection of leading key performance indicators. *Safety Science, 73*, 146–166.

Rajendran, S. (2013). Enhancing construction worker safety performance using leading indicators. *Practice Periodical on Structural Design and Construction, 18*(1), 45–51.

Reiman, T., & Pietikäinen, E. (2012). Leading indicators of system safety–monitoring and driving the organizational safety potential. *Safety Science, 50*(10), 1993–2000.

Reinartz, W. J., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Market Research, 26*(4), 332–344.

Saksvik, P. Ø., Nytrø, K., Dahl-Jørgensen, C., & Mikkelsen, A. (2002). A process evaluation of individual and organizational occupational stress and health interventions. *Work & Stress, 16*(1), 37–57.

Salas, R., & Hallowell, M. (2016). Predictive validity of safety leading indicators: Empirical assessment in the oil and gas sector. *Journal of Construction Engineering and Management, 142*(10).

Sawacha, E., Naoum, S., & Fong, D. (1999). Factors affecting safety performance on construction sites. *International Journal of Project Management, 17*(5), 309–315.

Sheehan, C., Donohue, R., Shea, T., Cooper, B., & Cieri, H. (2016). Leading and lagging indicators of occupational health and safety: The moderating role of safety leadership. *Accident Analysis and Prevention, 92*, 130–138.

Shields, M. D. (1995). An empirical analysis of firms' implementation experiences with activity-based costing. *Journal of Management Accounting Research, 7*(1), 148–165.

Sievänen, M., Nenonen, N., & Hämäläinen, P. (2013). Economic impacts of occupational health and safety interventions – a critical analysis based on the nine-box model of profitability. *Proceedings of the 45th Annual International Conference of the Nordic Ergonomics and Human Factors Society NES 2013. August 11–14, 2013, Reykjavik, Iceland*.

Sinelnikov, S., Inouye, J., & Kerper, S. (2015). Using leading indicators to measure occupational health and safety performance. *Safety Science, 72*, 240–248.

Stemn, E., Bofinger, C., Cliff, D., & Hassall, M. E. (2019). Examining the relationship between safety culture maturity and safety performance of the mining industry. *Safety Science, 113*, 345–355.

Swuste, P., Theunissen, J., Schmitz, P., Reniers, G., & Blokland, P. (2016). Process safety indicators, a review of literature. *Journal of Loss Prevention in the Process Industries, 40*, 162–173.

Tangen, S. (2005). Demystifying productivity and performance. *International Journal of Productivity and Performance Management, 54*(1), 34–46.

Tappura, S., Jääskeläinen, A., & Pirhonen, J. (2021). Performance Implications of Safety Training. In P. M. Arezes & R. L. Boring (Eds.). *Advances in Safety Management and Human Performance. AHFE 2021. Lecture Notes in Networks and Systems* (Vol. 262, pp. 295–301). Cham: Springer.

Tappura, S., & Nenonen, N. (2016). Categorization of effective safety leadership facets. In P. Arezes & P. Carvalho (Eds.), *Ergonomics and Human Factors in Safety Management* (pp. 367–383). Boca Raton: CRC Press.

Tappura, S., Sievänen, M., Heikkilä, J., Jussila, A., & Nenonen, N. (2015). A management accounting perspective on safety. *Safety Science, 71*, 151–159.

Tung, A., Baird, K., & Schoch, H. P. (2011). Factors influencing the effectiveness of performance measurement systems. *International Journal of Operations & Production Management, 31*(12), 1287–1310.

Van Aken, E. M., Letens, G., Coleman, G. D., Farris, J., & Van Goubergen, D. (2005). Assessing maturity and effectiveness of enterprise performance measurement systems. *International Journal of Productivity and Performance Management, 54*(5/6), 400–418.

Vinodkumar, M. N., & Bhasi, M. (2009). Safety climate factors and its relationship with accidents and personal attributes in the chemical industry. *Safety Science., 47*(5), 659–667.

Vredenburgh, A. G. (2002). Organisational safety: Which management practices are most effective in reducing employee injury rates? *Journal of Safety Research, 33*, 259–276.

Waterson, P., Jenkins, D. P., Salmon, P. M., & Underwood, P. (2017). 'Remixing Rasmussen': The evolution of Accimaps within systemic accident analysis. *Applied Ergonomics, 59*, 483–503.

Wentzel, K. (2002). The influence of fairness perceptions and goal commitment on managers' performance in a budget setting. *Behavioral Research in Accounting, 14*(1), 247–271.

Wettstein, T., & Kueng, P. A. (2002). A maturity model for performance measure systems. In C. Brebbia & P. Pascola (Eds.), *Management Information Systems: GIS and Remote Sensing* (pp. 113–122). Southampton: WIT Press.

Vinodkumar, M. N., & Bhasi, M. (2010). Safety management practices and safety behaviour: Assessing the mediating role of safety knowledge and motivation. *Accident Analysis and Prevention, 42*, 2082–2093.

Wu, T. (2005). The validity and reliability of safety leadership scale in universities of Taiwan. *International Journal of Technology and Engineering Education, 2*(1), 27–42.

Wu, T. C., Chen, C. H., & Li, C. C. (2008). A correlation among safety leadership, safety climate and safety performance. *Journal of Loss Prevention in the Process Industries, 21*(3), 307–318.

Zohar, D. (2010). Thirty years of safety climate research: Reflections and future directions. *Accident Analysis & Prevention, 42*(5), 1517–1522.

Zwetsloot, G., Leka, S., Kines, P., & Jain, A. (2020). Vision zero: Developing proactive leading indicators for safety, health and wellbeing at work. *Safety Science, 130*, 104890.

Yorio, P. L., & Wachter, J. K. (2014). The impact of human performance focused safety and health management practices on injury and illness rates: Do size and industry matter? *Safety Science, 62*, 157–167.

Aki Jääskeläinen, D.Sc. (Tech.), is an Associate Professor at Tampere University and adjunct professor at Lappeenranta-Lahti University of Technology. His main research interests relate to performance measurement, management control, operations management and purchasing and supply management. Currently, he is the responsible director of SafePotential project, which investigates safety performance measurement and develops models and practices for better utilization of safety performance information. He has around 100 publications and his h-index in Scopus is 11. He is a three-time winner of Emerald's Outstanding Paper Award.

Sari Tappura is a project manager and senior researcher at the Center for Safety Management and Engineering, department of Industrial Engineering and Management at Tampere University, Finland. She has over 10 years' experience in managing safety-related research projects and teaching at the university. Her interests include occupational health and safety, safety management, safety leadership and safety performance measurement. Currently, she works as a researcher in the SafePotential project, which investigates safety performance measurement and develops models and practices for better utilization of safety performance information. She has published over 60 scientific or professional articles related to safety during the last 10 years.

Julius Pirhonen, M.Sc. (Tech.), works as a project researcher at the unit of Industrial Engineering and Management at Tampere University, Finland. His research topics relate to safety performance measurement, safety culture and maturity models.