Safety Science 83 (2016) 48–58

Contents lists available at ScienceDirect

Safety Science

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

Exploring the state of health and safety management system performance measurement in mining organizations

Emily Joy Haas *, Patrick Yorio

National Institute for Occupational Safety and Health (NIOSH), Office of Mine Safety and Health Research (OMSHR), 626 Cochrans Mill Rd, Pittsburgh, PA 15236, United States

ARTICLE INFO

Article history:
Received 22 May 2015
Received in revised form 17 August 2015
Accepted 7 November 2015

Keywords:
Health and safety management system
Lagging indicators
Leading indicators
Mining
Performance measurement
Multiple method measurement

ABSTRACT

Complex arguments continue to be articulated regarding the theoretical foundation of health and safety management system (HSMS) performance measurement. The culmination of these efforts has begun to enhance a collective understanding. Despite this enhanced theoretical understanding, however, there are still continuing debates and little consensus. The goal of the current research effort was to empirically explore common methods to HSMS performance measurement in mining organizations. The purpose was to determine if value and insight could be added into the ongoing approaches of the best ways to engage in health and safety performance measurement. Nine site-level health and safety management professionals were provided with 133 practices corresponding to 20 HSMS elements, each fitting into the plan, do, check, act phases common to most HSMS. Participants were asked to supply detailed information as to how they (1) assess the performance of each practice in their organization, or (2) would assess each practice if it were an identified strategic imperative. Qualitative content analysis indicated that the approximately 1200 responses provided could be described and categorized into *interventions*, *organizational performance*, and *worker performance*. A discussion of how these categories relate to existing indicator frameworks is provided. The analysis also revealed divergence in two important measurement issues; (1) quantitative vs qualitative measurement and reporting; and (2) the primary use of objective or subjective metrics. In lieu of these findings we ultimately recommend a balanced measurement and reporting approach within the three metric categories and conclude with suggestions for future research.

Published by Elsevier Ltd.

1. Introduction

Health and safety management systems (HSMS) are broadly characterized as a set of institutionalized, interrelated, and interacting elements strategically designed to establish and achieve occupational health and safety (H&S) goals and objectives (ANSI/AIHA Z10, 2012). The goals and objectives of HSMS activities center on occupational injury, illness, and loss prevention. Numerous consensus standards provide guidance across industrial sectors as to the types of activities important for inclusion in an organization's HSMS (e.g., ANSI/AIHA Z10-2012; British Standards Institute, 2007; U.S. National Mining Association's [NMA] CORESafety). Although HSMS frameworks offer varying types and numbers of elements, each system is grounded in the now institutionalized Deming/Shewhart plan-do-check-act cycle (Haight et al., 2014). In addition, they each include activities designed to develop an organization's internal infrastructure in order to enhance the effectiveness of the activities within the cycle (e.g., employee involvement, leadership, organization and allocation of resources, etc.).

HSMS performance measurement and monitoring activities are used to determine whether the system is functioning as designed and to help evaluate the system's overall effectiveness (ANSI/AIHA Z-10, 2012). Performance indicators are the fundamental building blocks to the measurement and monitoring process of an HSMS. The information provided via performance indicators can be used to facilitate strategic H&S management decision-making and the implementation of appropriate risk management actions on behalf of the organization. However, theorists and researchers continue to debate the theoretical and practical perspectives of occupational HSMS performance measurement. This issue was recognized as early as when Petersen (2001) stated "Measuring the effectiveness of an organization's safety system has been a particularly difficult problem for all organizations" (p. 54). Indications that HSMS performance measurement remains a current concern can be gleaned from more recent publications. For example, Juglaret and colleagues (2011) argued that, although HSMS is an established tool to manage

* Corresponding author. Tel.: +1 412 386 4627
E-mail addresses: EJHaas@cdc.gov (E.J. Haas), PYorio@cdc.gov, patyorio@yahoo.com (P. Yorio).

http://dx.doi.org/10.1016/j.ssci.2015.11.009
0925-7535/Published by Elsevier Ltd.

occupational safety and health, how to effectively measure the performance and control of these systems remains a question in the literature. In response to the lack of empirical guidance regarding HSMS performance measurement, health and safety managers were asked how they commonly measure HSMS effectiveness via a variety of practices. This paper reports on common performance measures utilized by a sample of health and safety managers of mining organizations and, based on the data, proposes a new HSMS performance measurement framework that can be used to assess the performance of occupational health and safety initiatives.

2. Literature review

To date, there appears to be little consensus among researchers and practitioners in regard to the terms used to categorize the types of performance indicators used to assess the effectiveness of HSMS elements and practices (e.g. Körvers and Sonnemans, 2008; Laitinen et al., 2013; Reiman and Pietikäinen, 2012). Several indicator frameworks exist in the literature however, and most adhere to the leading/lagging indicator typology.

2.1. A brief review of performance indicator categorizing frameworks

Several indicator frameworks are similar in practice but utilize different terms to describe their areas of focus. For instance, Körvers and Sonnemans (2008) and Laitinen et al. (2013) synonymously refer to *proactive/reactive* indicator types. They argue that *proactive* indicators can be subcategorized as either *predictive* or *monitoring. Predictive proactive* indicators supply information on the types of managerial actions that have been taken to reduce workplace risk. *Monitoring proactive* indicators include H&S related outcomes observed prior to the occurrence of a major incident such as small releases of hazardous substances or near misses, the results of safety inspections and behavioral observations, the results of safety audits, and safety attitudes. In a similar argument, Laitinen et al. (2013) suggest that the *proactive activity* indicators capture the managerial activities being done in organizations (e.g., number of audits completed; number of workers trained). They describe *proactive activity* metrics as indications of what activities are being done in the organization rather than information about the results of those activities. In contrast, *proactive outcome* indicators, such as personnel knowledge, focus on results and observable outcomes rather than mere activities.

Similarly, Reiman and Pietikäinen (2012) suggest that performance indicators could be categorized as *lag outcome*, *lead monitor* indicators, and *lead drive* indicators. They indicate *lag outcomes* do not merely capture harm associated with traditional *lagging* indicators but capture information related to the temporary end result(s) of a continuous process. *Lead drive* indicators reflect the workplace activities aimed at improving safety and include, for example, measures of supervisory activity and practices related to physical hazard control. They are ''measures of the fulfillment of the selected safety management activities'' (p. 1995). *Lead monitor* indicators measure the potential of the organization to achieve safety and include measures related to worker safety motivation, awareness, and knowledge.

As is evident from the brief review, despite varying terminologies, most theoretical categories integrate the leading/lagging indicator typology. This framework is most prevalent in mining as well (Industrial Council on Mining & Metals, 2012). Therefore, we focus on the nuances of this framework in the subsequent sections as well as potential drawbacks of applying such measures in a mining environment.

2.2. Limitations of current indicator frameworks

The terms *leading* and *lagging* originated within the economics discipline as a way to describe key indicators of economics and business cycle performance over time-dependent phases (Diebold and Rudebusch, 1994; Shishkin, 1961). Perhaps because of the introduction and widespread acceptance of HSMS and its cyclic similarity to economic and business cycles, the *leading/lagging* and related terms (e.g., *proactive/reactive*; *leading/trailing*; *upstream/downstream*) were adopted to help identify and describe the types of indicators important to assess the effectiveness of HSMS performance.

Given the appeal and seeming utility of the *leading/lagging* indicator categories, numerous efforts have been undertaken to develop the theory and application underpinning the framework. Strictly based on their definitions, *lagging* indicators represent information related to significant safety incidents such as injuries, illnesses, and major property losses, while *leading* indicators can conceptually span the plan, do, and proactive checking phases of the management system cycle. The purpose of *leading* indicators is to understand and manage the organizational circumstances thought to precede undesired occupational H&S outcomes (International Council on Mining and Metals, 2012). With the exception of Körvers and Sonneman's (2008) discussion related to *predictive proactive* indicators, however, most work on *leading indicators* seems to neglect management practices related to the risk management planning phase of the system cycle. This suggests that indicator frameworks grounded in the leading/lagging nuance may not be optimally positioned to assess the actual effectiveness of the full breadth of HSMS activities that take place before and after the occurrence of occupational injuries (i.e., practices involved with preventing and investigating safety incidents and implementing corrective actions) or management review activities.

Additional limitations of the leading/lagging terms have been articulated (e.g., Janicak, 2011; Juglaret et al., 2011; International Council on Mining and Minerals, 2012; Wachter, 2012; Reiman and Pietikäinen, 2012; O'Neill et al., 2013; Payne et al., 2009; Hale, 2009; Hopkins, 2009). First, an imprecise link between the two terms has been noted. One of the more practically beneficial features of the *leading/lagging* framework centers on the premise that assessing and managing well-chosen *leading* indicators minimizes the potential for poor performance as assessed through *lagging* indicators. The most effective *leading* indicators within an organization can then be determined through rank order of the magnitude of the correlations (presumably negative) linking the list of *leading* indicators to the identified *lagging* indicators. Although attractive theoretically, in practice there are complex causal pathways between *leading* and *lagging* indicators as well as time-lagged linkages (Janicak, 2011; International Council on Mining and Minerals, 2012; Wachter, 2012; O'Neill et al., 2013).

This complex causal pathway argument is consistent with the idea of statistical moderation—i.e., the relationship between two variables is contingent upon the level of a third, perhaps unidentified, variable. Because the possible moderating contingencies within an organizational context are numerous, this argument suggests that it may be difficult to make strategic decisions regarding HSMS activities merely based on the correlation between the *leading* and *lagging* indicators chosen—especially in the case of a null or small correlation between them. Relatedly, in some cases, the effect of *leading* on *lagging* indicators may take time to actualize; some have suggested years perhaps (Wachter, 2012). This, again, increases the difficultly and imprecision of strategic management decision-making through the use of *leading/lagging* indicator correlations (i.e., how long should a *leading* indicator with no observable effect be tracked and managed before new indicators are chosen and the stopwatch starts again?). These issues may be one reason why Laitinen and colleagues (2013) suggested that "a lack of effective proactive indicators is probably the single biggest problem faced in occupational safety and health management today" (p. 69).

Second, many have argued that the *leading*/*lagging* terms do not sufficiently capture the complexity of the phenomena of organizational health and safety. For example, given that system *lagging* indicators often shape *leading* indicators, ideas of reverse causation have been articulated between the two terms (i.e., *leading* indicators can lag behind *lagging* indicators just as *lagging* indicators hypothetically lag behind *leading* indicators). More specifically, Payne et al. (2009) suggested that the occurrence (or non-occurrence) of occupational accidents and injuries can shape perceptions of safety climate. Based on a literature review, the authors indicated that safety climate—a traditional *leading* indicator—can also be viewed as a *lagging* indicator. They argued that positioning safety climate as a *lagging* indicator recognizes the "continuity of understanding that employees have regarding safety in the workplace" (p. 738). The authors concluded that, "Safety climate is both a leading and a lagging indicator of safety-related events, which should continue to be recognized in the safety literature" (p. 738). Based on this line of reasoning, traditional *leading* and *lagging* indicators seem to be involved in a complex pattern of interdependence that is difficult to disentangle.

2.3. Alternative indicator frameworks

There has been at least one effort to develop an indicator framework void of the leading/lagging nuances. Juglaret et al. (2011) suggested that because HSMS are made up of numerous interacting sub-processes, or elements (e.g., change management, leadership development, contractor management, emergency preparedness, etc.) that comprise a unitary cycle, indicators should simply be selected to provide information on the distinct effectiveness of each. Juglaret et al. (2011) suggests that this approach may provide a more specific list of potential performance indicators which could, in turn, allow for the improved monitoring of each of the sub-process activities within the system. They further argue that performance knowledge corresponding to each of the system's sub-processes allows for precise and focused alterations which can be used to consciously improve overall system performance.

2.4. Research objectives

Given the lack of consensus regarding a critical aspect of H&S management, the current research set out to empirically explore the types of performance indicators currently being used in mining organizations using the approach proposed by Juglaret et al. (2011). We sought to answer the following research questions based on data collected from mine health and safety managers:

- What performance measures are being used by mining practitioners to measure HSMS effectiveness?
- What performance indicators cover the breadth of the P-D-C-A cycle?
- Based on the above results, what are the possible barriers and solutions involved with the practice of HSMS performance measurement?

3. Methods and materials

3.1. Survey instrument framework and development

A survey was designed to elicit textual data to explore the research objectives presented above. The survey, approved by the NIOSH Institutional Review Board, included a list of HSMS elements and practices and requested that participants provide information regarding optimal ways to measure the effectiveness of each practice, within each element. Given the mining-specific context of the current investigation, the U.S. National Mining Association's (NMA) CORESafety HSMS was used to inform the empirical investigation. The NMA CORESafety HSMS is comprised of 20 different elements to address the specific risks and hazards associated with mining (NMA CORESafety Handbook, nd). The NMA CORESafety handbook indicates that the 20-element system is consistent with the ANSI/AIHA Z-10 and OHSAS 18001 consensus standards and aligns its 20 elements under the traditional HSMS cycle as shown in Fig. 1.

The committee responsible for the development of NMA CORESafety HSMS developed a list of specific management practices corresponding to each of the 20 elements shown in Fig. 1. This HSMS framework informed the development of the survey for the current study for several reasons. First, CORESafety is a recently developed

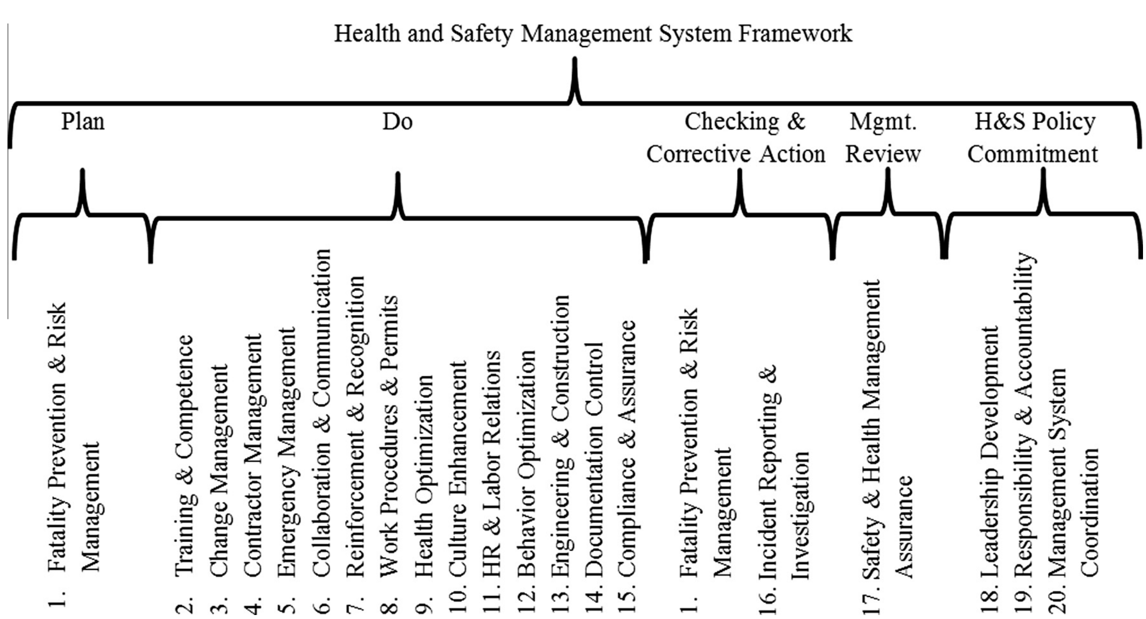

Fig. 1. Organization of CORESafety's 20 elements under the generalized HSMS system cycle.

Table 1
Example health and safety practices and corresponding elements included within the survey.

Element	Practice examples
Behavior optimization	• Educate employees as to the causes of safe and unsafe behavior • Encourage employees to intervene when they see a co-worker behaving unsafely • Ensure that employees' behavior is consistent with the critical behavior inventory • Ensure that employees' behavior is consistent with SOPs
Work procedures and permits	• Develop standard operating procedures (SOPs) for routine jobs from job risk assessments • Develop SOPs for non-routine, repeatable jobs from job risk assessments • Implement a work permit procedure for high-risk work • Review SOPs often and update when needed
Change management	• Define operational changes which require site leadership review (e.g. changes in process, equipment, physical environment, procedures, etc.) • Conduct risk assessments on proposed changes • Communicate changes to employees and contractors • Conduct pre-start-up safety reviews after changes are complete
Training and competence	• Conduct training needs assessments • Develop and deliver job specific training programs to new employees • Develop and deliver job specific training programs to employees transferring to different jobs • Develop and deliver ongoing refresher training
Occupational health	• Conduct industrial hygiene exposure assessments • Communicate health hazard assessments to employees in a timely manner • Utilize continuous industrial hygiene monitoring where appropriate • Implement medical surveillance for employees where appropriate
Leadership development	• Develop and communicate H&S leadership competencies for site leadership • Develop and offer H&S leadership development training programs • Require that site leadership develop personal development plans which include H&S metrics • Tie personal development plans to the performance management system

HSMS, meaning the content was deemed important and currently applicable for health and safety management in mining organizations. Second, CORESafety was developed in collaboration among mining stakeholders. As a result, mine organizations throughout the United States utilize this framework at a site-level. Last, this framework contained more specificity regarding the elements and practices advocated than alternative HSMS, allowing us to ask participants to provide examples of performance measurement around a more exhaustive list of health and safety management practices.

Generally, one CORESafety HSMS element is comprised of five to nine practices. We included all 20 elements and each of the practices that corresponded to the elements, which totaled to 133 complementing practices. These 133 practices, categorized under their respective elements, were compiled and provided to each participant. Corresponding to each of the 133 practices provided, each participant was asked the open-ended question, '*How do you measure the performance (i.e., effectiveness) of each of the following practices in your organization?*'. Participants were allotted a text box for each practice provided to type their feedback about how they measure performance. Table 1 provides examples of practices within various elements in which participants were asked to provide metrics throughout the survey. A list of the complete 133 practices that were used in the survey to probe performance measurement is available in the CORESafety handbook (coresafety.org).

However, we also acknowledged that not all participants may have implemented the full 20-element, 133 practice HSMS, even if they were aligned with the management strategy. Given the possibility that some of the practices provided were not identified as a strategic imperative within a particular participant's organization, the single question previously posed was qualified with: '*If you do not currently use the practice within your organization, please provide your opinion as to the best possible metric for the given practice.*'

The survey also asked participants to rate the degree to which each of the 20 elements and 133 practices is essential to prevent accidents and injuries. Participants were asked to rate each element and practice on a scale from one to five. These quantitative results, which informed the "most essential" HSMS elements and practices, are reported elsewhere (Yorio and Willmer, 2015). The qualitative performance indicators provided by participants were the focus in the current study and were qualitatively analyzed in an effort to further understand and organize HSMS performance measurement in this context.

3.2. Recruitment and sample

A convenience sampling approach was used to recruit subject matter experts at operating U.S. mine sites (i.e., practicing executives, managers, and professionals) (Miles and Huberman, 1994). Potential participants were identified via their publicly available affiliation with the CORESafety or similar HSMS framework. Once identified these individuals were contacted for participation. Recruitment occurred consistently and was ongoing between June 2013 and June 2014. Individuals were contacted via personal telephone and email communications from our personal contacts database. Although it took varying amounts of time for individuals to respond to queries, eventually every individual who was contacted either agreed to participate or identified someone within the organization who would be more appropriate to complete the survey questions.

Upon expressing interest, potential participants were e-mailed a secure internet link to access the survey. Before consenting to participate, we assured each individual that their demographic information and the mine's specific demographic information would be kept private. After opting to participate via a consent page, participants could access and complete the survey. Participants could break from completing the survey and pick up where they left off at a later time so they could critically consider each element and practice presented. In general, the survey took participants approximately two hours to complete. The sample ($n = 9$) included experienced individuals who were identified as health and safety managers/professionals for their specific mine organization and responsible for developing and implementing their HSMS at a site-level. Among the nine participants, they represented an

Table 2
Example of behavioral observations as a common performance indicator across practices and elements.

HSMS element & participant	HSMS practice	Examples of performance indicators provided by participants to assess the respective practice
Behavior optimization, participant 3	Ensure that employees' behavior is consistent with the critical behavior inventory	*"Document that an observation checklist is used on incidents at the operation and is used for employee observations."*
Behavior optimization, participant 5	Ensure that employees' behavior is consistent with SOPs	*"Job safety observation. Use SOPs for 'job-specific' observations."*
Work procedures & permits, participant 4	Develop SOPs for routine jobs from job risk assessments	*"Job observation to determine if SOP is used and followed."*
Work procedures & permits, participant 5	SOPs contain warnings about the potential consequences of deviation	*"Observation of employees if a warning is in place and acknowledged."*
Contractor management, participant 5	Ensure contractor work adheres to compliance obligations	*"Regularly scheduled inspections/job observations."*
Leadership development, participant 6	Develop and communicate H&S leadership competencies for site leadership	*"Observation of behaviors."*
Safety accountability & responsibility, participant 5	Use positive reinforcements (carrots) specific to individual employee S&H roles and responsibilities	*"Observe for increased safety behaviors and compliance."*
Safety accountability & responsibility, participant 3	Provide employees with the tools and equipment needed to fulfill their S&H roles and responsibilities	*"Evidence from observations or injury reports to determine if lack of proper tools & equipment is a factor in injuries or risks."*

even distribution from underground coal, surface coal, and metal/non-metal mining operations.

Recruitment was ongoing until an appropriate sample size was reached to answer the research questions. In the current study, the Delphi technique was used to initially recruit an appropriate sample. The Delphi method is a technique used to determine if consensus in expert opinion exists on a particular issue (Hasson et al., 2000). This technique has been used in previous health and safety research to refine and/or condense a large body of information. For example, Donaldson et al. (2013) used this method to identify health and safety management system practices most applicable to sports management organizations and Hardison et al. (2014) used this method to prioritize supervisor competencies for effective site safety. Due to the related nature of the data, this method was deemed appropriate to assess the similarities and differences of these data as it was received. Additionally, an examination of the data revealed that qualitative saturation of the feedback was occurring early in the responses, with similar information being provided by participants (Corbin and Strauss, 2008). Data was repetitive and themes could be clearly distinguished by the time the ninth data set was reviewed. At this point the researchers determined a sample size of 9 was appropriate for the exploratory analysis and recruitment ended.

3.3. Data analysis

The responses were considered using a qualitative content analysis approach. Qualitative content analysis (QCA) subjectively and formatively interprets the "content of text data through the systematic classification process of coding and identifying themes or patterns" (Hsieh and Shannon, 2005, p. 1278). More specifically, QCA is a method that helps systematically describe and categorize qualitative material – which is classified around a categorizing coding framework that emerges from the data during initial analysis (Schreier, 2012). QCA was deemed an appropriate data analysis method because this approach allows for intense focus on selected aspects of the data which in this case, were the types of the performance measures provided by participants among the various HSMS practices. The sense-making process involved in the QCA helped us identify core meanings in the data and understand the realities of the HSMS experts in a subjective, scientific manner (Patton, 2002; Berg, 2001). For more information about QCA as a data analysis method, refer to Krippendorff (2007). The analysis steps are detailed below.

3.3.1. Initial and focused coding by HSMS element

A primary document that contained all of the survey responses was saved in ATLAS.ti.7 to display, browse, organize, and eventually code the textual data into specific groups and categories. We coded the data together in order to discuss and resolve any discrepancies immediately. Working together also ensured consistency of the coding, the development of coding categories, and the definitions and rules for assigned codes (Weber, 1990). First, an initial coding of the textual data, in which each piece of text was read word by word (Charmaz, 2006) allowed us to account for each performance indicator that participants provided within each of the 20 HSMS elements. In total, we coded 1215 responses as part of the analysis process because some participants provided more than one performance measurement for a practice provided.

Next, we used a focused coding approach to begin identifying potentially useful concepts provided in response to each HSMS practice (Charmaz, 2006). During focused coding, we used the constant comparative method (Glaser and Strauss, 1967; Strauss, 1987; Glaser, 1992) to reflect on each performance indicator, refine performance indicators that continued to emerge across the survey, and further examine whether certain indicators were similar to one another that appeared in a different element throughout the survey responses. During this process we began to collapse the performance indicators across the practices within each element, and across the elements. As Table 2 illustrates, during focused coding, various specific performance indicator categories surfaced that were applicable across numerous HSMS elements. The table shows that participants' suggested the *results of behavior observations* could provide performance information regarding the practices included across a variety of the HSMS elements, including, but not limited to: Behavior Optimization, Work Procedures and Permits, Contractor Management, Leadership Development, and Safety Accountability and Responsibility.

3.3.2. Refining performance indicators and indicator categories

We continued to apply the constant comparative method to more systematically compare and understand the theoretical properties and purposes of each performance indicator pattern that emerged in the data. This phase was particularly important to help prevent drifting into an idiosyncratic sense of what the indicators meant with respect to each HSMS element (Schilling, 2006). Rather, we were able to remain open to the performance indicator data, apart from the elements. As we continued to code and analyze the indicators, it became possible to begin linking specific indicators together around central categories (Glaser and Strauss, 1967). Specifically, similar performance indicator trends emerged

Table 3
Results of the 22 performance measurements and their respective indicator categories.

Indicator category	22 Specific performance measurements
Interventions	Number of communications and meetings Number of investigations and reviews Number of corrective actions completed Number of disciplinary actions taken Number of risk management studies (e.g., hazard inspections and audits) Number of hazards or suggestions reported Number of medical surveillance or substance abuse testing activities Number of behavioral observations Number of rewards allocated or withheld Amount of training (frequency, hours) Workforce participation information (rates, count) Number of interviews that probe H&S issues Number of surveys that probe H&S issues
Organizational performance	Number of and type of citations/compliance rates Number of and types of injuries and illnesses Number of and types of near misses Results of root cause analysis of injuries and illnesses Results of behavior observations Results of performance evaluations Results of medical surveillance or substance abuse testing activities Results of risk management studies (e.g., hazard inspections and audits) Results of workforce H&S knowledge assessments Interviews that probe performance Surveys that probe performance
Worker performance	Number of and type of citations/compliance rates Number of and types of injuries and illnesses Number of and types of near misses Results of root cause analysis of injuries and illnesses Results of behavior observations Results of performance evaluations Results of medical surveillance or substance abuse testing activities Results of H&S knowledge assessments Interviews that probe performance Surveys that probe performance

throughout participants' responses to each of the 133 practices provided in the survey. Upon coding all of the performance indicators discussed for each practice, by each participant, we were able to clearly distinguish and group some of these key indicators across participants and their respective HSMS.

Throughout the process of content analyzing, coding, and further comparing and contrasting the performance measurements provided for each of the 133 practices, we identified and narrowed the list of over 1200 performance measures to 22 unique indicators. Upon identifying these 22 overarching indicators, we re-reviewed the results of our coding to characterize measurement and methodological patterns among the 22 indicators (Charmaz, 2006). These 22 unique performance indicators (listed in Table 3) are further discussed in the results section as they relate to specific measurement methods for assessing HSMS effectiveness.

4. Results

Results are reported based on the themes and patterns that emerged throughout the data. Individual responses are not reported in an effort to protect the anonymity of participants. However, because the primary goal of this study was to better understand the types of effective performance indicators being used to assess an HSMS, noting the primary responses and trends within the data is the most useful tactic for answering the research objectives.

4.1. A new categorizing framework

Although the purpose of the current study was to identify methods that mine health and safety managers use to assess the effectiveness of their HSMS, a closer examination of these performance indicators also supports the exploration of a new framework for HSMS performance measurement. Upon completing the final phase of the analysis – further analyzing and categorizing the 22 performance indicator types – three general types of performance measurements emerged. These three overarching performance measures ultimately revealed a new categorizing framework: *organizational performance, worker performance*, and *interventions.* All 22 performance indicators that surfaced during the analysis fit within these three categories (Table 3).

4.1.1. Organizational and worker performance

Organizational performance and *worker performance* emerged as two primary indicator categories, each encompassing eleven of the measures that could be used to evaluate the individual and/or the organizations' performance. Participant responses suggested that *worker performance* was measured through information related to, for example, the results of behavioral observations, and assessments of health and safety knowledge through tests, standardized performance evaluations, or the results of medical surveillance. From an *organizational performance* perspective, measures provided included the results of root cause analyses, hazard inspections and audits, and the number of injuries and illnesses within the organization.

The examples of *organizational* and *worker performance* indicators outlined above and in Table 3 are designed to answer questions such as 'are workers behaving safely?'; 'do workers have adequate H&S knowledge?'; and 'is the workplace free from hazardous conditions?'. Thus, as opposed to only tracking the number of, for example, hazard inspections, these indicators may help provide a more in-depth picture of conclusions based on results of such inspections.

4.1.2. Interventions

Interventions also emerged as a distinct indicator category that focused on the application of activities to measure performance. This category encompassed the types of specific measures related to the number of focused HSMS-related activities that take place within the organization. Thirteen specific performance measures were included within the *interventions* category. These consisted of specific activities that could be prompted by organizational leaders, managers, supervisors, or workers and included, for example: the number of disciplinary actions taken, the number of hazard inspections and audits performed, the number of behavior observations, the amount and type of training provided, and the number of corrective actions completed. Participant feedback suggests that *intervention* performance indicators could be used to assess the full range of elements within the HSMS system cycle.

The *intervention* activities had less to do with the outcomes or the evaluation of specific activities. Rather, *interventions*, as discussed in the survey responses, captured the number/frequency/amount of health and safety actions/behaviors executed within the organization. For instance, participants noted that communication with workers could vary on topics from informing new rules to correcting an unsafe behavior observed on site. In this case, the act of communicating/meeting/reviewing with workers is the action/behavior executed by organizational leaders. Also, participants noted that the number of hazards or health and safety suggestions from the workforce could be used as an indicator of effectiveness for various elements within the HSMS. In this case, a specific form of worker health and safety behavior is tracked as an *intervention* rather than merely actions prompted by management.

5. Discussion

The three performance indicators, *interventions, organizational performance,* and *worker performance,* share similarities to aspects of current literature; however, there are also important differences. The following sections discuss overlap, differences, and strengths in comparison to other indicator frameworks.

5.1. New organizing framework's overlap with other indicator frameworks

On its face, the three indicator categories found in the current empirical exploration share commonalities with aforementioned frameworks. For example, *predictive proactive* indicators (Körvers and Sonnemans, 2008), *proactive activity* indicators (Laitinen et al., 2013), and *lead drive* indicators (Reiman and Pietikäinen's, 2012) share commonalities with the *interventions* indicator category that emerged in the current study. All capture the frequency or number of specific types of activities designed to achieve safety and health within the organization. Also, the *monitoring proactive*, *proactive outcome,* and *lead monitor* indicators proposed in previous frameworks share similarities with the *organizational* and *worker performance* indicator types found in the current study in that they all provide information related to levels of safety performance or glimpses into the current state of safety within the organization. A closer look within the indicator types, however, reveals subtle differences that illuminate the strengths of the proposed framework.

5.2. Possible strengths of the three indicator categories

First, the major difference between previous categorizing frameworks and the categories found in the current study is that they do not integrate the *leading/lagging* nuances. In this way, the categories are not prone to limitations associated with the use of these terms and are consistent with previous arguments (Hollnagel, 2004; Dekker, 2005; Reiman and Pietikäinen, 2012) that suggest the *leading/lagging* framework may not optimally capture the nature of health and safety as an emergent property of the entire organizational system. In contrast, the current indicator categories are neither explicitly categorized based on time nor causally related to any other general or specific indicator category. Therefore, one category should not be understood to precede another as in a causal framework.

Of course, even given this framework, theories and hypotheses of causation can still be made between and within the indicator categories. However, freeing the category definitions from explicit causation hypotheses also eliminates the problem of reverse causation that can be argued for nearly any hypothesized time-dependent, causal pathway. In a sense, then, every general indicator category, along with all of the associated specific performance indicators are needed to provide a snapshot picture of the state of safety and health in the organization—thereby acknowledging the complexity with which organizational safety emerges. Thus, the three indicators that emerged in this analysis support a "systemic and dynamic view of organizational safety" (Reiman and Pietikäinen, 2012, p. 1994) in that they provide a collective view of how the HSMS is operating at a certain point in time and guidance in what aspects of the entire system need improved.

Further, because the NMA CORESafety HSMS incorporates 'employee involvement' practices within many of its elements, the *interventions* category uncovered in the current study explicitly includes activities prompted by people who hold positions within each of the levels of the organization (i.e., from top organizational leaders to front-line supervisors and workers). This stands in distinction from the *predictive proactive*, *proactive activity,* and *lead monitor* indicators which are theoretically designed to measure managerially-driven activities. This nuance may seem trivial, but as the importance of employee involvement has been noted in nearly all HSMS consensus standards and applicable performance indicator categories, may be increasingly critical in HSMS effectiveness.

On a more practical level, the proposed categorizing framework may contribute to the advancement of HSMS performance measurement while helping to somewhat cognitively simplify the assessment process. Given that each of the categories that emerged span across activities included in each phase of the HSMS Plan-Do-Check-Act cycle, these three indicator types could be used continuously and interdependently to monitor, for example, the performance of risk management planning activities as well as management review and incident investigation activities (i.e., activities that occur prior to and after safety incidents). Further, reducing measurement efforts from hundreds to 22 specific metrics significantly helps reduce data overload and may enhance the quality of data analysis and decision-making on behalf of health and safety managers (Podgórski, 2015; Hwang and Lin, 1999).

Another strength of the current results indicate that the new indicator categories may enhance practitioners' propensity to follow the Plan-Do-Check-Act (P-D-C-A) cycle when choosing performance measures for HSMS activities. To illustrate how practitioners can begin choosing performance indicators that may be appropriate for a specific activity within their P-D-C-A cycle, we use activities associated with safety climate and/or its enhancement as an example of a specific health and safety management activity. We use these climate enhancement activities as an example based on participants' discussion of surveys as a useful metric and also due to the arguments in the literature for safety climate as both a leading and lagging performance indicator. As illustrated in Fig. 2, the use of all three indicator categories is applied when determining and responding to an organization's safety climate or enhancement. The *intervention* indicators help measure and evaluate goals associated with the planning phase of understanding and measuring safety climate as part of an HSMS activity (i.e., behavioral observations), as well as those associated with evaluating the activity goals related to results of a safety climate assessment including follow-up corrective actions, and management reviews (i.e., number of corrective actions completed). Similarly, *organizational* and *worker performance* indicators capture the results/outcomes of activities that span all of the phases of the cycle.

Therefore, within each indicator category, practitioners can look at the specific performance indicators and choose ones that are aligned with their current stage of the health and safety management process. This step helps practitioners ensure they are progressing through the complete P-D-C-A cycle and not overlooking necessary planning or response to prevent or mitigate a health/safety issue. In other words, this process of aligning indicators and activities within the cycle helps ensure balance and representation while helping to minimize data overload by using too many metrics at an inopportune stage of a specific activity.

5.3. Areas of divergence among participant responses: the need for balanced and representative measurement

Revealing these three categories and respective indicators is only part of the challenge to HSMS performance measurement. As observed in the current study and consistent with previous research, a critical need also includes guidance and measurement criteria that can be used to assess and improve performance (Robson and Bigelow, 2010; Robson et al., 2012; Podgórski,

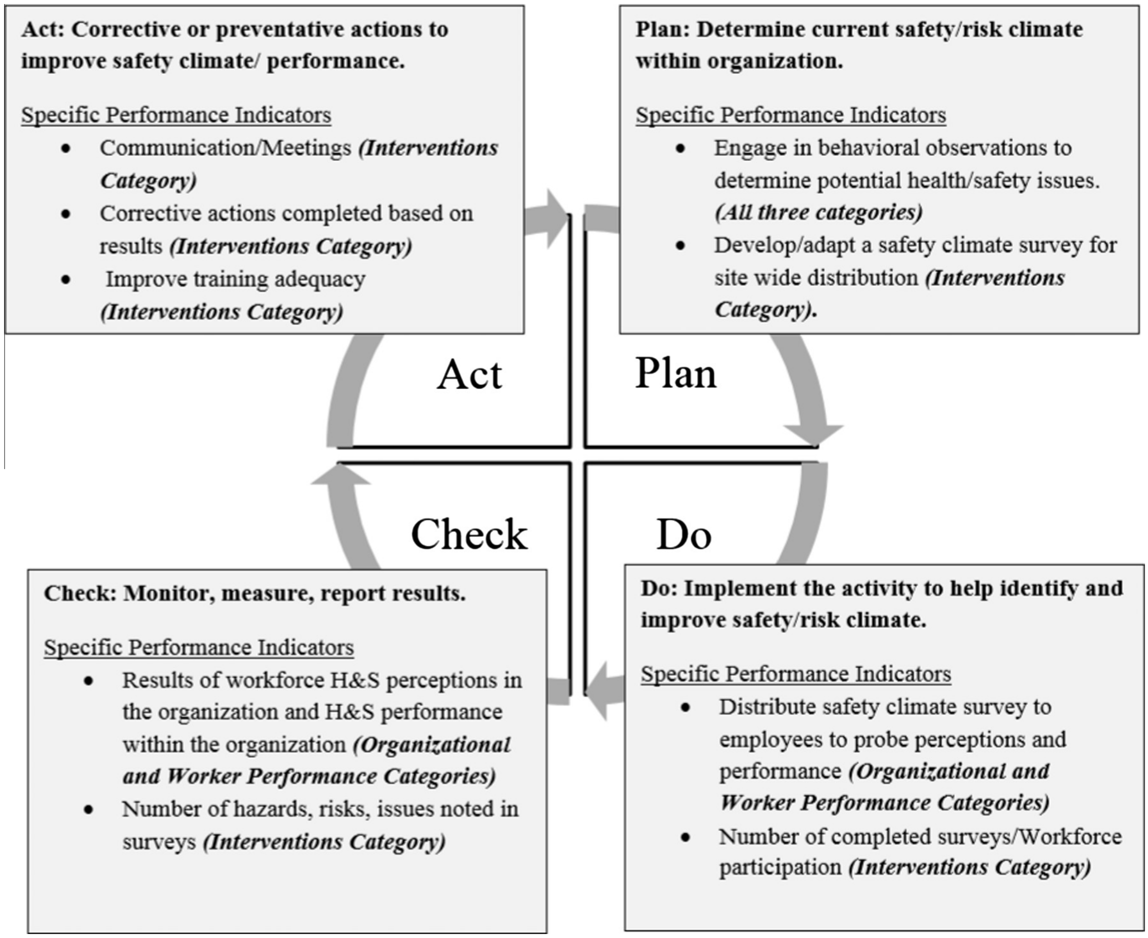

Fig. 2. Use of specific performance indicators within three categories to engage in safety climate activities across P-D-C-A cycle.

2015). Although similarities across participant responses allowed for the development of the three primary indicator categories, two primary, interrelated differences were noted: (1) objective versus subjective metrics; and (2) quantitative versus qualitative measurement. These results suggest a need for (1) the value of balancing the use of both quantitative (e.g., the number of safety audits conducted or corrective actions provided) and qualitative performance indicators (e.g., results of safety interviews); and (2) the importance of objective measurements but also not neglecting subjective assessments when using performance indicators. Fig. 3 illustrates how a mixture of measurement methods and objective/-subjective indicators can be consciously integrated into an HSMS performance management program.

5.3.1. Quantitative vs qualitative measurements

First, the results of the current analysis support a consistent and balanced use of both quantitative and qualitative measurement methods. In other words, it should be a best practice to both numerically represent and describe when it comes to interpreting and following through on performance measurement. As numeric representations can facilitate ease in reporting, they are often limited by a lack of context (e.g., the nature or severity of hazards found and/or the process complicating circumstances). Relying only on numeric representations of performance can, in some cases, lead to overlooking important mitigating and/or process complicating circumstances which may be the key to determining and correcting root causes to poor HSMS performance.

Further, although a quantitative measurement can be used to reveal causal relationships, HSMS performance information derived through qualitative methods can help understand the more intangible and tacit aspects of the HSMS including attitudes, beliefs, and behaviors (International Council on Mining and Metals, 2012). For example, referencing the example earlier, Fig. 3 depicts performance indicators that may be used to directly assess safety climate or the activities associated with its enhancement. As shown, these indicators represent a mix of qualitative and quantitative measurements. In the planning phase, qualitative behavioral observations to inform the development of a survey would be appropriate, followed by the distribution and analysis of the survey data, a quantitative measurement. Moving into the action phase of the cycle, engaging in increased training or corrective actions could be both quantitative and qualitative. Therefore, applying a balance of quantitative and qualitative measurements is an aspect of performance measurement that should be consciously considered when choosing how to evaluate health and safety processes. Used simultaneously, a mixed-methods approach could offer a more comprehensive vantage point of health and safety performance.

5.3.2. Objective vs subjective measurements

Another way to ensure an accurate viewpoint of an HSMS is to engage in indicators that reveal objective and subjective performance. While some of the responses reflected the use of objective knowledge assessments and hazard assessments, for example, many advocated the use of psychological perceptions gathered through interviews and surveys for the same or similar practices.

Regarding objective assessments, one participant said, "Indicators should be styled to specific processes, but be objectively measured and reported (i.e., injuries, illnesses, near misses,

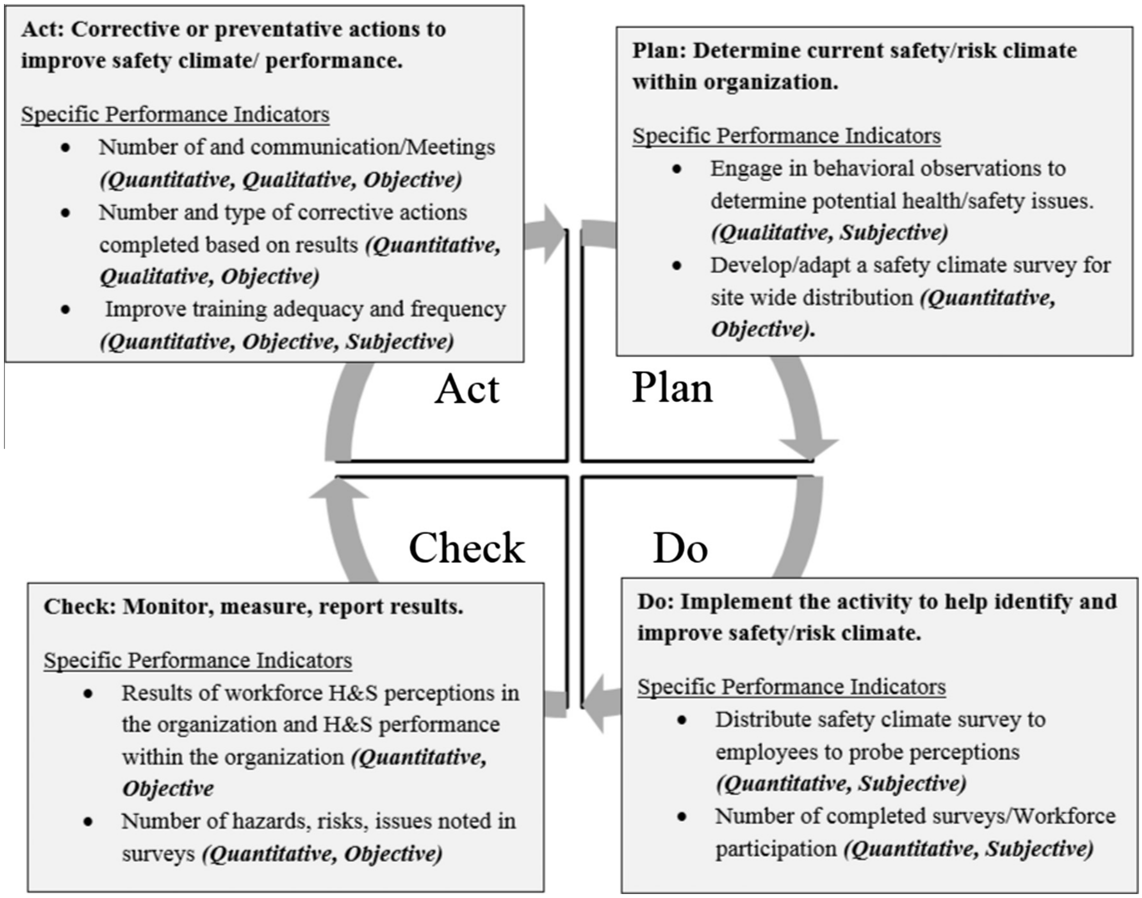

Fig. 3. Triangulated use of multiple methods and measures to engage in safety climate activities across P-D-C-A cycle.

standardized observation count/results, % participation, % trained, etc.).” An advantage of objective performance indicators is that they defy interpretation. Participant responses suggested that objective worker performance can be measured through information related to, for example, the results of behavioral observations, assessments of H&S knowledge through tests, standardized performance evaluations, or the results of medical surveillance. From an organizational performance perspective, objective measures provided by participants included the results of root cause analyses, hazard inspections and audits, and the number of injuries and illnesses within the organization. Participants identified this indicator category as useful to assess practices corresponding to each of the elements within the P-D-C-A cycle.

However, results also revealed the importance of subjective perceptions to evaluate the performance of the HSMS. In general, participants expressed the value in understanding how leaders, managers, supervisors, and workers view health and safety within the mining organization. Primarily, participants noted using interviews and/or surveys as a way to measure these subjective perceptions. Many responses advocated the use of deriving quantitative scores from a survey of questions chosen to measure the effectiveness of the practice in question or to measure psychometrically tested items corresponding to theoretical constructs. For example, they noted that subjective perceptions could be used to assess whether or not workers were provided with adequate tools and equipment they need to execute their job tasks in a safe and healthful manner (i.e., corresponding to the responsibility and accountability element). Further, participants suggested their use for an initial baseline measurement and to track trends in desired safety culture characteristics within the organization. Participants also noted, for example, that subjective perceptions could be used to evaluate the effectiveness of the behavior optimization element by asking questions related to workers’ willingness and ability to work safely and encourage others to do so. In all, throughout the 1215 performance indicators, it was evident that participants felt that survey and/or interview questions could be asked of leaders, managers, supervisors, and workers in order to evaluate how well each element within the HSMS was performing.

Although some have argued that the validity of the information provided through subjective perceptions can be lower and its link to performance measurement is questionable (e.g., Kongsvik et al., 2010), if one of the primary objectives of the HSMS is to engage workers, then using subjective perceptions to evaluate its performance should be embraced by practitioners. Other theory and research also suggests that subjective perceptions may play a unique and important role in H&S performance measurement. For example, Bergh et al. (2014) developed a performance indicator for psychosocial risk in which subjective perceptions played a unique and key role. Also for example, Yorio et al. (2015) suggest that worker perceptions and interpretations of the HSMS can be an important indicator of the effectiveness of the set of practices chosen.

6. Limitations, conclusions, and directions for future research

A few limitations of this empirical investigation need to be noted. First, the participant sample was mining-specific, and the results arguably generalize to that industry. Second, although grounded in the traditional P-D-C-A cycle, the NMA CORESafety framework—an HSMS consistent with ANSI/AIHA Z-10 and OHSAS 18001—was used to develop the practices *a priori* specified in the survey. Therefore, the results may be further restricted to mining organizations who align their HSMS strategy with those frameworks. Thus, we believe future research may consist of similar inductive, exploratory designs in distinct industries and HSMS strategies.

Additionally, although the results of this study reduced the number of key performance indicators to 22, more research using these indicators to determine their sufficiency for continuous assessment purposes within health and safety management is necessary (Podgórski, 2015). All of these indicators may not be feasible for any one practitioner. As one example, engaging in criteria to help practitioners identify specific performance indicators, may be helpful. The SMART (i.e. Specific, Measurable, Achievable, Relevant, and Time-bound) criteria is one tool that can be applied to further select key performance indicators for measurement (e.g., Carlucci, 2010; Kjellen, 2009). For a review of criteria for selection of key performance indicators, see Podgórski (2015). Specifically, this article helped add to information about the specific measurement criteria for identifying performance indicators – or the "S" and "M" in the SMART framework.

Given the noted limitations, in lieu of the ongoing, open-ended discussions related to HSMS performance measurement in organizations, the goal of the current research was to identify performance measures that may contribute to assessing HSMS effectiveness in the future. We interjected a creative approach to better identifying and understanding these phenomena. Through this inductive effort, however, we also uncovered a novel way to look at and think about performance indicators in organizations. This study reveals important criteria to consider in terms of specificity, measurement, and pragmatism when choosing performance indicators. The results provide a precise link back to the original P-D-C-A cycle which helps guide the selection of fitting performance indicators at the right phase of continuous assessment and improvement. Additionally, choosing indicators within the P-D-C-A cycle helps bring a sense of simplicity back to HSMS assessment for practitioners. Regardless, future research would benefit from assessing these identified indicators against additional criteria to help further measure performance. The efficacy and utility of this framework is subject to debate and further empirical testing. To that end, we need to ask ourselves if the existing metric frameworks, such as the *leading/lagging* categories, enhance our understanding of and the practical use of organizational health and safety performance management, or if the terms invoke unnecessary and inefficient theoretical and practical confusion.

The results also show the importance of balancing the types of measurement and methods used when choosing performance indicators. Specifically, engaging in a consistent selection of both quantitative and qualitative performance indicators when proceeding through the P-D-C-A cycle may help ensure a representative view of the health and safety problem. Also taking into account objective and subjective measures allows for more information that can inform future improvements in health and safety processes.

In conclusion, consistent with the notions of other H&S academics and theorists, there is no question that performance management of an HSMS is a critical and pressing issue for organizations in addition to understanding the most effective ways to implement the system (i.e., we need to understand more about the most effective ways to enhance the behavioral execution of H&S strategy). Although this study provided an overarching view of general performance indicator categories that can be utilized, future research is needed that continues to build upon and inform health and safety managers and practitioners how these indicators can be developed and implemented over time. In response, subsequent studies should plan to move beyond the "*what*" and focus on the "*how*" in order to more specifically inform HSMS implementation and evaluation. As Bowen and Ostroff (2004) argue, the visibility of the strategic practices in question (i.e., the degree to which the HSMS are salient and readily observable) is a basic prerequisite for worker interpretation and sensemaking of the strategy. Therefore, an important aspect of implementation is the use, promotion, and effective management of a well-chosen set of performance indicators. This argument is consistent with Kaplan and Norton's (1996) assertion that despite the best intentions, lofty statements of those at the top and verbal calls to action "don't translate easily into operational terms that provide useful guides to action at the local level. For people to act on the words in vision and strategy statements, those statements must be expressed as an integrated set of objectives and measures" (p. 4). Thus, we encourage future research into the subject and hope that the results of this research spark new and creative ways to approach the underpinning theory as well as the practical utility of performance indicators for health and safety management.

References

ANSI/AIHA Z-10:2005. American National Standards Institute, 2012. American National Standard for Occupational Health and Safety Management Systems. American National Standards Institute, Washington, DC.

Berg, B.L., 2001. Qualitative Research Methods for the Social Sciences. Allyn and Bacon, Boston.

Bergh, L.I.V., Hinna, S., Leka, S., Jain, A., 2014. Developing a performance indicator for psychosocial risk in the oil and gas industry. Safety Sci. 62, 98–106.

Bowen, D.E., Ostroff, C., 2004. Understanding HRM–firm performance linkages: the role of the "strength" of the HRM system. Acad. Manage. Rev. 29 (2), 203–221.

British Standards Institute, 2007. Occupational Health and Safety Assessment Series18001, Occupational Health and Safety Management system Standard – Require-ments. British Standards Institute, London.

Carlucci, D., 2010. Evaluating and selecting key performance indicators: an ANP-based model. Meas. Bus. Excellence 14 (2), 66–76.

Charmaz, K., 2006. Constructing Grounded Theory: a Practical Guide Through Qualitative Analysis. Sage Publications, Thousand Oaks, CA.

Corbin, J.A., Strauss, A., 2008. Basics of Qualitative Research, third ed. Sage, Thousand Oaks, CA.

CORESafety, nd. CORESafety framework. Handbook can be retrieved from http://www.coresafety.org/coresafety-framework/handbook/.

Dekker, S., 2005. Ten Questions about Human Error, a New View of Human Factors and System Safety. Lawrence Erlbaum, New Jersey.

Diebold, F.X., Rudebusch, G.D., 1994. Measuring Business Cycles: A Modern Perspective (No. w4643). National Bureau of Economic Research.

Donaldson, A., Borys, D., Finch, C.F., 2013. Understanding safety management system applicability in community sport. Safety Sci. 60, 95–104.

Glaser, B.G., 1992. Emergence vs Forcing: Basics of Grounded Theory Analysis. Sociology Press, Mill Valley, CA.

Glaser, B.G., Strauss, A.L., 1967. The discovery of grounded theory: strategies for qualitative research. Aldine, New York.

Haight, J.M., Yorio, P.L., Rost, K.A., Willmer, D.R., 2014. Health and Safety Management Systems – a Comparative Analysis of Content and Impact". Professional Safety – J. Am. Soc. Safety Eng. 59 (5), 44–51.

Hale, A., 2009. Why safety performance indicators? Safety Sci. 47, 479–480.

Hardison, D., Behm, M., Hallowell, M.R., Fonooni, H., 2014. Identifying construction supervisor competencies for effective site safety. Safety Sci. 65, 45–53.

Hasson, F., Keeney, S., McKenna, H., 2000. Research guidelines for the Delphi survey technique. J. Adv. Nurs. 32 (4), 1008–1015.

Hollnagel, E., 2004. Barriers and Accident Prevention. Ashgate, Aldershot.

Hopkins, A., 2009. Thinking about process safety indicators. Safety Sci. 47, 460–465.

Hsieh, H.-F., Shannon, S.E., 2005. Three approaches to qualitative content analysis. Qual. Health Res. 15 (9), 1277–1288.

Hwang, M.I., Lin, J.W., 1999. Information dimension, information overload and decision quality. J. Inform. Sci. 25 (3), 213–218.

International Council on Mining & Metals, 2012. Overview of leading indicators for occupational health and safety in mining. November, 2012. Retrieved from: http://www.icmm.com/leading-indicators.

Janicak, C.A., 2011. Safety Metrics: Tools and Techniques for Measuring Safety Performance, 2nd ed. Government Institutes, Scarecrow Press Inc, Lantham, MD.

Juglaret, F., Rallo, J. M., Textoris, R., Guarnieri, F., Garbolino, E., 2011. New Balanced Scorecard leading indicators to monitor performance variability in OHS management systems. In: Proceedings of the Fourth Resilience Engineering Symposium.

Kaplan, R.S., Norton, D.P., 1996. Using the balanced scorecard as a strategic management system. Harvard Bus. Rev., 75–85 (Jan-Feb., 1996).

Kjellen, U., 2009. The safety measurement problem revisited. Safety Sci. 47 (4), 486–489.

Kongsvik, T., Almklov, P., Fenstad, J., 2010. Organisational safety indicators: some conceptual considerations and a supplementary qualitative approach. Safety Sci. 48 (10), 1402–1411.

Körvers, P.M.W., Sonnemans, P.J.M., 2008. Accidents: a discrepancy between indicators and facts! Safety Sci. 46 (7), 1067–1077.

Krippendorff, K., 2007. Content analysis, an Introduction to its Methodology, second ed. Sage Publications, Thousand Oaks, CA.
Laitinen, H., Vuorinen, M., Simola, A., Yrjänheikki, E., 2013. Observation-based proactive OHS outcome indicators–validity of the Elmeri+ method. Safety Sci. 54, 69–79.
Miles, M., Huberman, A.M., 1994. Qualitative Data Analysis. Sage Publications, Thousand Oaks, CA.
O'Neill, S., Martinov-Bennie, N., Cheung, A., 2013. Issues in the measurement and reporting of work health and safety performance: A review. A report co-funded by Safe Work Australia, Safety Institute of Australia and CPA Australia. Retrieved from: http://www.safeworkaustralia.gov.au/sites/swa/about/publications/pages/issues-measurement-reporting-whs-performance-review.
Patton, M.Q., 2002. Qualitative Research and Evaluation Methods. Sage, Thousand Oaks, CA.
Payne, S.C., Bergman, M.E., Beus, J.M., Rodriguez, J.M., Henning, J.B., 2009. Safety climate: leading or lagging indicator of safety outcomes? J. Loss Prev. Process Ind. 22, 735–739.
Petersen, D., 2001. The safety scorecard: using multiple measures to judge safety system effectiveness. Occupational Hazards 5, 54–57.
Podgórski, D., 2015. Measuring operational performance of OSH management system – a demonstration of AHP-based selection of leading key performance indicators. Safety Sci. 73, 146–166.
Reiman, T., Pietikäinen, E., 2012. Leading indicators of system safety–Monitoring and driving the organizational safety potential. Safety Sci. 50 (10), 1993–2000.
Robson, L.S., Bigelow, P.L., 2010. Measurement properties of occupational health and safety management audits: a systematic literature search and traditional literature synthesis. Can. J. Public Health 101 (2), S34–S40.
Robson, L.S., Macdonald, S., Gray, G.C., Van Eerd, D.L., Bigelow, P.L., 2012. A descriptive study of OHS management auditing methods used by public sector organizations conducting audits of workplaces: implications for audit reliability and validity. Safety Sci. 50 (2), 181–189.
Schilling, J., 2006. On the pragmatics of qualitative assessment: designing the process for content analysis. Eur. J. Psychol. Assess. 22 (1), 28–37.
Schreier, M., 2012. Qualitative Content Analysis in Practice. Sage, Thousand Oaks, IL.
Shishkin, J., 1961. Signals of Recession and Recovery, NBER Occasional Paper #77. NBER, New York.
Strauss, A.L., 1987. Qualitative Analysis for Social Scientists. Cambridge University Press, Cambridge, United Kingdown.
U.S. National Mining Association's CORESafety Handbook. Retrieved from: http://www.coresafety.org/download/CORESafety-Handbook.pdf.
Wachter, J.K., 2012. Trailing safety indicators. Professional Safety 57 (4), 48–60.
Weber, R.P., 1990. Basic Content Analysis. Sage Publications, Newbury Park, CA.
Yorio, P.L., Willmer, D.R., 2015. Explorations in pursuit of a risk-based health and safety management systems. In: Society for Mining, Metallurgy, and Exploration Annual Meeting, Feb. 15–18, 2015, Denver, CO.
Yorio, P.L., Willmer, D.R., Moore, S.M., 2015. Health and safety management systems through a multilevel and strategic management perspective: theoretical and empirical considerations. Safety Sci. 72, 221–228.