UNISINOS UNIVERSITY
GRADUATE PROGRAM IN PRODUCTION ENGINEERING AND SYSTEMS
DOCTORAL DEGREE

RODRIGO FRANK DE SOUZA GOMES

BEYOND EFFICACY IN THE EVALUATION OF OCCUPATIONAL HEALTH AND SAFETY MANAGEMENT SYSTEMS (OHSMS): A PERSPECTIVE BASED ON EFFICIENCY ANALYSIS

São Leopoldo
2022

RODRIGO FRANK DE SOUZA GOMES

BEYOND EFFICACY IN THE EVALUATION OF OCCUPATIONAL HEALTH AND SAFETY MANAGEMENT SYSTEMS (OHSMS): A PERSPECTIVE BASED ON EFFICIENCY ANALYSIS

A thesis submitted to the UNISINOS University for the doctoral degree in Production and Systems Engineering.

Supervisor: Prof. Daniel P. Lacerda, D.Sc.

Co-supervisor: Prof. Ana S. Camanho, Ph. D.

São Leopoldo

2022

G633b Gomes, Rodrigo Frank de Souza.
Beyond efficacy in the evaluation of occupational health and safety management systems (OHSMS) : a perspective based on efficiency analysis / by Rodrigo Frank de Souza Gomes. – 2022.
196 p. : il. ; 30 cm.

Thesis (doctoral degree) — UNISINOS University, Graduate Program in Production Engineering and Systems, São Leopoldo, RS, 2022.
Supervisor: Daniel P. Lacerda, D.Sc.
Co-supervisor: Ana S. Camanho, Ph.D.

1. Occupational health and safety management system (OHSMS). 2. Work. 3. Accidents. 4. Efficiency analysis. 5. Data envelopment analysis. I. Title.

CDU: 65:331.45

Cataloging in Publication (CIP):
Librarian Alessandro Dietrich - CRB 10/2338

DECLARATIONS

- Section 3 is based on the paper "***Safety at Work: a complex or an exceedingly simple matter?***". The paper is co-authored with Leandro Gauss, Fabio Sartori Piran, and Daniel Pacheco Lacerda.

- Section 4 is based on the paper "***Performance measurement in Occupational Health and Safety Management Systems: an analysis of critical elements comprising assessment instruments***". The paper is co-authored with Daniel Pacheco Lacerda, Leandro Gaus, and Ana P.C. Ermel.

- Section 5 is based on the paper "***Measuring Efficiency of Safe Work Environment from the perspective of the Decent Work Agenda***". The paper is co-authored with Fabio Sartori Piran, Daniel Pacheco Lacerda, Debora Oliveira, and Ana Maria Camanho.

- Section 6 is based on the paper: "***Application of Association Rules to identify patterns of occurrence in work-related accidents: an analysis of the effect on efficiency***". The paper is co-authored with Leandro Gauss.

ACKNOWLEDGMENTS

First and foremost, I would like to thank God, who has granted me countless blessings, knowledge, and opportunity to be able to finally accomplish this thesis.

Second, I am grateful to acknowledge that I have counted on the contribution of many people over the years to develop this work and that I could not have undertaken this without my loved family, especially **my wife Josy**, **my daughters Lívia and Julia,** and **my parents Francisco and Lucy** who generously understood my limited time and attention over this period.

Third, I would like to express my gratitude to my supervisors, **Professor Daniel Pacheco Lacerda** and **Professor Ana S. Camanho** for guiding me throughout this journey with their generous sharing, encouragement, and commitment to seeking excellence.

I also would like to show my deepest appreciation to my friends and professors **Leandro Gauss** and **Fabio Antonio Sartori Piran** for their bright minds, generosity, and incredible support all over this amazing endeavor.

Special thanks to my colleagues and researchers from GMAP for the beautiful collaboration, and for creating the best reference of teamwork I ever had in my life.

Finally, I dedicate this thesis to **my sister Luciana K. Gomes (in loving memory)**, a brilliant professor and researcher in Education that quietly inspired me to someday get a Doctoral degree.

.

RESUMO

Os Sistemas de Gestão em Saúde e Segurança do Trabalho (SGSST) cumprem um papel fundamental para as organizações. Segundo o *International Labour Organization* (ILO, 2020), apesar da maior atenção ao tema por parte dos governos e empresas, a quantidade de acidentes e doenças relacionadas ao trabalho, assim como o respectivo impacto econômico permanecem significativos. Aspectos técnicos, comportamentais e econômicos são afetados pelo desempenho desses sistemas e, por consequência, o resultado geral das organizações. Neste contexto, a avaliação de performance do SGSST representa um mecanismo importante para revisão das medidas implementadas com respeito aos resultados obtidos. Esse campo de estudo, entretanto, tem sido negligenciado na medida em que divergências conceituais são frequentemente observadas nos instrumentos de avaliação propostos, comprometendo a qualidade da análise dos resultados. Além disso, as organizações avaliam o desempenho do SGSST predominantemente através da eficácia (cumprimento dos objetivos previamente definidos), sem levar em consideração os recursos utilizados, e sem avaliar o impacto de cada iniciativa em relação aos resultados obtidos. Portanto, defende-se a tese de que a performance de SGSST deve ser avaliada de forma complementar por meio da análise da eficiência, e que a *Data Envelopment Analysis* (DEA) pode contribuir para esse objetivo. Essa pesquisa segue o modelo de tese baseada em artigos e tem por objetivo geral a avaliação da eficiência de Sistemas de Gestão de Saúde e Segurança do Trabalho (SGSST) por meio do uso da DEA. O estudo contribui para preencher lacunas teóricas acerca da precisão conceitual das medidas de performance utilizadas na literatura de SGSST, propondo uma base conceitual para essa linha de estudo. Além disso, o modelo de avaliação proposto contribui diretamente para uma maior assertividade na tomada de decisão por parte das organizações, ao agregar a avaliação integrada da eficiência técnica e da eficácia do SGSST. As principais limitações do estudo são a ausência de aplicação do modelo em benchmarking externo, e no fato de que os dados utilizados serem provenientes de um caso único.

Palavras-chave: Sistema de Gestão em Saúde e Segurança do Trabalho; SGSST; Análise de Eficiência; Análise Envoltória de Dados.

ABSTRACT

Occupational Health and Safety Management Systems (OHSMS) play a fundamental role within organizations. According to the International Labor Organization (ILO, 2020), despite greater attention to the issue by governments and companies, the quantity of work-related accidents and illnesses, as well as their respective economic impact, remain significant. Technical, behavioral, and economic aspects are affected by the performance of these systems and, consequently, the general result of organizations. In this context, the performance evaluation of the OHSMS represents an important mechanism for reviewing the implemented measures concerning the results obtained. However, this field of study has been neglected as conceptual divergences are frequently observed in the proposed assessment instruments, compromising the quality of the analysis of the results. Furthermore, organizations evaluate the performance of the OHSMS predominantly through its efficacy (i.e. achievement of objectives previously defined), without taking into account the resources used, and without evaluating the impact of each initiative concerning the results obtained. Therefore, we defend the thesis that the performance of OHSMS should be complementarily evaluated through the analysis of efficiency, and that the Data Envelopment Analysis (DEA) can contribute to this objective. This research follows the paper-based thesis (PBT) model and has as its primary objective the evaluation of the efficiency of Occupational Health and Safety Management Systems (OHSMS) using DEA. The study contributes to filling theoretical gaps about the conceptual accuracy of performance measures used in the OHSMS literature, proposing a conceptual basis for this field of study. In addition, the proposed model directly contributes to greater assertiveness in decision-making by organizations, by aggregating the technical efficiency and the efficacy on the evaluation of the OHSMS. The main limitations of the study are the lack of application of the model in external benchmarking and the fact that the data used come from a single case.
Keywords: Occupational Health and Safety Management System; OHSMS; Efficiency analysis; Data Envelopment Analysis.

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AIHA	American Industrial Hygiene Association
BS	British Standard
DDF	Directional Distance Function
DEA	Data Envelopment Analysis
DMU	Decision-Making Unit
DSR	Design Science Research
DWA	Decent Work Agenda
DWMF	Decent Work Measurement Framework
EU-OSHA	European Agency for Safety and Health at work
H&S	Health and Safety
ISO	International Standard Organization
ISRS	International Safety Rating System
OHS	Occupational Health and Safety
OHSMS*	Occupational Health and Management Standard
OHSMS	Occupational Health and Management System
OSHA	Occupational Safety and Health Association
PBT	Paper-based Thesis
PDCA	Plan-Do-Check-Act
RE	Realistic Evaluation
SGSST	Sistema de Gestão de Saúde e Segurança do Trabalho
SIF	Serious Injuries and Fatalities
SME	Small and mid-size enterprise
TOC	Theory of Constraints
VPP	Voluntary Protection Program

TABLE OF CONTENTS

1 INTRODUCTION

Occupational health and safety (OHS) at work are vital components of decent work. The physical conditions and mental demands of the workplace determine to a great extent workers' conditions. (ILOSTAT, 2021).

OHS management has become a global issue, and solutions to enhance its performance have been urgently required in modern work environments (Wang et al., 2020). According to the International Labour Organization (ILO, 2020), more than 2.8 million deaths per year result from occupational accidents or work-related diseases. When considering non-fatal work-related injuries, this number increases to approximately 376.8 million a year. This global perspective based on official statistics of non-fatal occupational injuries and work-related fatalities per 100,000 workers is broken down in Figure 1.

The issue cannot be considered as a particularity of economies in transition and developing economies since the impacts of the number of accidents can be perceived in the majority of countries, including developed economies such as Switzerland, Germany, and Denmark. However, some Latin American economies have a significantly higher rate of occupational fatalities. While European developed countries such as Germany and Denmark have one fatality per 100,000 workers, Brazil and Mexico report seven times more.

Although the growing importance of occupational health and safety is evidenced by international standards and scientific publications, the economic losses incurred from work-related accidents and occupational illnesses are still a challenge. This pragmatic reality, retrieved from poor occupational health and safety management systems (OHSMS), accounts for economic losses estimated at 3.94% of the global Gross Domestic Product (Brocal et al., 2018; ILO, 2020; Wang et al., 2020). This significant impact suggests a potential lack of efficiency in terms of OHS management within the organizations. Also, the concepts used in OHSMS performance measurement should be reviewed by considering the specialized literature (Coelli et al., 2005; Piran et al., 2020), as well as the assessment instruments available to evaluate the OHSMS.

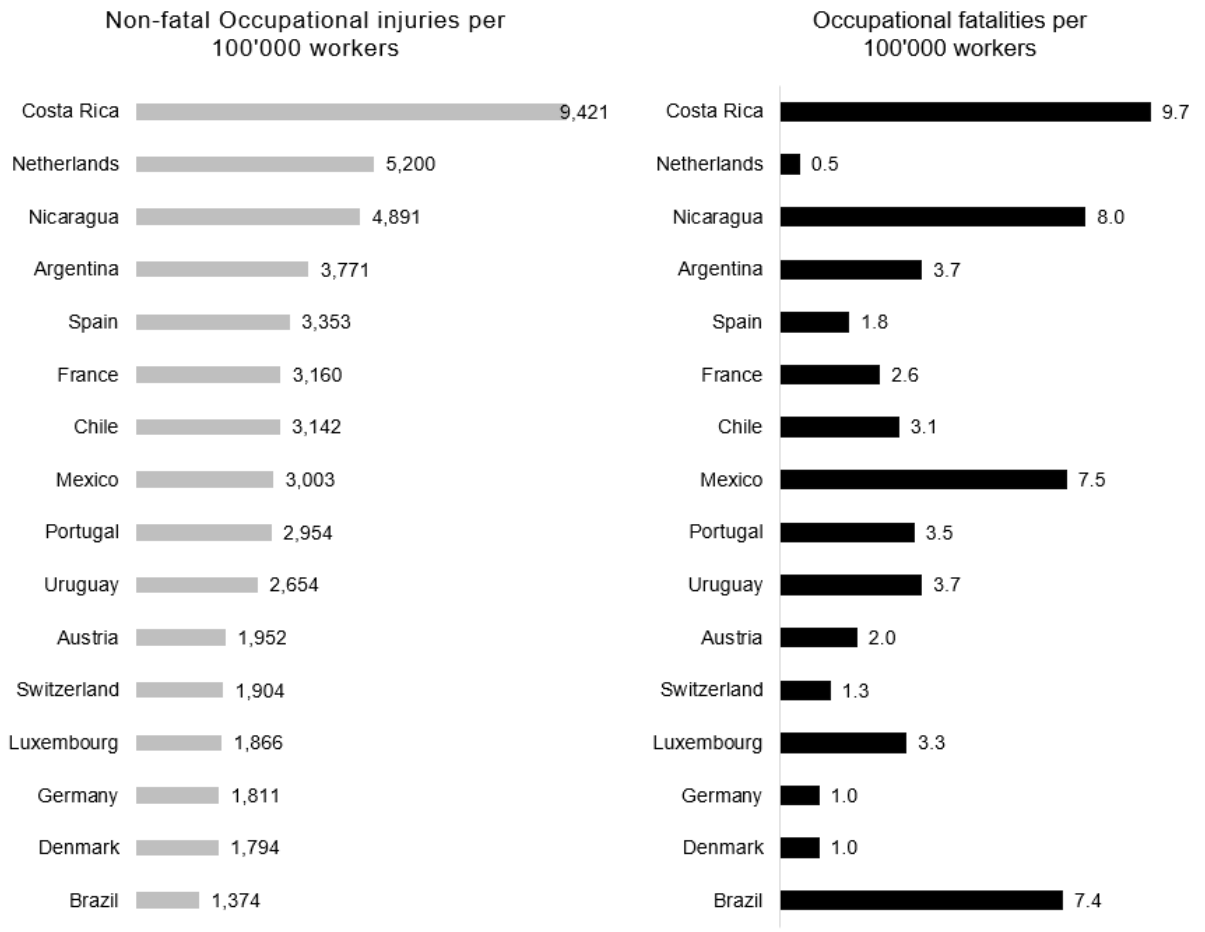


Figure 1 – Occupational Health and Safety indicators

Source: International Labour Organization (202). ILOSAT database. Accessed on March, 31st 2021. Available from https://ilostat.ilo.org/topics/safety-and-health-at-work/

Theoretical roots explain the historical evolution of the OHSMS (Figure 2). Initially, in the so-called Heinrich's root, the accident causation analysis, and safety audit instruments were developed as a primary initiative to prevent work-related accidents and losses. The concept of OHSMS was not yet defined and the proposed audit instruments focused on operational occupational health and safety issues. A second root is based on quality management systems developed in the United States in the early twentieth century, notably due to the tools and techniques of process statistical control (Redinger et al., 2011). These techniques got wide acceptance with the integration of the Japanese improvement philosophies, and their earliest application to OHS resulted from the work of Bird and Germain, 1976 based on causation analysis of accidents reported by the steel industry. This Quality-based root was fundamental to the creation of the ISO quality standards that contributed to the earliest conceptual base of OHSMS, and to establishing means of evaluation.

The evaluation of OHSMS is mainly represented on the called Michigan's root due to the studies conducted in the School of Public Health of the University of Michigan (Dyjack et al., 2003; Dyjack and Levine, 1996a, 1996b, 1995; Redinger et al., 2011, 2002a, 2002b; Redinger and Levine, 1998). Despite the attention given to the development and implementation of OHSMS, there has been relatively little attention given to the measurement of its effectiveness (Robson et al., 2007). The OHSMS assessment instruments represent methods and tools designed to structure the data and provide appropriate analysis. Moreover, those instruments shall be comprehensive for academic researchers and practitioners, and provide more evidence about the effectiveness of interventions aimed at protecting worker's health, employing language and formats also suitable to non-scientific audiences (Bigelow and Robson, 2005a).

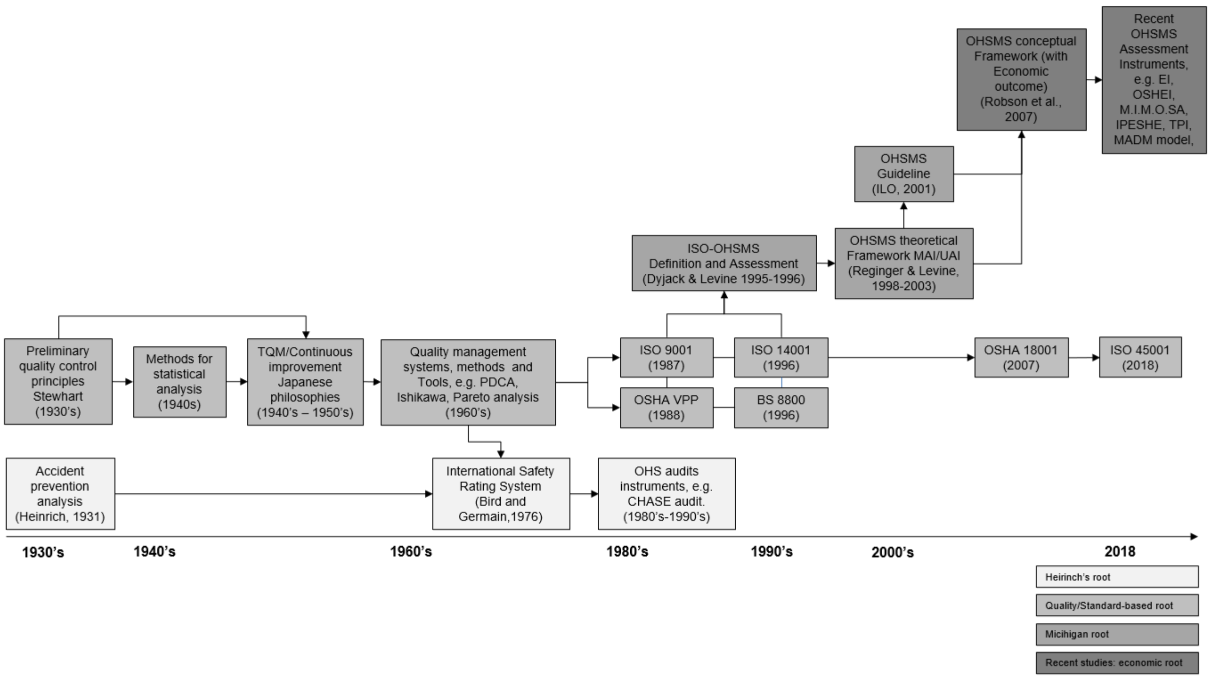


Figure 2 – Theoretical roots of OHSMS

Historically, the studies about OHSMS assessment instruments found in the literature are sparse. Bigelow and Robson (2005) reviewed eleven OHS audit instruments based on measurement concepts, such as interrater reliability, responsiveness, and content validity, among others. The OHS management audit method developed by Diekemper and Spartz (1970) appears to be the earliest in the literature it was structured in five sections, and raters were required to categorize each activity on a four-category scale rating, from poor to excellent. Other instruments, notably before 1995, a time when the concept of OHSMS had not yet been proposed, focused only on specific areas of interest, instead of a broader perspective. However, the assessment instruments initially proposed in 1995 at the University of Michigan showed strong evidence of content validity, in contrast with preceding ones, where the level of detail in the published information did not allow a judgment. A possible reason for that is because the concept of an occupational health and management system only emerged in 1995 with the studies of Professor Steven Levine and his Ph.D. students (Dyjack, 1996; Dyjack and Levine, 1996a; Redinger, 1998; Redinger and Levine, 1998) in partnership with the American Industrial Hygiene Association (AIHA).

The discussion on the need for OHS assessment tools emerged at the University of Michigan when the researchers raised the question of whether the international community should also consider the development of an ISO[1] 9000-compatible occupational safety and health management standard (OHSMS[2]). They intended to evaluate the advantages and disadvantages of developing an ISO standard compatible with the necessary requirements in the area of OHS (Dyjack and Levine, 1995).

The ISO 9000 series was first published in 1987 to encourage companies to implement quality management systems. It has received over the years considerable attention from the public and private sectors due to its potential to increase performance (Terziovski et al., 2003; Terziovski and Samson, 1999). ISO 9000 standards count on a fundamental feature of generalizability, being applied as a tool to evaluate the system performance of any type of organization (Naveh and Marcus, 2005).

In 1994, the ISO Technical Committee for Environment Management (TC-207[3]) was drafting additional standards to account for environmental aspects of corporate operations. These drafts received the 14000 series designation and encompass environment management systems, i.e. ISO 14001 (ISO, 1996), environmental performance evaluation, and auditing, among other related matters. In parallel,

---

[1] ISO is a nongovernamental organization headquartered in Geneva, Switzerland. It was created in 1946 to promote the development of international standards.
[2] The terminology OHSMS was preliminarily defined as Occupational Health and Safety Management Standard (instead of System) due to the initiatives for developing an ISO-compatible standard the encompass necessary requiremetns in the area of OHS. Later on, OHSMS gained a broader concept, being called Occupational Health and Management System.
[3] The ISO Technical Management Board created Technical Committee 207 (TC-207) in 1992 to develop internationally recognized environmental management standards. (Dyjack and Levine, 1996a)

practitioners initiated discussion on the merits of developing an international occupational health and safety standard (OHSMS), to complement the existing ISO 9000/14000 series.

On one side, numerous advantages were mapped with a potential ISO-OHSMS in place, e.g. the integration between prevention-oriented OSH and environmental programs, compatibility with the scope of ISO 9001 2nd version (ISO, 1994), and harmonization with the entire ISO 9000/14000 series (ISO, 1996, 1994). Besides, organizations may benefit from access to shared visions from health and safety processes, that are reflected in the standards. On the other side, disadvantages were also raised. An ISO-OHSMS could be expensive for small companies, and redundant for those already engaged with health and safety programs. A stronger influence could be perceived by dominant countries, and the potential for fraud through unethical registrar conduct also represents a challenge for evaluation.

Based on emerging interest from agencies and practitioners, the OHSM concept came up with the formal resolution from the ISO TC-207 in May 1994, where the ISO Technical Management Board was requested to determine whether there was a need to evaluate the desirability for standardization in the area of OHSM. Also, the British Standard Institute, regarded as a major force for the development of ISO 14000, published its draft guide to health and safety management systems in December 1994, later on, designated as BS 8800 standard (BSI, 1996).

In 1996, (Dyjack and Levine, 1996a) were pioneers in publishing a broad concept of an Occupational Health and Safety Management System (OHSMS) as “an orderly arrangement of interdependent activities and related procedures that drives an organizational’s occupational health performance” (Dyjack and Levine, 1996, p.932). This definition supported the idea that an OHSMS assessment instrument would need to evaluate the system features, and it should be based on the assumption that an OHSMS assessment instrument should resemble those found in the structure and scope of ISO 9001 and emerging ISO 14001. This approach that integrates quality, process safety, and the environment was previously discussed by Berkey et al. (1993) and brings some advantages as it could reduce audit fatigue associated with multiple site assessments and place H&S alongside quality and other business aspects as equal in organizational priorities.

Critical features for the OHSMS assessment instrument were then identified, such as structure and scope, nested statutory requirements, synchronic reliability, predictive validity, auditor bias, continuous improvement metrics, beneficial aspects, implications to organizations of modest resources, and differences between public and private assessment instruments. Those features became an initial reference for creating a valid, reliable OHSMS assessment instrument to support continuous H&S system improvement efforts.

A theoretical framework and an universal assessment instrument (UAI), also called the Michigan assessment instrument (MAI) were proposed in 1998 to measure the effectiveness of a wide range of occupational health and safety management systems (Redinger and Levine, 1998). The MAI was structured based on the review of four previous and consolidated systems that were selected because they provided the most comprehensive management system approaches and contained the essential elements of all of the models reviewed (OSHA VPP, BS 8800:1996, AIHA OHSMS, and ISO 14001:1996).

The proposed framework (Figure 3) was a pioneer in defining boundaries, constructs, and elements that comprise an OHSMS. Besides, the framework was primarily structured by considering elements of initiation, e.g. management commitment, OHS processes, and outputs, such as the number of work-related injuries or accident rate, as well as encompassing the evaluation process to support management review and continuous improvement. Also, the structure of the Michigan Assessment Instrument came to light based on this conceptual approach and resulted in 5 categories, 27 sections, 118 principles, and 486 management criteria.

One of those categories is the management review. It encompasses OHSMS principles and measurement criteria to evaluate the overall performance. Different from a single system evaluation that focuses on a particular OHS component, the management review assesses whether the initiatives and OHS processes are in the correct direction to reach the intended objectives and outcomes. Also, it provides the link between the OHSMS, the organization, and the environment external to the organization and differs from system evaluation (Redinger and Levine, 1998, p.580). Thus, an overall performance evaluation approach plays a fundamental role in a successful system and is a key attribute of strong management commitment to OHS.

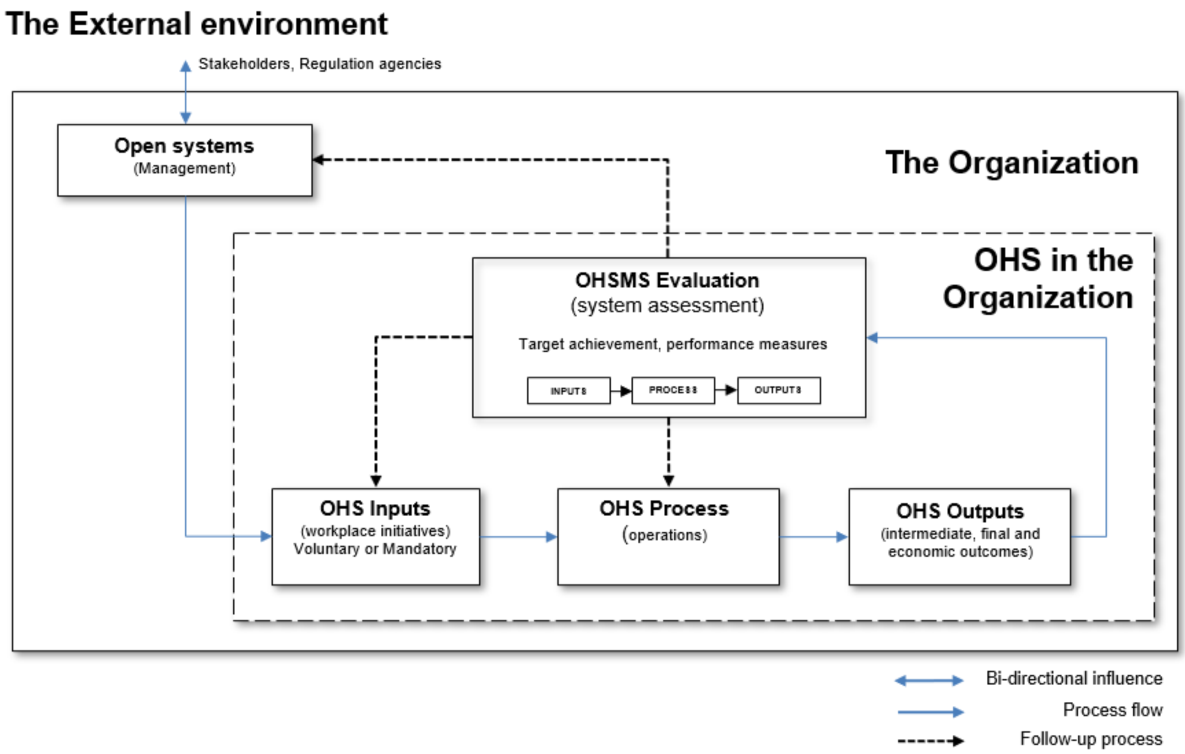


Figure 3 – Theoretical OHSMS framework. Adapted from Redinger and Levine (1998)

The universal OHSMS assessment tool developed in Michigan was pilot-tested to evaluate its effectiveness in three sites with a particular focus on the Initiation category (Redinger et al., 2002b, 2002a). Results of the scale scores were found to be consistent with a qualitative assessment of the three sites, giving preliminary construct validity to the audit instrument (Bigelow and Robson, 2005a). Dyjack et al., 2003 continued the work for validating the universal assessment instrument and examined four auditable sections from different categories (employee participation, training, controls, and communications). A modified test-retest method using a different scoring system was used, and the findings suggested opportunities for improving the reliability of the instrument and the existing audit process.

Despite that theoretical root addressed by the researchers from the University of Michigan, and their initial tentative to develop and validate an universal instrument to evaluate any OHSMS, not much progress on theory-based models has been achieved in the literature. Also, economic losses incurred from work-related accidents and occupational illnesses are still a challenge at the organizational level.

Another conceptual framework where economic outcomes are outlined was proposed in the work of Robson et al., 2007. Three different types of outcomes, intermediate (e.g. safety climate), final (e.g. injury rates), and economic (e.g. insurance premiums) were presented as a broader perspective for studies in OHSMS assessment, evidenced in the research of Ramli et al. (2011) and Saracino et al. (2015). Yet, this so-called Economic root based on models and empirical studies recently published in the literature has attested that the economic aspects in the inputs and outputs are still ignored, being prevalent within the organizations the assessment of safety management systems based on the standards (e.g. ISO), and tools such as the Plan-Do-Check-Act (PDCA) cycle (Deming, 1986).

In this context, a new standard ISO 45001 (ISO, 2018) was recently released to focus on the OHSMS requirements and to help organizations improve employee safety, reduce workplace risks, and create better working conditions (Brocal et al., 2018). ISO 45001 is getting growing adherence within the organizations, and it suggests that OHSMS should be based on the PDCA reasoning. In the same way as previous standards, it does not deepen the concept of performance evaluation (Check), and it is limited to superficial guidance on performance assessment implementation.

Performance measures are well-defined in operations management. Among the key concepts used in production systems' assessment, productivity, efficacy, and efficiency appear as the most relevant dimensions.

Productivity is the output-to-input ratio, corresponding to a performance indicator that organizations wish to maximize. Efficacy relates to a direct measure associated with the achievement of a defined objective, without taking into account the resources used. Efficiency, in turn, is a relative measure of performance. A production system is considered 100% efficient if performance observed in others does not show that is possible to improve some of the inputs or outputs without worsening the other input and outputs (William W. Cooper et al., 2011a). In simple terms, efficiency is doing things right and efficacy is doing the right thing (Druker, 2006).

Those concepts of performance measures were synthesized by (Piran et al., 2020).

A well-recognized technique for efficiency analysis is the Data Envelopment Analysis (DEA) developed by Charnes et al. (1978). DEA is a non-parametric technique used to assess the relative efficiency of Decision Making Units (DMUs) (Piran, 2021a).

DEA has been widely used to evaluate the efficiency of productive systems in different sectors of the economy. However, despite DEA being considered one of the main techniques used to measure the efficiency of the systems, the majority of studies focus on the evaluation of technical efficiency, corresponding to considering only the quantities of resources used (Piran, 2021a). Studies that consider economic efficiency (input cost minimization for the outputs produced) to evaluate health and safety management systems are scarce, notably in using DEA.

In the analysis with the use of DEA, the system is called DMU and the relative efficiency of DMUs is assessed through a benchmarking procedure. Using the DEA it is possible to calculate the efficiency of an OHSMS by considering it as the entity under assessment. Technical efficiency takes into account physical quantities used in the production process, while economic efficiency (cost, revenue, and profit) also takes into account the prices of the factors used in that process (Färe et al., 1994, 1985). Furthermore, the OHSMS can be comparable to themselves in previous periods (corresponding to an internal benchmarking exercise) or with other organizations at a given moment in time (corresponding to an external benchmarking exercise).

Studies using DEA in the field of occupational health and safety are rarely found in the literature when compared to other areas. Nahangi et al. (2018), for instance, proposed a DEA-based method for identifying the efficiency of construction sites. Nissi and Rapposelli (2012) examined the safety performance of fifteen European countries, in three economic sectors-manufacturing, construction, and distribution. Other research addressing safety performance in the construction and industrial sectors using DEA can be also found in the works of Beriha et al. (2011); El-Mashaleh et al. (2010); Kang et al. (2020); Said et al. (2013), and Yeh (2017).

Worth mentioning that other applications associating DEA with safety are available, although the majority of these studies are out of the scope of this thesis, e.g. traffic accident prevention, and road safety performance (Amini et al., 2019; Antić et al., 2020; Chang et al., 2020; Ganji et al., 2020; Krmac and Djordjević, 2018; Omrani et al., 2020; Seyedalizadeh Ganji et al., 2019; Seyedalizadeh Ganji and Rassafi, 2019; Shah et al., 2018; Shen et al., 2010, 2020).

Following that reasoning, existing literature indicates a lack of studies concerning the use of DEA for OHSMS assessment from the perspective of efficiency analysis. Only a few studies tackle a piece of such an economic approach, although not targeting to address a robust OHSMS efficiency assessment (Babajani et al., 2019; Shirali et al., 2018). In this context, DEA can support this process by seeking an original approach to close this gap. Also, by implementing a cost-efficiency model integrated with technical efficiency measurement, practitioners are given relevant data analysis towards achieving better safety outcomes, such as identifying internal benchmarks, critical inputs, and robust evaluation to drive continuous improvements.

The discussion carried out exposes, in general terms, the convergences between the analysis of operational productivity and OHSMS assessment. Thus, it is argued that the use of the DEA-based model is feasible as an assessment instrument to evaluate the efficiency of the occupational health and safety management system as proposed in this work. Also, this study expands the literature by discussing and proposing conceptual definitions of the performance measures used in OHSMS, and by creating an artifact where it is possible to aggregate the OHSMS technical and allocative efficiency assessment. Finally, considering the context presented, and seeking to contribute to the analysis of OHSMS performance in organizations, this thesis is adherent to the topic of OHSMS evaluation. The next section is presented the object and research question.

1.1 Object and Research Problem

The problem this research poses is how to evaluate the efficiency of Occupational Health and Safety Management Systems.

Over the years, assessment instruments have been proposed to evaluate OHSMS as a mechanism for checking the outcomes and to promote continuous improvement (Dyjack and Levine, 1996b; Ghahramani, 2017; Holvad et al., 2004; ISO, 2018; Lehtola et al., 2008; Mohammadfam et al., 2017; Redinger and Levine, 1998; Robson et al., 2007; Skład, 2019). Yet, although greater attention has been given to OHSMS since the end of the 90s, the number of work-related accidents and occupational illnesses, as well as their respective economic impact, remain significant (Brocal et al., 2018; ILO, 2020; Wang et al., 2020).

On the academics side, a possible reason for such a level of ineffectiveness is the stagnation of safety science due to the lack of reality-based safety studies, and an excess of untested models, reductionist categories, and proxy measurements, rather than direct observation and sophisticated analysis of real people doing real work in real organizations (Rae et al., 2020).

On the practitioners' side, several aspects might contribute to this undesirable reality. The first one (1) is not considering the granularity issue when defining initiatives or assessment methods, i.e., ignoring the scale or level of detail present in a particular phenomenon observed, for instance, in a job site, an SME organization, a multinational company comprised by business units, or even in a country. Granularity, in general terms, contributes to defining an adequate tool for solving a particular problem. Fall protection systems, for example, can be designed for a specific service, or a class of problems. In both cases, despite dealing with the same problem, they have different levels of granularity and potentially different solutions.

The second aspect (2) is associated with financials. Due to limited resources, small and mid-size organizations (SMEs) do not undertake significant efforts in health and safety management, limiting themselves to mandatory requirements. Thus, in countries where small businesses prevail, the efficiency of H&S initiatives is potentially ignored, and results do not consistently improve. This is consistent with statistics retrieved from (ILOSTAT, 2021), where countries such as Brazil still struggle to decrease the social and economic impacts of poor H&S management.

The following issue (3) is a misunderstanding that investing in health and safety incurs reduced productivity or increased costs. Although the cost associated with accident prevention is far less than the cost of accidents themselves, organizations still seem to doubt that the investment is worth it under claims of cost and reduced productivity. Such lack of comprehension might be associated with inadequate methods to assess OHSMS, without taking into consideration the proper inputs and outputs related to the system.

Next aspect (4) is the lack of cohesion between global and local H&S policies to drive H&S improvements. Large companies often define global policies to be followed by business units. In this context, initiatives of general scope may not deliver the expected results as they were not structured to solve specific problems in different contexts. For example, technical labor qualification is not an issue in the American elevator industry. This is because the Union regulates and assesses the training of all workers in this sector. On the other hand, in Brazil, companies need to train their personnel to ensure quality and safety. Therefore, a global policy that does not focus on technical training may negatively impact Brazil and have a low effect in the United States.

The fifth aspect (5) is related to the OHSMS assessment. In the circumstances where the OHSMS is established, organizations predominantly assess OHSMS performance just by checking its efficacy, such as the percentage of target achievement (e.g. 10% reduction in the number of lost time accidents compared to the previous year). In these cases, two circumstances can arise. On the one hand, when the result is positive, organizations are satisfied and define new annual goals without even understanding the existing relationships between the resources used and the results obtained. On the other hand, when the results are frustrating, it rises a gap in understanding the reasons why the results were below the objective, leading managers to make important decisions based on fragile data and unsupported arguments.

The next aspect (6) is not to consider the assumption that initiatives in the H&S field might present delays to deliver results. Delays between either an action or initiative and a perceived result can often occur in the occupational health and safety field. This is because part of the initiatives involves, e.g. behavioral changing, and initiatives aimed at strengthening the safety culture. Those kinds of initiatives do not generate immediate results since they require time to mature within the organizational culture. The risk of not acknowledging the existence of delays is prematurely to consider an action as ineffective because no short-term results were observed. In this case, resources are wasted and new initiatives are established without a proper understanding of cause-effect.

Finally (7), isolating initiatives in H&S to observe their results is a difficult task and requires a proper methodology and highly skilled personnel to carry out this type of analysis. Therefore, even for companies that assess the effectiveness of their OHSMS, understanding causality in-depth is a significant challenge for both safety experts and managers.

All these aspects are consistent with the fact that H&S management at the organizational level is usually responsive to regulatory requirements rather than evidence of what works. In this context, evaluating the efficiency of OHSMS grounded in rigorous observations of existing practice is fundamental for the creation of a virtuous cycle toward the direction of a more promising scenario.

OHSMS can be represented as a typical system comprised of initiatives (inputs), processes, and outcomes (outputs) within a context. However, it is encompassed into a complex structure as depicted in the framework shown in Figure 4. This is a side contribution of this thesis since it offers a wide understanding of OHSMS, exploring the relationships between some constructs, organizational elements, and OHSMS components.

Health and Safety are about the well-being of people. It is ultimately about stopping the likes of employees, visitors, and customers from being subject to workplace incidents so that they can enjoy freedom from illness and injury (Duroe, 2021). Some constructs are fundamental for giving meaning to that: ethics and social responsibility, corporate social responsibility, and organizational culture.

According to some management literature, ethics and social responsibility have distinct concepts and views under different perspectives. This thesis, which has no intention of going deeper into the topic, adopted the most common view of the relationship between ethics and social responsibility: *"(…) social responsibility has various dimensions, one of which is ethics"* (Fischer, 2004, p. 4). This view is grounded in Carrol's pyramid (Carroll, 1999) and it is consistent with the approach taken by many ethics and management texts and considers that there are four dimensions of corporate social responsibility: economic, legal, ethical, and philanthropic. This is the starting point for understanding OHSMS as depicted in Figure 4: H&S is a social responsibility topic, and therefore, should be an integral part of businesses of all shapes and sizes.

Over recent years the importance of corporate social responsibility (CSR) has become more apparent (Duroe, 2021). The term CSR gained notoriety in the 1980s to describe how businesses could contribute to the achievement of a sustainable society (Oliveira, 2008). However, CSR is a dynamic concept, for which many definitions have been proposed since the beginning of the last century (Oliveira et al., 2019). It was initially referred to corporate obligations to pursue policies and lines of actions with value to society, with focus on pursuing honest profit (Bowen, 1953). For this thesis, it is adopted the widely accepted definition provided by the World Business Council for Sustainable Development (WBCSD), which describes Corporate Social Responsibility as *"(…) the continuing commitment by business to contribute to economic development while improving the quality of life of the workforce and their families as well as of the community and society at large"* (WBCSD, 2011, p.3). Besides the variety of definitions, CSR is often interpreted as the responsibility of businesses to successfully integrate economic, environmental, and social issues into organizational practices (Belu and Manescu, 2013).

Notably, there is an expectation from society that businesses are held accountable for their ethical practices and that companies consider safety as a top priority. In that reasoning, this is expected to be reflected in the core values and behaviors widely perceived within an organization, i.e., in its culture.

Culture is an ambiguous term. On one side, in a formal approach, it is defined as the shared values, beliefs, and assumptions that govern behavior. On the other side, informally, culture is expressed as the way organization's employees do things around there" (Krause and Bell, 2015). Also, according to (Schein, 1990; Schein and Bennis, 1965) culture is represented by a set of observed behaviors, and it sustains organizational performance (Krause and Bell, 2015).

In the business environment, both formal and informal approaches previously mentioned are consistent with other definitions found in the specialized literature. Also, the market segment influences the organizational culture, e.g. the level of safety requirements in the aviation industry is not comparable with those patterns in construction. This influence impacts qualification, process execution, operational discipline, and other elements of how an organization "does things". Therefore, as a concept, organizational culture can be defined as *"(…) (a) a pattern of basic assumptions, (b) invented, discovered, or developed by a given group, (c) as it learns to cope with its problems of external adaptation and internal integration, (d) that has worked well enough to be considered valid and, therefore (e) is to be taught to new members as the (f) correct way to perceive, think, and feel in relation to those problems"* (Schein, 1990, p.111). This concept outlines the degree of internal consistency of an organization and associates it with the strength of its culture. Furthermore, organizational culture plays a fundamental role in the OHSMS' effectiveness. This is because health and safety need to be intrinsically part of the core values, beliefs, and behaviors spoken and practiced by top management and by the operational workforce. Besides, if H&S is present as a core value in the organizational culture it's also outlined in the organizational strategy.

The framework shown in Figure 4 outlines that organizational strategy is defined by management. Thus, the business purpose, mid/long-term goals, key performance indicators, and the factors of competitive advantage are critical elements that comprise the organizational strategy defined by managers. In general terms, the definition of goals and KPIs either govern or (at least) influence the behavior of managers and the operational workforce. For instance, if an industry measures the occurrence of serious accidents and near misses with the same level of importance, it is possible to expect not much effort in understanding the root causes of each category, and thus, the likelihood of re-occurrence of serious accidents remains high. Acknowledging the argument that organizations with safety-oriented culture state their purpose of business growth and profitability sustained in decent work practices, the values, and perceived behaviors should be consistent with the well-being at work.

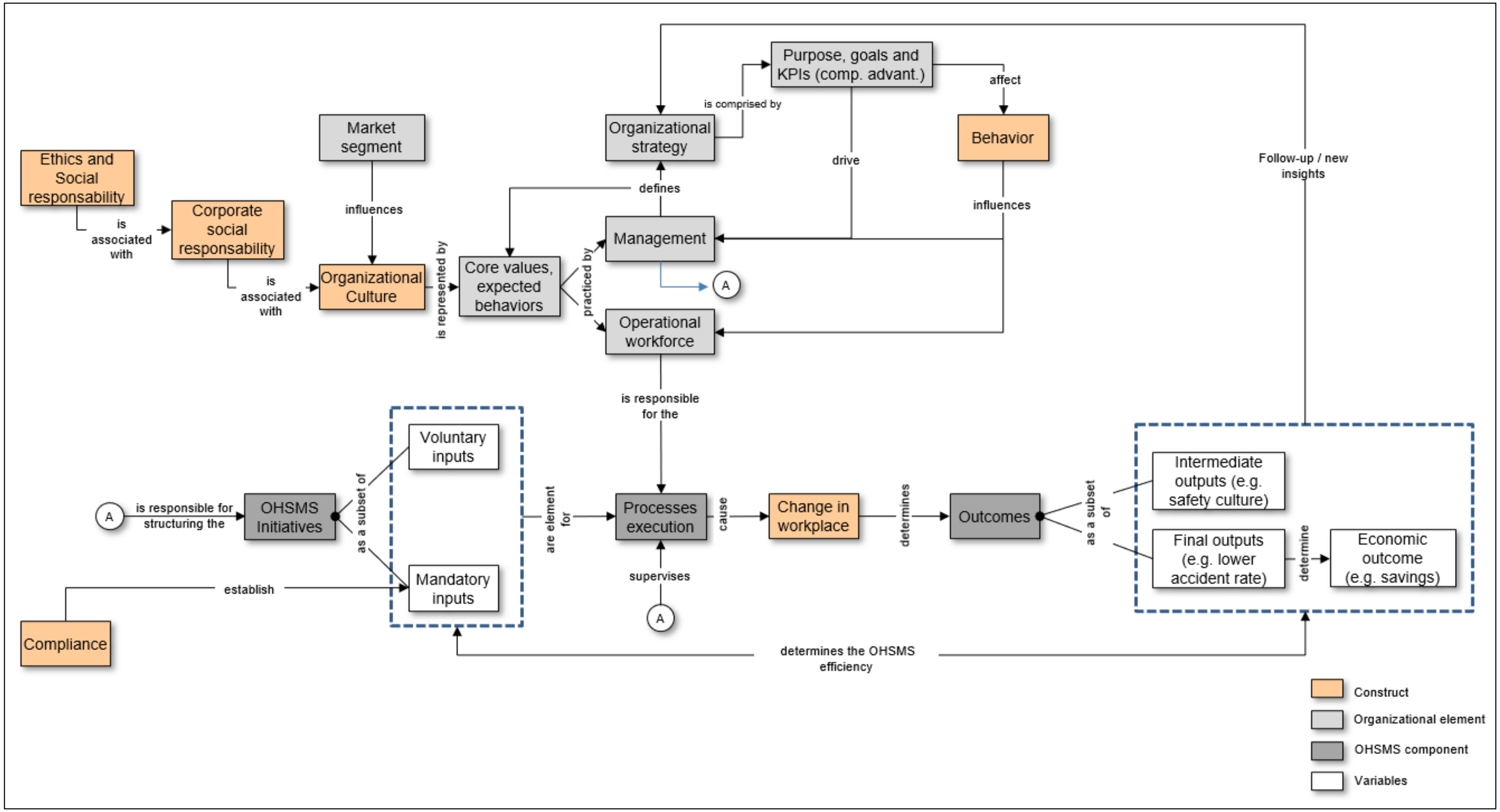


Figure 4 – Relationships between constructs, organizational elements, and OHSMS components

In this context, an occupational health and safety management system is supposed to be based on voluntary inputs, beyond mandatory requirements. As previously referred, OHSMS is a system comprised of initiatives (inputs), processes, and outcomes (outputs) within a context. Examples of inputs of this system are the number of safety experts, training hours, personal protective equipment (PPE), top-management engagement, standardized processes, labor costs, insurance costs, fees, and fines due to non-compliant procedures. Outputs are the results of the impact of those initiatives within the context, e.g. number of lost-time accidents, absenteeism, and cost/savings. The relation between those inputs and outputs is the baseline of the efficiency assessment.

Outputs also play the role of following up and offering insights into the organizational strategy promoting a virtuous cycle of continuous improvement in H&S management. Yet, a common approach to evaluate OHSMS is to focus on the final outcomes of the system, e.g. number of work-related accidents (Robson et al., 2007). This type of evaluation only addresses the level of efficacy resulting from the initiatives implemented, compared to the targets defined at the management level. For instance, accident rate-oriented organizations are of utmost interested in reaching the target. However, the usage of resources employed to reach the results is not properly evaluated. This approach does not offer any insight concerning the potential optimization of the resources utilized. Also, it does not come up with any direction for managers that need to address unsuccessful results. Finally, whether the comparison between different organizations or business units within the same organization, the lack of information about the relationship between outputs and inputs prevents adequate analysis.

This input-output relation is well-addressed by measuring productivity and efficiency. Productivity means an absolute measure that considers the relationship between the results generated and the resources used: the ratio of outputs to inputs. Efficiency, in turn, is a relative measure that compares realized productivity with maximum productivity (Cummins and Weiss, 2013; Førsund, 2018; Kerstens et al., 2019). It can be expressed as technical and allocative efficiency. On the one side, technical efficiency is related to the capacity of a process to produce a certain quantity of goods and services, consuming the lowest quantity of inputs. On the other side, allocative efficiency reflects the ability to minimize costs, using inputs in optimal proportions (Ferreira and Gomes, 2009). Economic efficiency, therefore, is a broader concept as it involves the optimized choice of the level and mix of inputs and outputs, taking into account costs (Piran et al., 2020).

Economic analysis provides a broad view of efficiency and may suggest actions initially considered counterintuitive for the analysis of technical efficiency. With the application of the DEA to analyze the economic efficiency of the Occupational Health and Safety Management System, the research design is shown in Figure 5. The model proposes the use of DEA to analyze technical, allocative, and economic efficiency for the OHSMS, and it assumes that an organization's OHSMS is composed of a set of subsystems existing in its business units. Thus, it would be possible to identify one or more efficient DMUs that would be considered an internal benchmark to promote continuous improvement based on the relationships between inputs and outputs. Similar to (Piran, 2021a), the modeled system should be analyzed preferably longitudinally due to its nature of cause-effect or with panel data (when possible), considering an internal benchmark as the central premise.

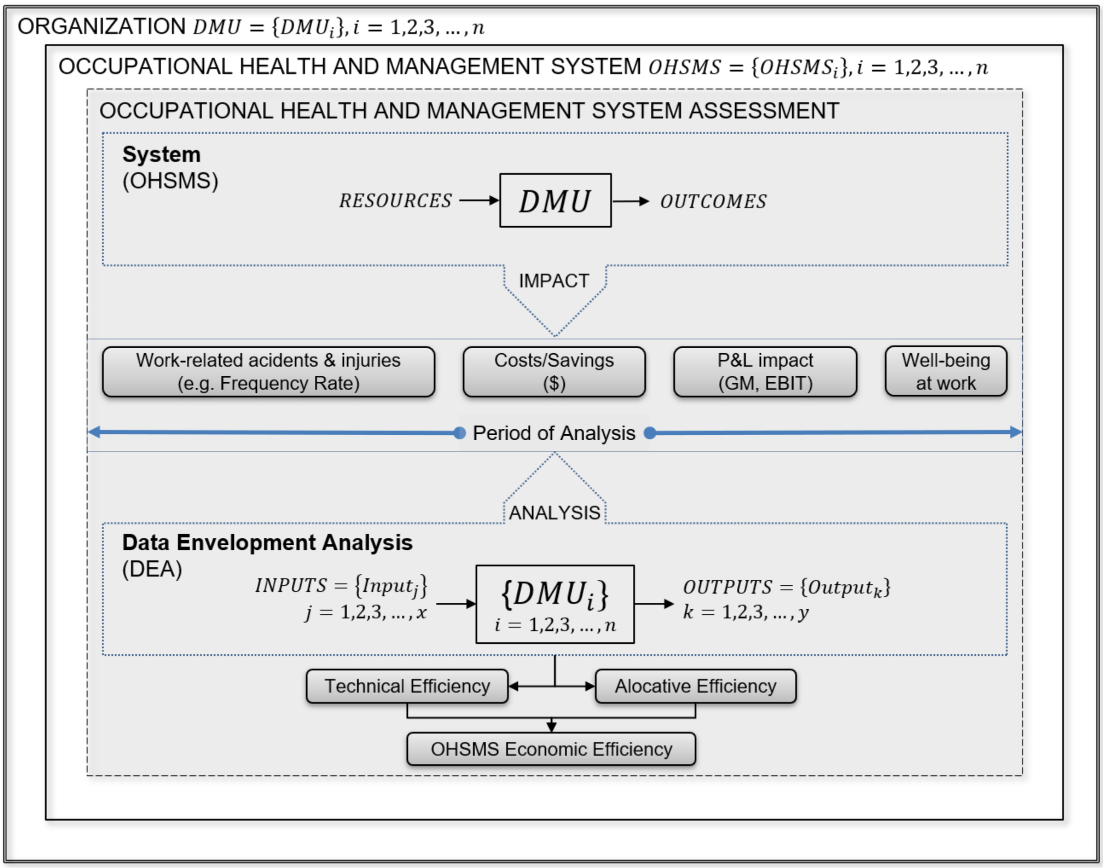


Figure 5 – Research Design

The research question was formulated based on the design proposition following the CIMO-logic (Denyer et al., 2008) as shown in Table 1.

Table 1 – Research problem structure based on CIMO-Logic

Component	Scope
C – Context	OHSMS evaluation at the organizational level
I – Intervention	Efficiency analysis
M – Mechanism	Data Envelopment Analysis (DEA)
O – Outcomes	DEA-based Model, benchmark identification, insights for non-efficient DMUs, and continuous improvement

This logic is primarily based on the technological rule (Bunge, 1967), stated as follows: “to achieve outcome *O* in context *C*, use intervention *I*”. However, a missing link with regard to causality was raised by the realistic evaluation (RE) proposed by (Pawson and Tilley, 1997). RE is a form of theory-driven evaluation that aims to advance an understanding of why these complex interventions work, how, for whom, in what context, and to what extent (Greenhalgh et al., 2015). It raises the issue of causality, i.e. by asking which generative mechanism(s) the intervention produces the outcome in the given context. Denyer et al., 2008 included the mechanism as a component of the design proposition in their CIMO-logic, defined as a general template wherein a problematic Context, an Intervention type is used to invoke generative Mechanism(s), to deliver Outcome(s).

Considering the aspects discussed, the research question was defined as: *how to evaluate the efficiency of OHSMS?* Moreover, the research’s boundaries were well-defined to reach the primary and specific objectives of this research, as discussed in the next section.

1.2 Objectives

This section introduces the primary and specific objectives of this research.

1.2.1 Primary objective

The primary objective of this research is to evaluate the efficiency of Occupational Health and Safety Management Systems (OHSMS) using Data Envelopment Analysis (DEA).

1.2.2 Specific objectives

The specific objectives of this research include:

1) Explore the causalities that govern safety at work based on the general method of theory building in applied disciplines.
2) Identify the constructs that govern safety at work and discuss the complexity of this phenomenon based on the Theory of Constraints.
3) Critically review the existing methods and assessment instruments for the evaluation of OHSMS, including the identification of critical elements and the conceptual foundation used in the examined instruments.
4) Propose a DEA-based model to evaluate the efficiency of OHSMS.
5) Conduct a real-world case-based application to critically analyze the OHSMS efficiency in a relevant context, e.g. Agenda 2030.
6) Compare the results of efficacy and efficiency in OHSMS assessments. Also, evaluate a combined measurement of efficacy and efficiency into the OHSMS effectiveness.
7) Combine the analysis of efficiency with data mining techniques to open new opportunities for future research in safety data science.

1.3 Justification

A healthy and safe work environment not only is desirable from the worker's perspective but also contributes to higher labor productivity and promotes economic growth (Heuvel et al., 2017). According to the European Agency for Safety and Health at Work (EU-OSHA), an effective OSHMS has the potential to increase the competitiveness and productivity of organizations by reducing costs resulting from work-related accidents and illnesses, and by enhancing motivation. Moreover, a decrease in accidents and illnesses at work relieves pressure on public and private social protection, insurance, and pension systems.

Thus, the impacts from the lack of OHS-oriented measures have consequences at individual, organizational and social levels (Heuvel et al., 2017). For individuals, the impacts are observed, e.g. in personal financials, dwindling career prospects, and social isolation. From the employers' perspective, the costs are high due to the loss of productivity, increase in insurance cost, and because of the risk of early retirement and long-term benefit dependency. Socially, this ultimately results in the burden of €2,680 billion (Elsler et al., 2017; EU-OSHA, 2021), equivalent to 4% of the global gross domestic product (GDP), and increased pressure on the social security systems.

Based on these arguments, the magnitude of insufficient H&S practices in workplaces becomes clear. In addition, from the academic perspective, safety science is stagnating and a key symptom of this problem is the lack of high-quality intervention research (Rae et al., 2020). At least three facts sustain this argument: (1) organizations are currently spending significant money on safety, and it is unknown whether and where this expenditure is effective or needed. (2) not much progress has been verified in decreasing the number of work-related accidents and serious injuries and fatalities (Cooper, 2019; ILO, 2020). (3) very few empirical studies of intervention evaluations based on strong methodology are found in the literature (Lehtola et al., 2008; Robson et al., 2007; Vilela et al., 2017).

Traditionally, the effectiveness of most safety initiatives is monitored by measuring, for instance, the number or rate of accident and injury incidents (Lingard et al., 2013). Based on specialized literature about productivity and efficiency, this type of approach should be distinguished from efficiency since the relationships between resources and outcomes are not analyzed.

The problem resulting from this type of “after event” performance indicator is that it does not support safety practitioners and managers in addressing the issues of “what to change?” and “to what to change?” (Goldratt, 1990). In short terms, the effect of the use of resources in the outcomes is out of scope and there is not much useful information to explain how to improve the occupational health and safety management system. For example, organizations that fail in advancing safety management to get better outcomes usually struggle to identify what are the critical factors, what business units are inefficient, and what to do to approximate those inefficient units to the internal benchmarks.

This thesis supports the assumption that OHSMS assessment based on efficacy, e.g. accident rate, severity rate, or absenteeism does not offer the necessary view about critical factors in H&S performance, thus, does not drive necessary improvements. This is because efficacy is adequate for tracking targeted KPIs but not sufficient for managing health and safety systems since it does not consider the resources used. Moreover, efficacy does not allow any association between inputs and outputs, for example, to explain what initiatives were relevant to achieve an outcome. Furthermore, the comprehension of causality is out of the scope of management systems predominantly based on efficacy. In the H&S field, despite its recognized complexity, putting light on what really works is critical to driving strategy reviews and causing the necessary changes. Therefore, it becomes clear that such efficacy measurement of any safety management system is not sufficient to support a consistent improvement of H&S within the organizations.

In general terms, to solve this practical issue, this thesis argues that it is possible to evaluate the performance of OHSMS by following a similar approach to productivity in operations. Also, seeking an original perspective, this study suggests that the performance of OHSMS should be evaluated through the analysis of efficiency and that frontier techniques like Data Envelopment Analysis (DEA) can contribute to this objective.

The assessment of H&S management systems from the perspective of efficiency represents a fundamental contributive factor to the organizational strategy, as previously mentioned and depicted in Figure 4. This is because it integrates different classes of initiatives and outcomes, rising the comprehension of the effects of those initiatives on technical and economic outputs. Particularly for continuous improvement in H&S systems, in-depth comprehension of the relationships between inputs and outputs is critical, since it drives appropriate actions to boost results. Thus, the analysis of OHSMS efficiency supports safety experts and managers toward a data-driven decision-making process, avoiding decisions that are taken based on superficial data and perception.

Based on this, the work seeks to contribute to the analysis of performance in OHSMS at the organizational level in three ways. First, it seeks to establish a theoretical discussion that harmonizes the concepts of performance measurement in safety science, e.g. efficiency, presenting an aggregative review of previous studies, and positioning each piece of research in an appropriate disciplinary context (Rae et al., 2020). This contribution aims to propose a new baseline related to OHSMS assessment concepts, highlighting the most diverse terms used in the literature in disagreement with experts in the field of productivity and efficiency. Also, it supports the understanding of possible scenarios found in the organizations, such as an inefficient OHSMS with a low accident rate or a non-zero accident case unit being the benchmark for the entire organization.

Second, it seeks to contribute to the area of research through an analysis of the economic efficiency of OHSMS with the use of Data Envelopment Analysis. A DEA-based model is designed and applied in an empirical case, where a discussion is carried out based on the existence of convergences between the methodology of analysis of operational productivity and OHSMS assessment.

Data Envelopment Analysis (Charnes et al., 1978; Farrel, 1957) has been used in several applications in operational management (Camanho and Dyson, 2008; Piran et al., 2020). However, in the field of health and safety management, only a few DEA applications can be found in the literature, such as the works of Kang et al., 2020 and Nahangi et al., 2019. These studies were focused to evaluate the performance of the construction sites in China and Canada, respectively. Variables in both studies predominantly consider technical and physical aspects, demographic data, incident reports, and safety climate factors (e.g. management commitment). Following a different perspective, this thesis comes up with a theoretical contribution to using DEA to analyze the efficiency of OHSMS. The study draws the design of an application model based on DEA, and it evaluates the advantages and disadvantages of the technique in a contemporaneous context, e.g. Agenda 2030 (Nations, 2022).

This reality-based application describes how the efficiency analysis of OHSMS can be carried out within the organization to support managers and safety professionals to achieve better results, mitigating significant losses due to inefficiencies in the H&S management system. Also, it is useful for the top management to identify internal benchmarks to drive improvement actions.

Finally, this work aims to analyze the combined efficacy and efficiency measurements to assess OHSMS and the impact on critical factors for competitive advantage. It is argued that the OHSMS assessment from the efficiency perspective is a fundamental driver towards the organizational strategy in promoting continuous improvement.

Existing literature indicates a lack of real-world studies concerning the use of DEA for OHSMS assessment from the perspective of efficiency analysis. Therefore, this thesis aims to propose an original approach to close this gap by proposing a DEA-based model to measure the economic efficiency of OHSMS within organizations, providing relevant data analysis toward better safety outcomes.

The next section draws the boundaries and research delimitation.

1.4 Research delimitation

This research is delimited by its focus on the evaluation of efficiency in occupational health and safety management system within organizations, and therefore, it presents some limitations. First, published studies that are essentially related to healthcare are not part of the corpus of analysis of this thesis. Only studies that take into account occupational health associated with safety aspects were considered in the literature review. This is because healthcare represents a vast field of study and, at times, it has no direct relationship with OHSMS. Second, when analyzing the relationships between the inputs and outputs of the system, the reasons that lead an organization to define its initiatives and its desired results will not be addressed.

The object of the research is the evaluation of the OHSMS. In that reasoning, the study focuses on identifying the efficiency frontier, identifying internal benchmarks, references to inefficient DMUs, and driving direction for improvements. External benchmarks are out of this research's scope. Third, the data collected and analyzed for proposing the DEA-based model comes from a single organization, structured in business units (DMUs). No comparative tests were carried out between DMUs from different organizations. Also, despite its aim, the model has not been applied to the analysis of broader occupational health and safety management systems, such as federal states or nations.

Furthermore, this research does not intend to advance in mathematical modeling in DEA. On the other hand, it seeks to expand the application of theoretical models already consolidated in operations management to a new field of research, which is the OHSMS, seeking to simplify the evaluation of a complex system through the use of a well-recognized technique.

1.5 Research structure

This research consists of a paper-based thesis (PBT) structured in seven sections. The articles composing this thesis were considered in a logical, cohesive, and coherent connection with the research problem to comprise unique research as proposed by (Kubota et al., 2021).

The framework presented in Figure 6 organizes the reasoning behind the research structure of this thesis: starting from a level "A" of comprehension about the research object, we aimed to achieve a level "B" in which condition the comprehension of the phenomenon is, at least, wider. Moreover, this thesis follows a journey A→B grounded in a robust methodological approach, and remarks are communicated through its cohesive papers.

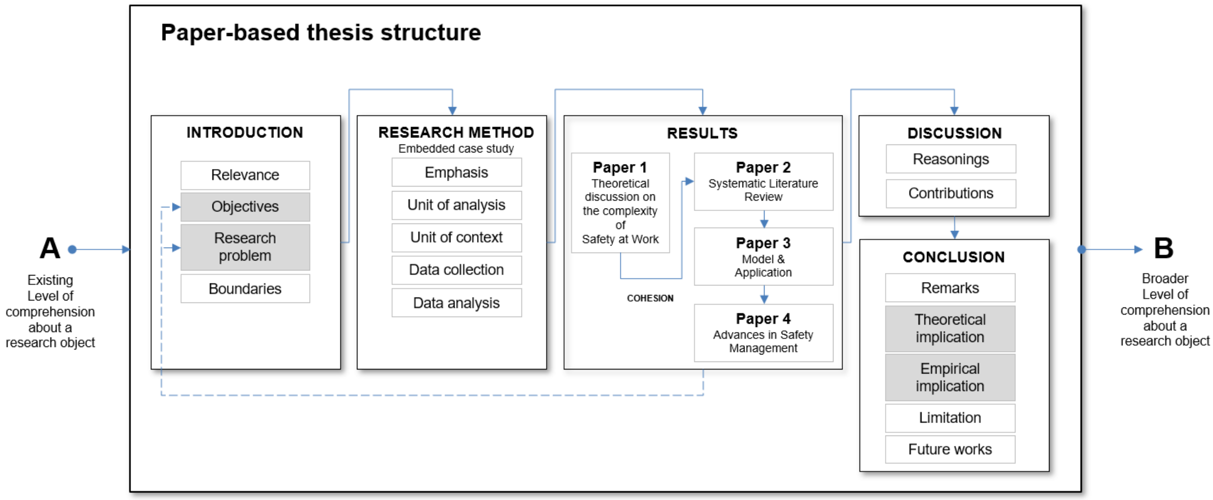


Figure 6 – Paper-based thesis structure

In the sequence of this Introduction, which presents initial research considerations, section 2 describes the research design along with relevant methodological issues.

Sections 3-6 are therefore composed of one out of the four articles comprising this thesis. Each article meets one or more research objectives, either general or specific (Table 2). The first article touches on whether safety at work is a complex or an exceedingly simple matter by following common strategies of theory building. The discussion about complexity is grounded in the Theory of Constraints, especially using the concept of inherent simplicity. The prevailing definition of complexity is confronted and a framework with the constructs that govern safety at work is explained.

The second article aims to investigate the concepts and instruments utilized to assess occupational health and safety management systems within organizations. More specifically, the study focused on the performance measures that have been associated with OHSMS assessment, and how the OHSMS initiatives (inputs), mechanisms, and outcomes (outputs) are linked. To do that, a systematic literature review was performed based on six previous reviews and twenty primary studies in compliance with the research scope. As a theoretical contribution from our findings, this research postulates a novel conceptual base for OHSMS assessment.

Through the outcomes from article 2, a model based on DEA to evaluate the efficiency of OHSMS is presented in article 3. The article is also focused on a real-world application in a contemporaneous context. It follows the structure of an embedded case study with longitudinal data conducted in a multinational company, characterized as a global player in the elevator industry. Besides the discussion about the findings, the applicability of DEA for OHSMS assessment is analyzed, as well as the use of efficacy and efficiency to compose a combined measure to offer a broader evaluation of OHSMS.

The fourth article addresses a business case study using a more advanced technique derived from safety data science. It aims to use association rules mining to unhide patterns of co-occurrence in work-related accidents and to evaluate its impact on efficiency before and after its implementation.

Finally, section 7 discusses the fulfillment of the research objectives, presents the limitations of the study, draws conclusions of this research, and encourages future research by presenting a set of directions.

Table 2 – Relationships between the objectives of the thesis, sections, and main contributions. Adapted from (Gauss, 2020)

		Specific objective	Section	Article	Scope	Main contributions
Primary objective	Evaluate the efficiency of Occupational Health and Safety Management Systems (OHSMS) using Data Envelopment Analysis (DEA).	1. Explore the causalities that govern safety at work based on the general method of theory building in applied disciplines. 2. Identify the constructs that govern safety at work, and discuss the complexity of this phenomenon based on the Theory of Constraints.	Section 3	Safety at work: a complex or an exceedingly simple matter?	Theoretical study to tackle the complexity of safety at work based on the concept of inherent simplicity stem from the Theory of Constraints. The study also provides an analysis of causal consistency between constructs and offers a conceptual postulation about whether safety at work is a complex issue.	• Identification of the constructs that govern safety at work. • Framework based on the theoretical propositions regarding the relationship between the identified constructs. • Analysis of causal consistency to explain the phenomenon based on the relationship of the constructs. • Conceptual definition for complexity in safety at work.
		3. Critically review the existing methods and assessment instruments for the evaluation of OHSMS, including the identification of critical elements and the conceptual foundation used in the examined instruments.	Section 4	Performance measurement in Occupational Health and Safety Management Systems: an analysis of critical elements comprising assessment instruments.	Systematic literature review comprised of 6 previous literature reviews (2007-2020) and 20 empirical studies published in peer-reviewed journals. Moreover, the study offers a theoretical proposition to drive new directions to the concepts of OHSMS performance measures, e.g. efficiency.	• Synthesis of OHSMS assessment instruments with regards to performance measurement. • Identification of the critical components comprising the examined literature: initiatives, outcomes, and mechanisms in both certified and non-certified contexts of OHSMS. • Analysis of the co-occurrence between contexts, initiatives, mechanisms, and outcomes. • Harmonization of concepts related to performance measures in OHSMS.

Primary objective	Evaluate the efficiency of Occupational Health and Safety Management Systems (OHSMS) using Data Envelopment Analysis (DEA).	4. Propose a DEA-based model to evaluate the efficiency of OHSMS. 5. Conduct a real-world case-based application to critically analyze the OHSMS efficiency in a relevant context, e.g. Agenda 2030. 6. Compare the results of efficacy and efficiency in OHSMS assessments. Also, evaluate a combined measurement of efficacy and efficiency into the OHSMS effectiveness.	Section 5	Measuring efficiency of safe work environment from the perspective of decent work agenda.	Case-based research with the application of a proposed model based on Directional Distance Functional (DDF), a similar frontier technique to Data Envelopment Analysis (DEA), to evaluate the efficiency of OHSMSs. The real-world case is characterized by an embedded longitudinal study carried out in a multinational organization to illustrate an empirical case.	• Definition of context, unit of analysis and variables for modeling. • A proposed DEA-based model for OHSMS assessment applied at the organizational level. • An empirical application to validate the model. • A critical analysis of the efficiency as a measure to evaluate safety performance. • A proposed combined measure to solve issues associated with the relative nature of the efficiency.
		7. Combine the analysis of efficiency with data mining techniques to open new opportunities for future research in safety data science.	Section 6	Application of Association rules to identify patterns of occurrence in work-related accidents: an analysis of the effect on efficiency	Business Case study with the application of association rules mining to identify patterns of occurrence in work-related accidents. Also, the study provides an architecture to implement the artifact in a real-world application, and analyze the results from the perspective of efficiency.	• Application of association rule mining based on the *Apriori* algorithm in a real-world case. • An artifact to be used in real-world applications is demonstrated in the existing systems of the organization. • A novelty approach of efficiency analysis combined with the use of mining techniques.

2 RESEARCH DESIGN

This thesis follows a case-based research strategy with an inductive approach. Case studies are appropriate when a phenomenon is broad and complex, when an in-depth investigation is needed, and when a phenomenon cannot be studied outside the context in which it occurs (Benbasat et al., 1987; Bonoma, 1985; Dubé and Paré, 2003; Yin, 1994). More specifically, an embedded longitudinal case study is conducted in an elevator industry to analyze the economic efficiency of the OHSMS in a set of business units, covering a period of 5 years. The use of single cases finds rationale when it represents a decisive case to confirm, contest or extend a well-formulated theory, or when it is considered a revealing case to observe and analyze a phenomenon.

This approach also is suitable for theory-generation emphasis since it fits the duality of being situationally grounded, but at the same time, seeking a sense of generality (Ketokivi and Choi, 2014). Moreover, according to Barratt et al. (2011), an important consideration for undertaking the theory-generation approach is to clearly articulate the rationable behind why such research is being conducted (Eisenhardt et al., 2007), e.g. the justification that research is explanatory (i.e., asking "how" and "why" types of questions) and the context and experiences of actors are critical, especially the experiences of managers so as to increase the practical relevance of the findings (Fisher, 2007). Central to this reasoning is induction, with empirical observations as driving force. The premise in theory-generating case research is, therefore, that in the context of the specific research question and empirical setting, explanation (theory) derives from exploration (analysis).

In the context of H&S management, assessment instruments can take a variety of forms. Thus, a case study supports the objectives of this research, which aims to evaluate the efficiency of OHSMS using a model based on Data Envelopment Analysis (DEA).

The study uses the method for conducting a case study design proposed by Cauchick-Miguel (2007), consisting of 5 main steps. Figure 7 illustrates the characteristics of each stage and the results generated. The continuous arrows indicate the order, while the dashed arrows show the feedback that can occur between the execution of the steps.

In general terms, the first stage accounts for the theoretical-conceptual structure where the problem is identified, and the awareness of the research problem is sought, reinforced by a systematic literature review. The literature review aims mainly to identify gaps to justify the research, as well as the constructs. As an outcome, the propositions of the work and its objectives are defined (Dresch et al., 2015b).

The next stage draws the planning case. The main step in this stage is defining the unit(s) of analysis and unit(s) of context. Also, the techniques for data collection, the means of analysis, and the research protocol are defined. Another prescribed stage is pilot testing. This stage is particularly important to verify the procedures, quality of the obtained data, and to establish improvements on the research protocol before the mass collection of data and analysis of results.

In the following stage, the data collection is conducted. Multiple sources of evidence might be taken into account for the research purpose, e.g. documents, performance reports, P&L, notes, interview records, and direct observations. The data collection is supposed to be completed when the amount of data is considered sufficient to address the research question.

Based on collected data, a narrative is produced to outline the analysis. In this stage, different approaches and techniques can be applied, such as quantitative analysis (e.g. DEA), content analysis, structural or aggregative. For more detail on methods of analysis see (Ermel et al., 2021). Finally, the results are communicated through a conclusion report.

Next section 2.1 presents the work method of the thesis.

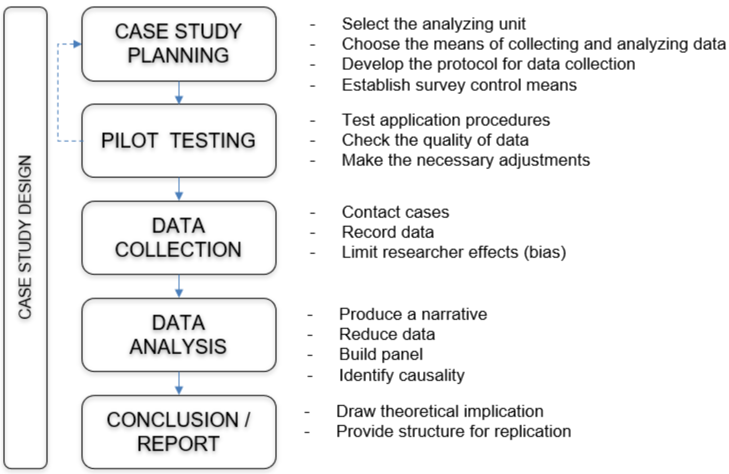


Figure 7 – Case study research (Cauchick-Miguel, 2007; Dresch et al., 2015b)

2.1 Work method

The work method of the thesis consisted of 43 procedures grouped in the 9 stages as shown in Figure 8.

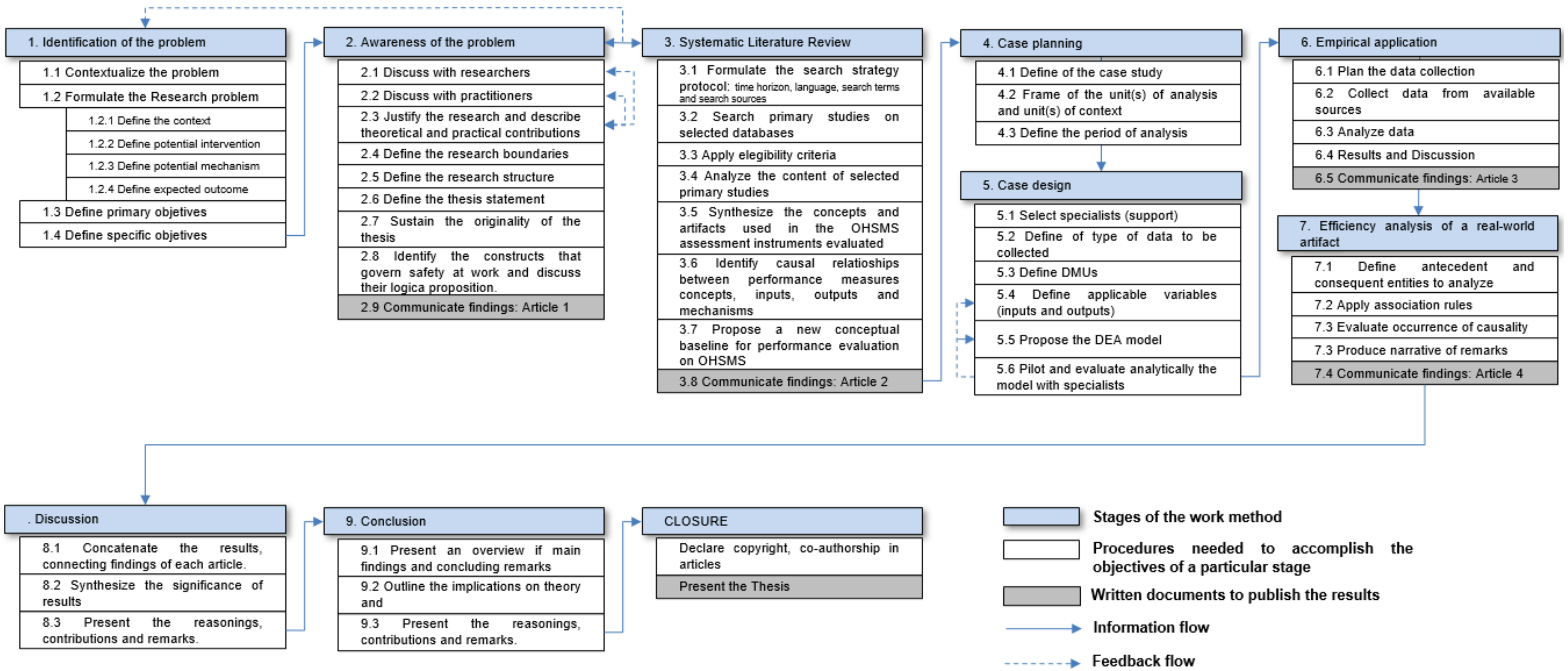


Figure 8 – Work method

The work method (Figure 8) was structured in stages, steps, and sub-steps (Figure 9). The stages follow the directives of the work method (Cauchick-Miguel, 2007; Dresch et al., 2015b; Piran, 2021a). The steps are procedures to be performed to accomplish a respective stage. Sub-steps are necessary routines to accomplish a specific procedure, evaluated as relevant to be highlighted in the work method.

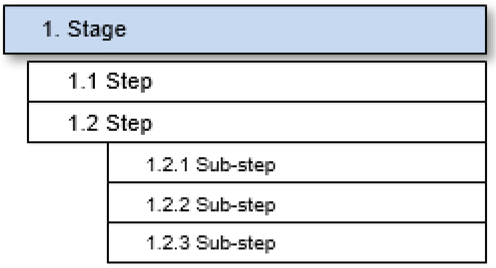


Figure 9 – Work method structure

The first stage of the work method consisted of identifying a problem to be solved. Four steps (1.1 to 1.4) were defined to address the research problem. Initially (1.1), it was sought to understand the context of the field of health and safety at work by reading International Labour Organization reports (ILOSTAT, 2021), including the main statistical data available on H&S, such as fatal and non-fatal occupational accidents reported by country region and economic activity, occupational disease reported by severity, risk factor, job characteristics, and time lost due to occupational injuries by economic activity. To deepen the context, the understanding of the statistical data was complemented by the reading of scientific articles that discuss the advancement (or not) of safety science, as well as peer-reviewed articles and some thesis addressing the performance evaluation of occupational health and safety management systems.

Secondly, CIMO logic was employed in the research problem formulation (step 1.2) as shown in Table 3. It is constructed as follows: “in this class of problematic Contexts, use this Intervention type to invoke these Mechanism(s), to deliver these Outcome(s)” (Denyer et al., 2008).

Table 3 – CIMO-logic. The components of design propositions

Component	Explanation
C – Context	The surrounding (external and internal environment) factors and the nature of the human actors that influence behavioral change. They include features such as age, experience, competency, organizational politics and power, the

Component	Explanation
	nature of the technical system, organizational stability, uncertainty, and system interdependencies.
I – Intervention	The interventions managers have at their disposal to influence behavior. For example, leadership style, planning and control systems, training, performance management. It is important to note that it is necessary to examine not just the nature of the intervention but also how it is implemented. Furthermore, interventions carry with their hypotheses, which may or may not be shared. For example, 'financial incentives will lead to higher worker motivation'.
M – Mechanism	The mechanism in a certain context is triggered by the intervention. For instance, empowerment offers employees the means to contribute to some activity beyond their normal tasks or outside their normal sphere of interest, which then prompts participation and responsibility, offering the potential of long-term benefits to them and/or to their organization.
O – Outcomes	The outcome of the intervention in its various aspects, such as performance improvement, cost reduction, or low error rates.

Source: (Denyer et al., 2008)

This technique was chosen to make sure that the right question is being made to solve the research problem (Ermel et al., 2021). After filling the elements of the technique, the research problem was formulated and presented in section 1.1.

Since the context was delimited, and the research problem is formulated, the primary objective (1.3) was defined as follows: "evaluate the efficiency of Occupational Health and Safety Management Systems (OHSMS) using Data Envelopment Analysis (DEA)". To reach this primary objective, seven specific objectives (step 1.4) were defined, as presented in section 1.2. Also, each objective was considered on one or more articles composing sections 3 – 6 of this thesis.

The second stage stood for raising awareness of the problem. Steps 2.1 and 2.2 were conducted with the support of specialists with theoretical and practical knowledge on topics related to efficiency, DEA, and Occupational Health and Safety Management. In addition, discussions were held with experts and managers to understand the needs of the market. The specialists and managers consulted and the justification for their choice are detailed in Table 4.

Table 4 – Specialists consulted to raise awareness of the research problem

Specialist/Manager	Reason for choice
Researcher in safety science	Dr. Andrew Rae, PhD. is a Senior Lecturer in the Safety Science Innovation Lab at Griffith University, where he teaches courses on research methods and safety engineering and manages the lab's research program. Drew's research uses a mix of ethnography, field experiments, and theory-building to investigate organizational safety practices. He is particularly interested in understanding the myths, rituals, and bad

Specialist/Manager	Reason for choice
	habits that surround the work of managers and safety practitioners, and how this work influences front-line operations. Drew co-hosts the Safety of Work podcast and is Associate Editor for the journal Safety Science.
Researcher in efficiency analysis	Dr. Ana M. Camanho Ph.D. is an Associate Professor at the Faculty of Engineering at the University of Porto, and author of more than 70 international indexed papers in management science. Her main research area is Operational Research and Data Science, with an emphasis on the development of models for evaluating efficiency and productivity evolution using the Data Envelopment Analysis Technique. Also, she has been involved in research projects in several areas, such as education, health, Corporate Social Responsibility, quality of life, and sustainable development of countries and cities.
Researcher, and consultant in economic efficiency with the application of the DEA	Dr. Fabio Piran is author of International publications and reference in conducting practical projects concerning economic efficiency using DEA within several organizations;
Senior executives in the Elevator Industry	The elevator industry has a well-consolidated OHSMS. Executives are open to applying the concepts of technical and economic efficiency with the use of DEA;
OSH Managers in different industries	Practitioners in H&S to support discussion involving OHSMS assessment, variables, and expected outcomes.
Executives of Brazilian Association of Machinery and Equipment (ABIMAQ)	Association represents the majority of machinery and equipment industries in Brazil and has a group dedicated to improving safety and health at work.

In step 2.1, it was sought to discuss potential methods for assessing OHSMS performance with specialists in efficiency analysis. The first specialist consulted is a senior lecturer, Ph.D. researcher, and associated editor of Safety Science, particularly interested in understanding the myths, rituals, and bad habits that surround the work of managers and safety practitioners, and how this work influences front-line operations. Through this support, the research question, constructs, boundaries, and originality of the thesis were discussed. The second specialist is an Associate Professor, an internationally well-known Ph.D. researcher in management science, with an emphasis on the development of models for evaluating the efficiency and productivity evolution using the Data Envelopment Analysis technique. From this discussion, it was found that the concept of efficiency as a measure of performance could be properly applied in the evaluation of OHSMS. Also, the originality of measuring the economic efficiency of OHSMS was confirmed. The third specialist holds a Doctoral degree in economic efficiency analysis and has conducted several practical works employing DEA for evaluating operational efficiency within organizations. Based on this discussion concatenated with the second specialist, DEA was defined as a potential mechanism for building the solution applied to OHSMS assessment.

The next step (2.2) was conducted to deepen the understanding of the OHSMS performance assessment from the practitioners' perspectives. OSH professionals from different industries and senior executives were consulted to inform how the H&S management is assessed in their organizations. It was clarified that organizations usually have H&S programs in place aimed at reducing accidents and that the topic has gained greater importance in strategic discussions. Also, it was captured that performance measurements are usually associated with 'number of days without accidents', 'number of work-related accidents or illnesses', frequency rate, and severity rate. In addition, senior executives confirmed that even considering health and safety as a core aspect of business performance, and great efforts have been made to promote safer working conditions within the organizations, results are not satisfactory.

From the discussion in step 2.2, It's argued that OHSMS performance assessment needs to be better addressed in a reality-based scenario. In general, the evaluation of the health and safety management system has not taken into account the resources used nor the relation between inputs and outputs. Therefore, efficiency is not analyzed and senior executives face a poor assessment instrument to drive continuous improvement on OHSMS.

As a result of step 2.2, the contributions of this study are presented in both theoretical and practical fields (section 1.3), and the research boundaries are highlighted in section 1.4. As an outcome of stages 1 and 2, the research design consisted of a paper-based thesis (PBT), structured in eight sections, and details are shown in section 1.5 (Table 2).

The research boundaries, main contributions, and structure used in this thesis are presented in steps 2.3 – 2.6. Steps 2.7 – 2.8 support the assumption efficacy does not offer the necessary view of critical factors in H&S performance and define the thesis statement. A deeper approach concerning the understanding of the research problem and its complexity is explored in step 2.8 and presented in step 2.9.

Stage 3 is related to the systematic literature review. The process started in step 3.1 with the formulation of the research protocol, as presented in Table 5. The research protocol synthesizes the research design (Thomé et al., 2016), and it is considered a mechanism whereby the research transparency and replicability are promoted (Higgins et al., 2021).

Table 5 – Research protocol

Research title	Occupational Health and Safety Management System: A Systematic Literature Review of Assessment Instruments	
Researcher:	Gomes, Rodrigo Frank de Souza	**Revision**: 00 **Date**: 23.05.2021
Stakeholders:	OHS practitioners; senior executives; public representatives	
Research questions:	(RQ1) which performance measures, concepts, and instruments have been used to assess an OHSMS? (RQ2) what are the mechanisms employed? (RQ3) how the initiatives (inputs), mechanisms, and outcomes (outputs) are linked?	
Research objectives:	To investigate the concepts and the instruments utilized to assess occupational health and safety management systems within the organizations. In addition, to identify how the OHSMS initiatives (inputs), mechanisms, and outcomes (outputs) are linked.	
Review scope:	**Amplitude**: Narrow **Deepness**: Deep	**Review type:** Aggregative
Theoretical framework:	Theoretical roots are presented in Figure 2, and the OHSMS framework is shown in Figure 3	
Time horizon	Not applied	
Search strings:	TITLE-ABS-KEY ("safety management system" AND "systematic review") TITLE-ABS-KEY (("occupational health and safety management system" OR "OHSMS") AND ("performance" OR "efficiency" OR "efficacy" OR "productivity" OR "effectiveness")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")) TOPIC ("safety management system" AND ("systematic review") TOPIC: (("occupational health and safety management system" OR "OHSMS") AND ("performance" OR "efficiency" OR "efficacy" OR "productivity" OR "effectiveness")) Refined by: DOCUMENT TYPES: (ARTICLE)	
Search sources:	Databases: Scopus and Web of Science	
Searching approach:	[X] Direct searching	
Eligibility criteria:	**Inclusion criteria:**	Articles reviewed per peers and published in the selected databases, in English
	Exclusion criteria:	Not accessible online; out of research scope based on Title/Abstract reading
Data analysis:	**Scientometric analysis:**	Not applied
	Bibliometric analysis:	Not applied
	Content analysis	Aggregative
Data synthesis:	Aggregative synthesis	

Source: Adapted from (Ermel et al., 2021; Morandi and Camargo, 2015)

Different from other broader literature reviews about OHSMS, this research is particularly interested in OHSMS performance evaluation, more specifically in the concepts and assessment instruments utilized. The next step (3.2) was to search primary studies on selected databases and to apply the eligibility criteria established in the research protocol (3.3). Figure 10 presents the generic flowchart of searching and eligibility of studies.

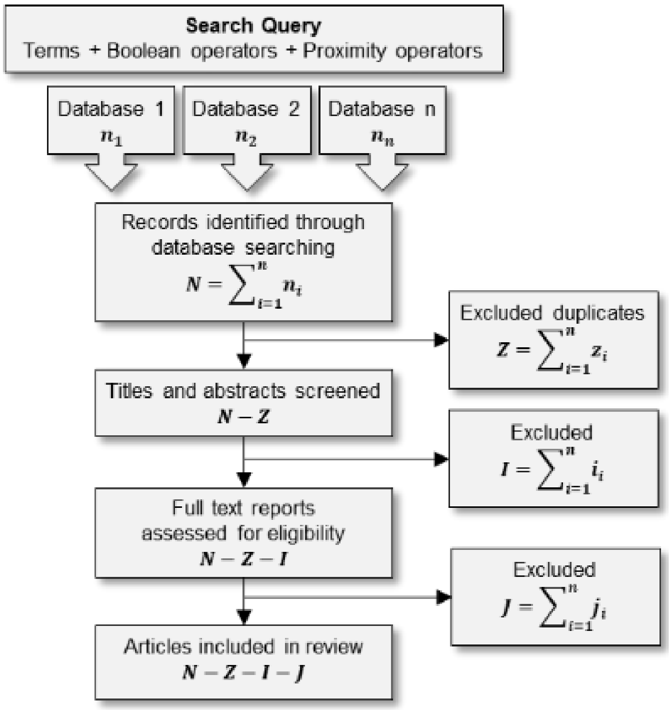


Figure 10 – Flowchart of searching and eligibility of studies (Gauss, 2020)

In step 3.4 the content of the corpus of analysis was analyzed in-depth (A. Bardin, 1993) by using qualitative data analysis software (ATLAS.ti (v.8.4.24), 2020). The outcome of this process was to synthesize and aggregate the concepts, classes of problems, and artifacts used in the OHSMS assessment instruments evaluated, as shown in Table 6.

Table 6 – Generic aggregative synthesis of OHSMS Assessment Instruments

Classes of problems	Design problem	OHSMS assessment		Performance Measurement concept	CIMO				Primary studies
		Techniques	Methods		C	I	M	O	
Cp_1	Pb_1	T_1	Mt_1	PMc_1	C_i	I_{ij}	M_i	O_{ij}	$R_1, R_2, R_4, \dots R_n$
	Pb_2	T_2	Mt_2	PMc_2	C_i	I_{ij}	M_i	O_{ij}	$R_3, R_6, \dots R_m$
	Pb_n	T_n	Mt_n	PMc_n	C_i	I_{ij}	M_i	O_{ij}	$R_5, R_7, \dots R_k$

Source: Adapted from (Gauss, 2020)

The classes of problems are defined as a grouped set of either theoretical or practical problems that comprise useful artifacts to be employed by organizations to solve problems (Dresch et al., 2015a). The classes of problems allow artifacts may be applied to general contexts, instead of only in a specific one, being useful for solving similar problems faced by researchers and practitioners.

Table 7 – Examples of classes of problems and OHSMS assessment instruments

Class of problem (C_p)	Example of assessment instruments (artifact)
OHS management	PDCA, Risk-based hierarchy
OHSMS performance	MAI (Michigan Assessment Instrument), E.I. (Efficacy Index)
Efficiency analysis	DEA, OEE
Strategy deployment	A3, Strategy & Tactics Tree (S&TT)

Artifact, such as an assessment instrument is, in turn, something artificially manmade to accomplish a purpose, e.g. a mathematical model or a technique, based on the internal environment (functional requirements) to achieve a satisfactory solution (objective) in the external environment (Simon, 1996). In the context of H&S management, the assessment instruments are artifacts of a class of problems that is OHSMS performance evaluation.

The OHSMS assessment instruments identified in the primary studies were grouped in step 3.5 based on the technique or model applied to solve a specific problem Pb_n (e.g. how to measure OHSMS efficiency?) of a defined class of problems Cp_n, which is the OHSMS performance assessment. In addition, it also verified the rigor of the theoretical concept of performance measurement by considering the specialized literature (Coelli et al., 2005; Piran et al., 2020). Finally, each primary study was associated with a context, its inputs and outputs variables considered in the performance evaluation, and the mechanisms invoked to achieve the outcomes.

The next step (3.6) accounted for identifying underlying causal relationships between performance measures concepts, inputs, outputs, and mechanisms considered in primary studies. To do that, association analysis represented in the form of association rules (Zhang and Zhang, 2002), such as $\{lhs\} \rightarrow \{rhs\}$, is used to uncover those hidden relationships, as shown in Figure 11. The left-hand side (lhs) of the rule is called antecedent, while the right-hand side (rhs) is named as consequent. Some measures indicate the strength of an association rule, which include: support, confidence, and lift. Support determines how often a rule applies to the corpus of analysis, while confidence determines how frequently the consequent items appear in relationships containing the antecedent items. Lift computes the leveraging of consequent items whenever the antecedent items occur (Ermel et al., 2021; Gauss et al., 2020).

```
lhs                rhs   support confidence coverage lift      count
[1] {O33}       => {O21} 0.05    1          0.05     2.50000  1
[2] {O12}       => {O21} 0.05    1          0.05     2.50000  1
[3] {M4}        => {O21} 0.05    1          0.05     2.50000  1
[4] {O24}       => {O21} 0.05    1          0.05     2.50000  1
[5] {I1b}       => {O21} 0.05    1          0.05     2.50000  1

lhs                rhs   support confidence coverage lift      count
[1] {O21}       => {O23} 0.20    0.5000000  0.40     2.500000 4
[2] {CL.2,O21}  => {O23} 0.15    0.5000000  0.30     2.500000 3
[3] {M5,O21}    => {O23} 0.10    0.6666667  0.15     3.333333 2
[4] {PM.3}      => {O23} 0.05    0.5000000  0.10     2.500000 1
[5] {I1b}       => {O23} 0.05    1.0000000  0.05     5.000000 1

   lhs            rhs     support confidence coverage lift      count
[1] {O22}       => {O34} 0.05    1.0000000  0.05     10.00000 1
[2] {PM.2}      => {O34} 0.05    0.5000000  0.10     5.000000 1
[3] {O14}       => {O34} 0.05    0.5000000  0.10     5.000000 1
[4] {I17}       => {O34} 0.05    0.5000000  0.10     5.000000 1
[5] {O22,PM.2}  => {O34} 0.05    1.0000000  0.05     10.00000 1
[6] {CL.1,O22}  => {O34} 0.05    1.0000000  0.05     10.00000 1
```

Figure 11 – Association rules. Retrieved from (R core team, 2018)

Based on a set of rules, e.g. association rule [3] involving the outcome O_{23} as a consequent factor, an inference shall be read as follows: the output O_{23} (reduction of the lot-time injury rate) is more likely to be achieved (66.67%) when it is associated with the output O_{21} (reduction of accidents) and with the mechanism M_5 (engagement). Another example is the association rule [6] concerning the outcome O_{34} shall be understood as: in the context of certified organizations (CL_1), the reduction on hazard exposure (O_{22}) is associated with the increase of economic KPIs, such as Sales and gross margin (O_{34}).

As an outcome of the content analysis and structural analysis, findings attested to a lack of conceptual rigor employed in defining different types of performance measures in the OHSMS context. Based on well-recognized concepts (Coelli et al., 2005; Farrel, 1957), an initial conceptual baseline is proposed in step 3.7, and shown in Table 8. Finally, results and contributions are presented in step 3.8 (section 3 article 2).

Table 8 – Performance concepts in OHSMS context

Concept	Definition
Efficacy	Efficacy is represented by the percentage of achievement of targeted outcomes, e.g., accident rate. For this particular evaluation, the resources used do not influence the level of the efficacy index.
Efficiency	Efficiency is a relative measure and has to do with the optimal use of resources. An efficient OHSMS focuses on maximizing its resources (inputs) to produce intermediate, final and economic outcomes. Thus, efficiency in OHSMS does not necessarily mean target achievements. It represents how close the system is to the efficiency frontier, which means the optimal ratio between all inputs and outputs.
Effectiveness	Effectiveness comprises both efficacy and efficiency. That means a multi-axial evaluation concerning target achievement and optimal use of resources. A highly

	effective OHSMS is one in which targets are reached under an optimized inputs/outputs ratio.

Stage 4 began with the case planning and scope definition. The main aspect of this stage is to properly address the imperative issues that are being investigated. At this point, the rationale for selecting the locations where the study will take place, the propositions between constructs, the hypotheses to be examined, and the broader relevance of the investigation are included (Yin, 1994). The selection of the case study (4.1) and the definition of the units of analysis and unit of context (also found in literature as sub-unit of analysis) is the starting point and should bring cohesion and coherence to the research problem (Figure 12). The unit of analysis needs to be associated with the research question that was previously defined. It is also called a primary unit of analysis. Besides, if there are contextual circumstances in which the unit of analysis has been studied, it is fundamental to define the unit of context. In that reasoning, the researcher needs to interpret results from the observations of the unit of context, and present reasonings related to the unit of analysis (represented as dashed arrows in Figure 12). For example, examining the occupational health and management system of an organization seems to be clear enough. However, it can be true in the case of a single entity. For larger organizations, the observations should be done at the business unit level, which is a particular context where the phenomena can be verified. This explains why defining the unit of analysis and unit of context is critical for case studies. The granularity of the analysis shall be carefully evaluated to reach the research question and further propositions for the research problem.

Finally, as the case study represents a profound deep dive into investigating a phenomenon, the period of analysis (4.3) is framed to precisely situate the time horizon when the study is conducted to support further inferences, such as propositions between constructs and variables.

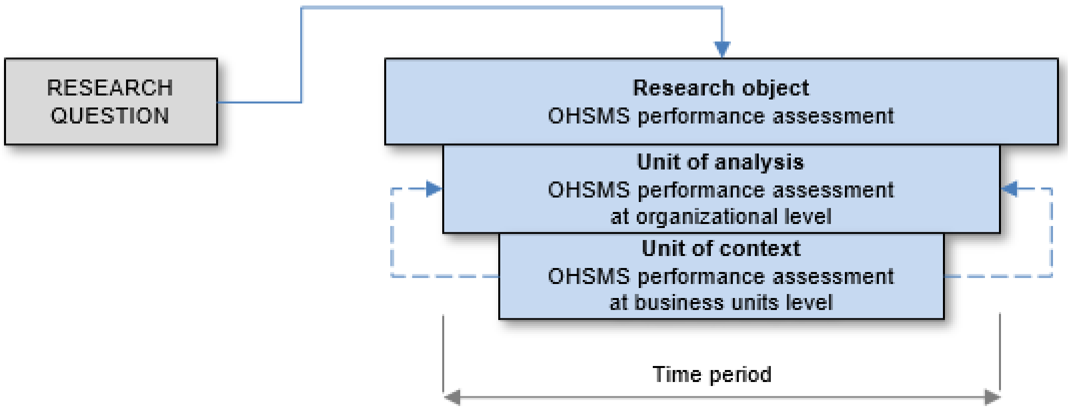


Figure 12 – An embedded case study approach. Adapted from (Piran, 2021b; Yin, 1994)

The following stage accounts for the proposition of the model to address the research problem previously declared in section 1.1. In steps 5.1 – 5.4, through the specialists (Table 4), the type of data, means of data collection, the decision-making units (DMUs), and variables for inputs and outputs are defined. In this study, the DMUs are the business units of an organization (O), i.e. $O = \{DMU_i\}, i = 1,2 \ldots, n$. This approach is adequate for efficiency analysis since recent literature has shown the potential of internal benchmarking to explore efficiency over time and the impact of management actions on business performance (Piran et al., 2021).

By following a similar approach to steps 2.1 and 2.2, the variables were defined with the support of specialists with theoretical and practical knowledge on topics related to efficiency (technical and economic), DEA, and Occupational Health and Safety Management. Figure 13 introduces an idea of how the model (as a possible satisfactory solution for the problem) is supposed to close the gap between the research problem and a satisfactory solution.

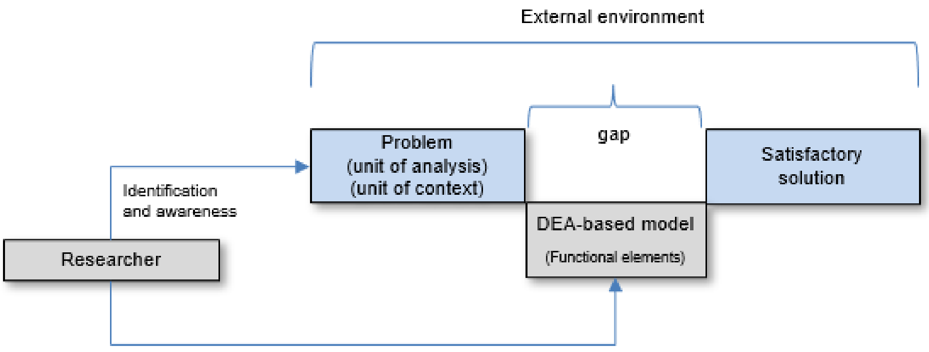


Figure 13 – Case design

In step 5.5, a generic solution is preliminarily defined based on the research problem, requirements (e.g. quantitative approach, applicable for internal benchmark, possibility of simulating different scenarios), and primary objective (Table 9). At this point in time, a DEA-based model is then formalized with its functional elements, e.g. variables for inputs, outputs, and model orientation, and formulations in terms of the returns to scale (CRS/VRS).

Table 9 – Proposition of solutions

Class of problem (C_p)	Research problem (P)	Requirements (Req)	Generic solutions
OHSMS performance	How to evaluate the efficiency of OHSMS within organizations?	(1) Req_1 (2) Req_2 (3) Req_3 (4) ... (5) Req_n	(1) S_1 (2) S_2 (3) S_3

Finally, in step 5.6, after the definition of DMUs, variables, and orientation, the proposed model is formalized through a mathematical model, generically represented by the CRS input-oriented model (1). It illustrates a formulation with constant returns to scale (CRS), and input-oriented. A DEA model also can be configured as output-oriented, and with variable return to scale (VRS). The main differences between configurations are presented in Table 10 and Figure 14.

$$Max\ e_{ff0}(OHSMS) = \frac{\sum_{j=1}^{m} u_j y_{j0}}{\sum_{i=1}^{n} v_i x_{i0}} \tag{1}$$

$$s.t.$$

$$\frac{\sum_{j=1}^{m} u_j y_{jk}}{\sum_{i=1}^{n} v_i x_{ik}} \leq 1, \forall k$$

$$u_j \geq 0\ , \forall j$$

$$v_i \geq 0\ , \forall v$$

eff_0 = efficiency of DMU_0 under analysis

u_j = weight calculated for the output $j, j = 1, .. m$

v_i = weight calculated for input $i, i = 1, .. n$

y_{j0} = quantity of output j for the DMU under analysis

x_{i0} = quantity of output j for the DMU under analysis

y_{jk} = quantity of output j for the DMU $k, k = 1, \dots, m$

x_{ik} = quantity of input i for the DMU $k, k = 1, \dots, n$

k = number of DMU under analysis

m = number of outputs

n = number of inputs

Table 10 – DEA modeling configurations

Formulation	Objective Function	Application	Input-oriented	Output-oriented
CRS	Linear	CRS is more attractive when DMUs are homogeneous, and all operate under similar conditions. Besides, in CRS the output increases by the same proportional change of the inputs, which reasoning has to make sense in the analysis.	Minimize inputs w/ constant outputs	Maximize outputs w/ constant inputs
VRS	Linear	A VRS assessment implies that DMUs are only compared to others DMUs of roughly similar size. It is recommendable when inputs and outputs vary greatly among DMUs. Also, under VRS, input and output-oriented analysis will give different measures of efficiency for inefficient DMUs.		

Source: Adapted from (Camanho, 2018; Piran, 2015)

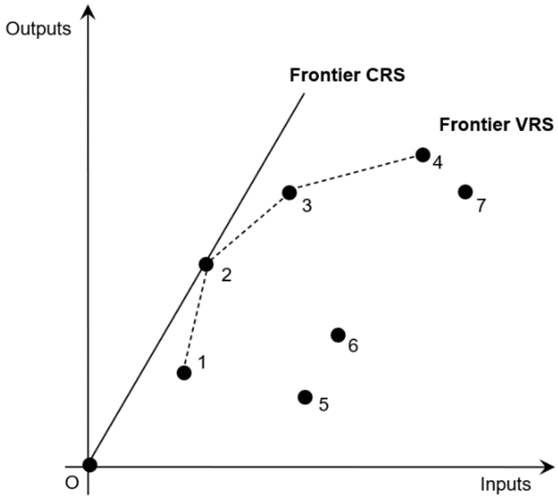


Figure 14 – CRS/VRS Models (Piran et al., 2020)

Furthermore, a computational application in R (R core team, 2018) was used to run the model as depicted in Figure 15. R is a language and environment for statistical computing and graphics. It is an integrated suite of software facilities for data manipulation, calculation, and graphical display. Also, it provides a wide variety of statistical (linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, clustering, etc) and graphical techniques, and is highly extensible (“What is R?,” 2021).

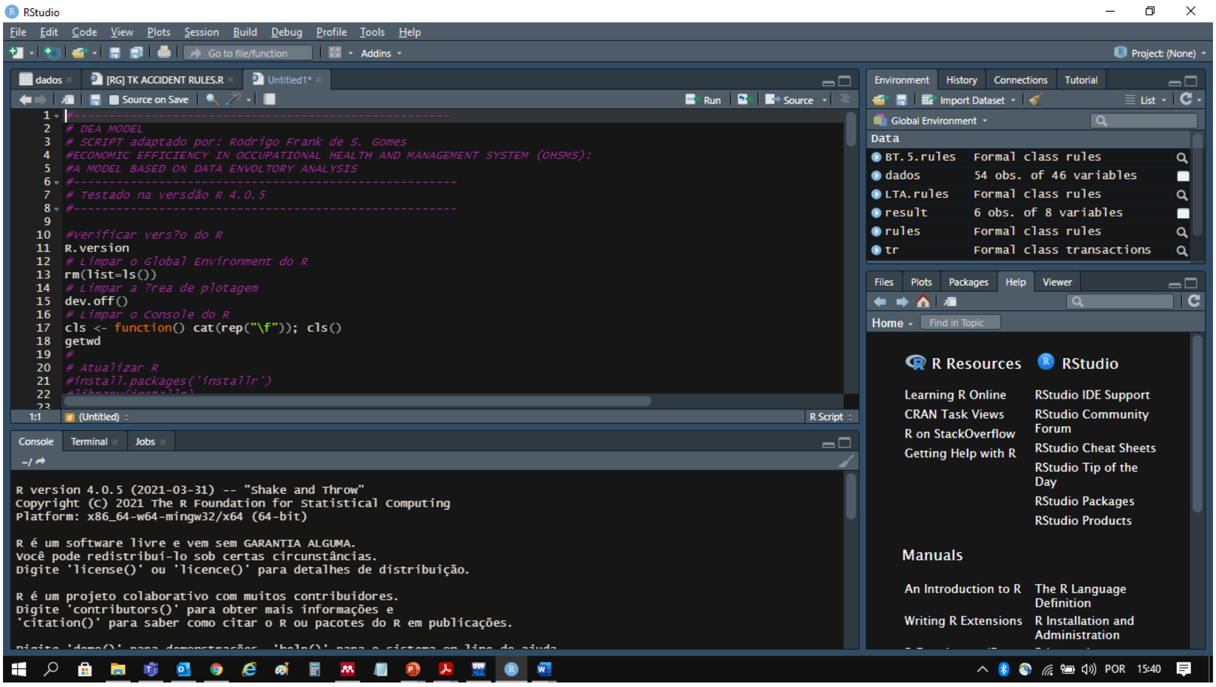


Figure 15 – Example of Functional stage in R Script (R core team, 2018)

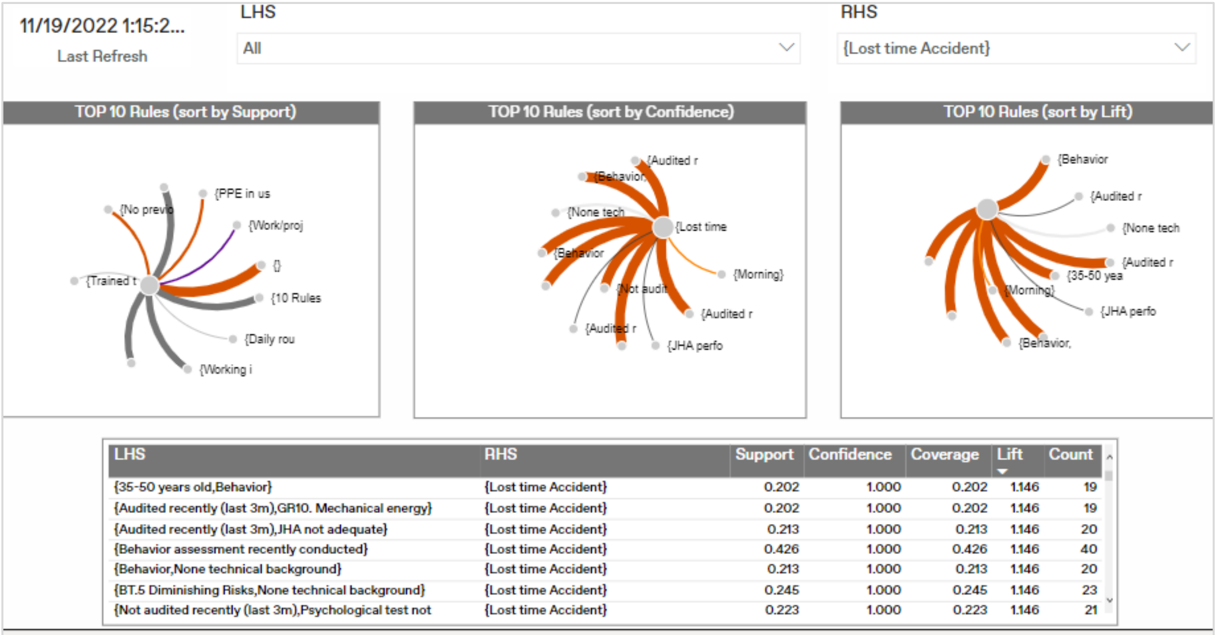
In an OHSMS, the inputs are represented by the resources employed to achieve the outcomes, such as labor cost, training hours, personal protective equipment, tools, etc. On the other hand, the outputs can be split between intermediate (e.g. safety climate), final (e.g. number of accidents), and economic outcomes, such as cost savings, insurance, etc (Robson et al., 2007).

Before the empirical application, a pilot is conducted to verify application procedures based on the protocol, aiming at its improvement. From this application, it is also possible to verify the quality of obtained data, to identify if they are associated with the constructs, and, thus, contribute to addressing research objectives (Dresch et al., 2015b). Since the model aims to be grounded in the reality-based application (Rae et al., 2020), practitioners are fundamental to accomplish this step. Also, DEA specialists supported this process to analyze the structural elements of the model, including the formulation and orientation chosen.

Since the learnings from pilot testing are incorporated into the model, the next stage is the empirical case itself. The case study was conducted based on the structure proposed by Cauchick-Miguel (2007) as previously mentioned in the research design (Figure 7).

In step 6.1, the data collection is planned based on the requirements previously defined in step 5.5 (Table 9). Next, once all necessary data is collected, the data is analyzed, and reasonings are presented based on the results in steps 6.2 – 6.4. The findings from the empirical application are communicated in step 6.5 corresponding to article 3 in section 5 of this thesis).

Stage 7 offers a real-world application using association rules mining to reveal patterns of occurrence in work-related accidents as shown in Figure 16 (steps 7.1 – 7.2). Results were combined with efficiency analysis in step 7.3, and causality was investigated in step 7.4. Finally, in step 7.4 the results are communicated corresponding to article 4 (section 6 of this thesis).



LHS	RHS	Support	Confidence	Coverage	Lift	Count
{35-50 years old,Behavior}	{Lost time Accident}	0.202	1.000	0.202	1.146	19
{Audited recently (last 3m),GR10. Mechanical energy}	{Lost time Accident}	0.202	1.000	0.202	1.146	19
{Audited recently (last 3m),JHA not adequate}	{Lost time Accident}	0.213	1.000	0.213	1.146	20
{Behavior assessment recently conducted}	{Lost time Accident}	0.426	1.000	0.426	1.146	40
{Behavior,None technical background}	{Lost time Accident}	0.213	1.000	0.213	1.146	20
{BT.5 Diminishing Risks,None technical background}	{Lost time Accident}	0.245	1.000	0.245	1.146	23
{Not audited recently (last 3m),Psychological test not	{Lost time Accident}	0.223	1.000	0.223	1.146	21

Figure 16 – Real-world case application of association rules mining

Finally, the results are concatenated and the findings of each article are connected in an integrated, comprehensive, and cohesive approach, discussing the outcomes of the research as a whole (steps 8.1—8.3). General remarks, implications to theory and practice, and trends on OHSMS performance assessment under the author's perspective are outlined in a conclusion section (steps 9.1 – 9.3).

3 SAFETY AT WORK: A COMPLEX OR AN EXCEEDINGLY SIMPLE MATTER?[4]

Abstract

This paper uses the concept of inherent simplicity stemming from the Theory of Constraints to explain whether safety at work is a complex or an exceedingly simple matter. In this context, the study seeks to explore the causalities that govern safety at work, identifying its constructs and presenting logic propositions based on the theory-building blocks: classification, correlation, and causal consistency. To support the research, a dataset composed of 46 work-related accident investigation reports from an elevator industry in Latin America was carefully analyzed using association rules. Moreover, direct observations grounded on inductive reasoning were used to speculate plausive causes concerning the effect of work-related accidents. The research strategy followed common strategies of theory building to reach common sense: theory-to-practice and practice-to-theory. As a result, a conceptual proposition is postulated based on the reasoning that safety at work is governed by very few constructs, and that its complexity is explained through the two elements from inherent simplicity: degrees of freedom (interdependencies between constructs) and harmony (conflicts resolution within the work environment). From the practitioners' perspective, the study also offers directions towards safety improvements at the organizational level by considering the impact of the interdependencies between constructs in safety at work.

Keywords: Inherent simplicity. Safety at work. Theory of constraints. Theory building. Causation.

3.1 Introduction

The field of safety science is advancing very slowly, despite an increasing volume of research activity and publication (Rae et al., 2020). On one side, a massive body of knowledge is available in literature in a form of cases, frameworks, mathematical models, and systematic literature reviews. On the other side, practitioners are struggling to improve safety practices within organizations without considering theories and published shreds of evidence. While this disharmony between theory and practice in safety science is verified, society remains to deal with social and economic impacts arising from ineffective safety management.

According to ILO (2021), more than 2.8 million deaths per year result from occupational accidents or work-related diseases. When considering non-fatal work-related injuries, this number increases to approximately 376.8 million a year. Moreover, the burden resulting from such ineffective safety management accounts for economic

---

[4] Gomes, R.F.S., Gauss, L., Piran, F.S., Lacerda, D.P., 2022. Safety at Work: a Complex or an Exceedingly Simple Matter? Reliab. Theory Appl. 17, 267–287. https://doi.org/https://doi.org/10.24412/1932-2321-2022-167-267-287

losses estimated at 3.94% of the global Gross Domestic Product (Brocal et al., 2018; ILO, 2020; Wang et al., 2020).

This pragmatic reality shall draw the attention of researchers and practitioners due to its impact on society. This is because a healthy and safe work environment not only is desirable from the workers' perspective but also contributes considerably to labor productivity and promotes economic growth (Heuvel et al., 2017). Furthermore, safety at work promotes worker motivation, increases productivity by reducing costs related to work-related health problems, and relieves pressure on public and private health systems.

Based on such a challenging scenario, a step back seems to be necessary. Rather than propose solutions to address just a piece of this issue, it is necessary to make sure that safety at work is well understood in academia and within organizations.

In this context, this paper aims at identifying the constructs and presenting propositions to explain the causalities that govern safety at work. In addition, this study explores how the definition of complexity should be understood in the field of safety science, and what is the prevailing definition. This is fundamental to draw attention to the main factors that affect safety, and how their interdependencies might increase or decrease the complexity of the system.

In that reasoning, the theoretical discussion of this study is structured on building blocks proposed by Whetten (1989), and consistent with the three stages of science proposed by Goldratt (1990): classification, correlation, and causation consistency. As a major theoretical outcome of this research, the causalities that govern safety at work and its complexity are explained through the two elements of inherent simplicity: degrees of freedom (interdependencies between constructs) and harmony (determined by the belief that every internal conflict can be removed by eliminating improper assumptions).

From a managerial's perspective, this study is useful for practitioners to put efforts into critical constructs that impact the overall safety management system to make it simpler and harmonious, instead of acting to reach local optima.

Finally, this study also has a side contribution in extending the applications of the Theory of Constraints (TOC) to the field of safety. Since literature is particularly lacking in investigative studies on the theoretical and practical implications of TOC principles (Ikeziri et al., 2018), this research contributes to closing this gap since no previous study is found connecting inherent simplicity and safety science.

This article is organized as follows: section 3.2 outlines a comprehensive review of the concept of inherent simplicity. The work method is described in section 3.3. In section 3.4 the results are presented, and a narrative of theoretical discussion is conducted. Finally, the main conclusions and limitations of the study are summarized in section 3.5.

3.2 Inherent simplicity

The concept of inherent simplicity is a principle from the Theory of Constraints (Goldratt, 1990) in which it is postulated that any part of reality is governed by very few elements and that any conflict can be eliminated (Goldratt, 2008). In its earliest stage, TOC focused on production system optimization before being recognized as an operations management theory to foster the process of ongoing improvement. Further on, TOC became a global management philosophy applied to various areas such as production, supply chain, project, and other fields (Ikeziri et al., 2018). In the theoretical field, TOC also satisfies the virtues of a good theory, such as uniqueness, parsimony, and generalizability (Naor et al., 2013).

Goldratt (1990) outlined that TOC is grounded in its practicability, and unlike in common sense, "theory in science must be practical, otherwise, it is not theory but just an empty scholastic speculation" (p.32). This is consistent with the assumption that the purpose of good theory shouldn't be other than describe and explain how things actually work, and in so doing to help us improve our actions in this world (Lynham, 2002).

The concept of inherent simplicity can also be understood as a practical way of viewing reality. However, reality usually looks complex to us, and Goldratt took for granted the foundation of modern science from Newton: "*Natura valde simplex est et sibi consona*" (nature is exceedingly simple and harmonious with itself). It does mean that if we deep dive enough into observing phenomena, we'll find that there are very few elements at the base that govern the whole system. Reality is, therefore, built in wonderful simplicity (Goldratt, 2008).

The interpretation of Goldratt from Newton's quote is also consistent with the principle of bounded rationality (Simon, 1957, pp. 198-199): "the capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world". In other terms, the key to simplification of the choice process is rather the goal of "maximizing", the goal of "satisfying", i.e. finding a course of action is good enough". This association of concepts was postulated by Eden and Ronen (2007) and in-deep described by Naor et al. (2013) for further readings.

The prevailing definition of complexity is that the more entities the system has, the more complex the system is. Thus, by following this approach to compare the complexity of the systems 'A' and 'B' represented in Figure 17, system 'B' is more complex than 'A' because the quantity of entities that comprise the system 'B' is higher than 'A'. However, since we are more interested in understanding, predicting, and controlling the system instead of just describing it, this study follows Goldratt's approach to define complexity by the following: the more degrees of freedom the system has, the more complex it is (Goldratt, 2008).

The concept of degrees of freedom might be clear for physicists or engineers but it is not under overall comprehension. In short, Goldratt explains that it means the minimum number of points (or entities) you have to touch in order to impact the whole system. For example, in the case of system 'B', by impacting the bottom circle, the whole system is impacted, i.e. it has only one degree of freedom. On the other hand, system 'A' has five degrees of freedom, which is harder to control and predict due to its magnitude. This becomes clear by observing the absence of arrows in the system, which means that there are no interdependencies between the entities. Figure 17 illustrates the reasoning of complexity based on inherent simplicity.

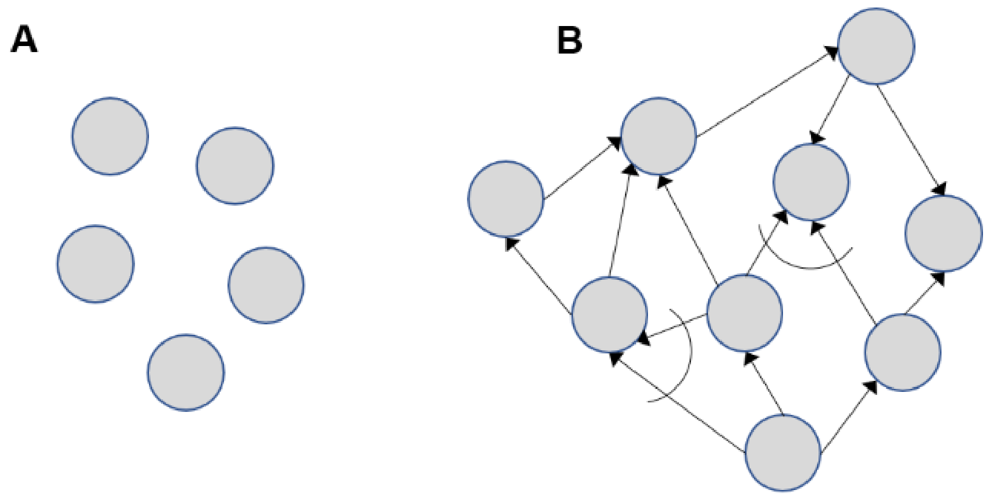


Figure 17 – The reasoning of complexity (Goldratt, 2008).

3.3 Research design

This study is based on 18 months of direct observations and primary data analysis concerning investigation reports of work-related accidents occurred in an elevator industry. The industry's activities are spread out over 12 countries across Latin America, covering one industrial facility in Brazil and more than 75 service operating units across the region. During this period, the first researcher had close contact with a reality-based source of data, in which scope it is included both manufacturing and service areas in the twelve countries where the organization has an operational presence.

The work method used both common strategies of theory building: theory-to-practice and practice-to-theory (Lynham, 2002; Swanson and Chermack, 2013) as shown in Figure 18.

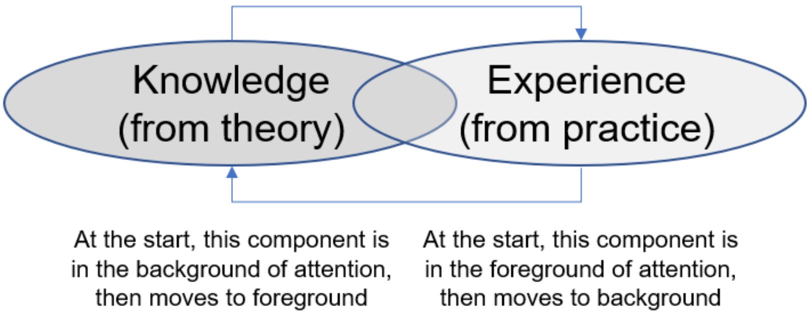


Figure 18 – General method of theory building in applied disciplines. Adapted from Lynham (2002) (Lynham, 2002).

Initially, the researchers observed an effect: the occurrence of work-related accidents as an issue with significant social and economic impacts worldwide. Then, following the stages proposed by Goldratt (1990), the focus moved to speculate plausive causes to explain this phenomenon. To do that, a research question was therefore defined, and awareness about the research problem was sought based on specialized literature.

The next step accounted for the use of a theory-to-practice approach to assume that very few constructs govern safety at work. In that reasoning, the principle of inherent simplicity derived from TOC was reviewed and the theory was framed in the field of safety. As a second stream, the research moved on to the practice-to-theory approach through reality-based data collection to analyze and come up with theoretical and practical contributions to safety science, exploring how and why the constructs that govern safety at work are interconnected and seeking to uncover underlying issues to explain its complexity. A detailed step-by-step of the work method is depicted in Figure 19.

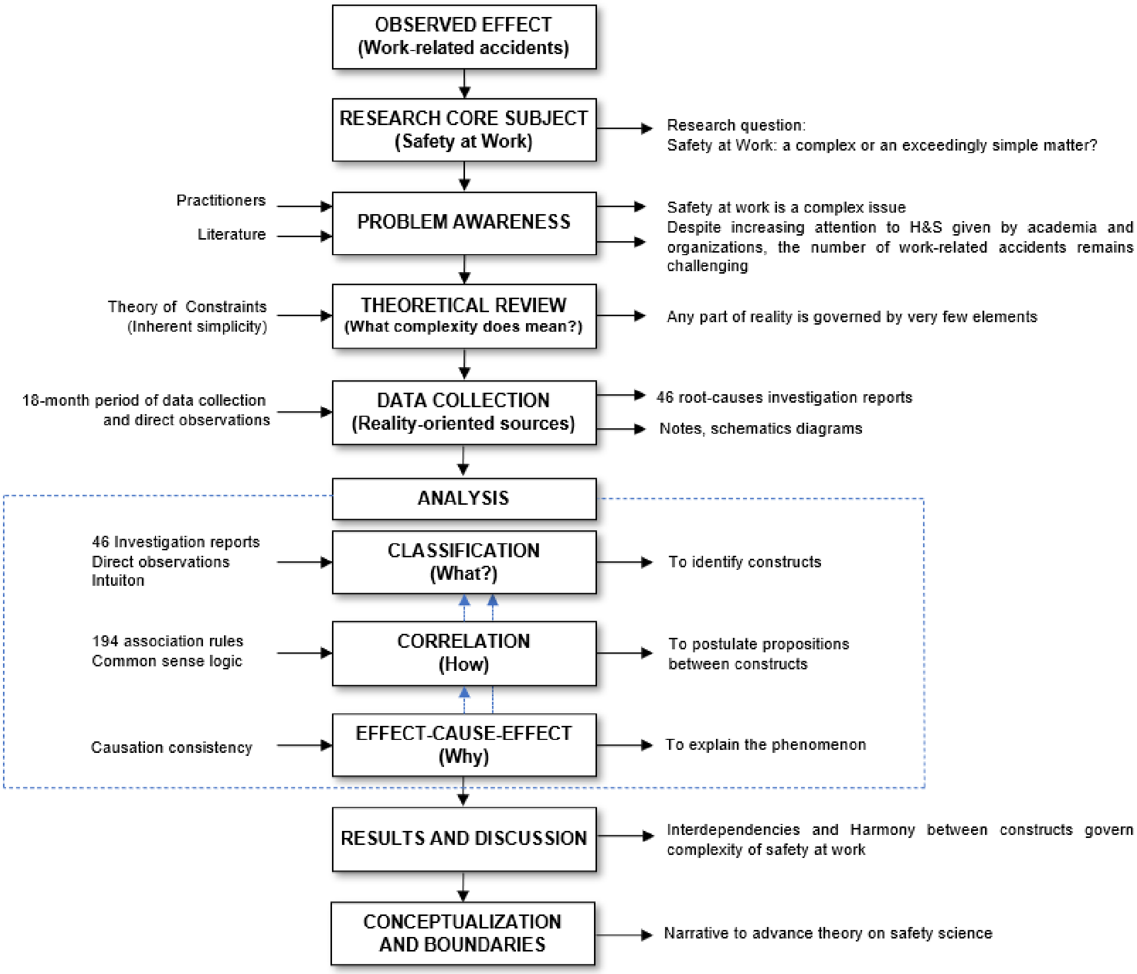


Figure 19 – Core research subject.

The first researcher examined in depth the existing body of documents in the occupational health and safety management system (OHSMS), the structural functioning of the case unit, and how health and safety (H&S) fits into the organization's strategic planning. Also, several job site visits were conducted to observe how the work is done, the resources available, level of technical knowledge, procedures, routine instructions, task planning, and personal protective equipment (PPE) usage.

Data retrieved from the OHSMS was studied through a business intelligence (B.I.) dashboard covering the period between oct-19 to mar-21. Forty-six root-causes investigation reports listed in Table 11 were collected and analyzed with the support of three specialists. The specialists are H&S managers in charge of the three main operations within the organization: the factory located in Brazil, field operations in Brazil, and field operations in other Latin American countries. In addition, an organizational psychologist supported the discussion when behavioral aspects were reported as contributive causes to the accidents.

Table 11 – Root-causes investigation reports

Country	Working hours	Root-causes investigation reports derived from lost-time accidents	
		Factory	Services
Argentina	403,000	-	1
Brazil	14,000,000	1	28
Chile	1,387,000	-	4
Colombia	1,256,000	-	2
Costa Rica	91,000	-	1
Mexico	978,000	-	3
Panama	2,918,000	-	2
Paraguay	372,000	-	1
Peru	1,049,000	-	1
Uruguay	163,000	-	2

Each root-cause investigation report followed a structured template based on 9 categories and 41 data fields (see appendix A 1). The outcome of this analysis was to identify and classify the most frequent factors that impacted work-related accidents.

Moreover, a data mining through the algorithm *Apriori* was powered to identify association rules between factors, i.e., what antecedent factors (named *lhs*) impact the other consequent ones (named *rhs*), and how strong is this correlation. It consists of a data mining algorithm that systematically controls the exponential growth of candidate itemsets (Zhang and Zhang, 2002). The parameters support (supp=0.5), and confidence (conf=0.8) were set up as thresholds based on adopted criteria from previous studies (Isa et al., 2018; Kouzis-loukas, 2014).

The parameter support determines how often a rule applies to a given dataset. Besides, it aims to identify the most relevant rules (Alves, 2020) in the dataset. Confidence, in turn, determines how frequently consequent factors [*rhs*] appear in relationships that contain antecedents [*lhs*]. It is used to measure the strength of an association rule, expressed as the times a specific itemset is found together with a specific item out of the total times this specific itemset is found in the entire dataset (Kouzis-loukas, 2014). In other words, the greater confidence of rule {X} ⇒ {Y}, the greater the probability of {Y} being present in events that contain {X} (Tan et al., 2018).

An additional measure used in that research is the 'lift'. The lift of an association rule is responsible for measuring the difference between the number of times {X} and {Y} co-occur and the expected frequency of such co-occurrence if they were statistically independent (Hahsler and Hornik, 2008). In that reasoning, high levels of lift mean that the consequent factor is scarcer within the population and more frequent within the specific itemset.

For this step, a script loaded in software RStudio was used for data processing (see appendix A 2). Additional explanations about the use of association rules can be found in the work of Zhang and Zhang (2002) and other mentioned literature. Furthermore, examples of how to explore cause-effect relationships using association rules in the H&S field can be found in the studies of Cheng et al. (2010), Mirabadi and Sharifian (2010), and Verma et al. (2014).

Through this technique, 194 associated rules were retrieved to support the correlation stage. The structure of rules is presented in Table 12 and can be interpreted as follows: based on a dataset with N events, the rule [n1], for example, associates the antecedent factor A to the consequent factor C. The support of this rule can vary between 0 – 1. A minimum support threshold is used to select the most frequent (and hopefully important) factors' combinations. Confidence, similarly, is understood as an estimate of the conditional probability of factors co-occur in a rule (0 – 1).

Finally, the lift value of 1 indicates that the factors are co-occurring in the database as expected under independence. Values greater than 1 indicate that the items are associated, and lower than 1 indicate an absence of association (Hahsler and Hornik, 2008).

Table 12 – Structure of association rules

Rule	lhs		rhs	support	confidence	lift	count
[n1]	{antecedent A}	=>	{consequent C}	0 - 1	0 - 1	0 - ∞	1 – N
[n2]	{antecedent A, antecedent B}	=>	{consequent D}	0 - 1	0 - 1	0 - ∞	1 – N

Besides the investigation reports, other general documents were carefully analyzed, e.g the strategic planning 2020-2025, OHSMS manual, and H&S policies.

From these documents, it was possible to situate expected management commitment as well as H&S in the strategic context of the organization, in order to check against reality through direct observations.

Direct observations were conducted in the course of the same period of the primary data collection. It followed as possible, a semi-structured approach as follows: (1) to verify the work being performed, such as the use of tools and personal protective equipment, printed instructions, work environment, etc; (2) to conduct an informal conversation to understand the task routine, capabilities required to the task, and capacity to foreseeing risks; (3) to verify the leadership commitment from the worker's perspectives, and possible behavioral impacts from externalities, such as COVID-19, personal issues. Yet, the informal approach was given to avoid the feeling of pressure when formal questions for interviews could bring up.

Moreover, additional factors were observed at the job sites beyond the technical field. The education level and behavioral aspects, such as lack of concentration and lack of awareness were considered as well. Also, the observations were not limited to job sites. Management meetings and reactions from the occurrence of accidents were also observed. Preliminary speculations from the direct observations were registered in notes and schematic diagrams to reach common-sense logic. Furthermore, confirmation questions were frequently used at the end of any informal approach: "if I understood well this effect was caused by this fact. Am I right?".

The relevance of the direct observations is based on the fact that it is rarely found whether in literature or in reality-based practices, pieces of evidence related to explain the safe work, i.e. a deep analysis of what went good, and the factors that led to a work environment in which safety culture is intrinsic. As an outcome of the use of both association rules and direct observations, a framework is proposed to explain the causalities that govern safety at work since it allows the researcher to observe, in practice, the effect-cause-effect stage.

Based on the framework elaborated, the first researcher was encouraged to use verbalized intuition with other researchers and practitioners (Goldratt, 1990) to practice simplicity, parsimony, and to reach common sense.

In that reasoning, principles of causal consistency derived from the Theory of Constraints Thinking Processes were also used to explain each proposition presented in the framework: causality existence, causality clarity, the sufficiency of cause, and additional cause (Ermel et al., 2021). As a result, a conceptualization of complexity in safety at work is postulated.

In the next section, results are discussed throughout a combined approach of the three main stages that every science has gone through (Goldratt, 1990; Whetten, 1989). The classification stage was associated with the ‘what’, correlation with the ‘how’, and effect-cause-effect with the ‘why’. Finally, the researchers sought to define limitations in time and context for the propositions. These contextual factors are critical to set the boundaries of generalizability in which the propositions are postulated.

3.4 Results and Discussion

3.4.1 Classification (building block ‘what’)

This stage sought to explore what constructs logically impact safety at work. In this context, the criteria of comprehensiveness and parsimony supported the researchers to determine whether a factor should be considered as a variable to explore the causalities of safety at work. In short, it was sought for relevance and value-added of each variable to explain phenomena (Whetten, 1989). One primary instance of identifying these constructs was based on an inductive approach and intuition. Initially, it was considered plausive factors that influence phenomena (safety at work). For instance, technical expertise is a plausive factor to impact positively safety. However, even in case of considering this example a common sense, it does not explain what is its level of importance, how this factor is connected to others, and what is its effect on the whole system.

In addition, the analysis and classification of primary data and the findings obtained through direct observations supported the researchers in that stage. Numerous factors came up with this process, including training, task planning, years of experience, education level, availability of proper tools, personal protective equipment usage, adequate instruction. However, at this point in time, no correlation was checked, and each factor was considered an independent one. In that reasoning, consistent with the concept of inherent simplicity, the system primarily seemed to be very complex (see Figure 20).

Figure 20 – Factors that impact safety at work.

The next step was to practice simplicity and parsimony, considering that theory should have a minimum of complexity and few assumptions. Each variable was considered as a potential factor to impact safety at work. Next, every variable was associated with a construct as a theoretical element wherein the variable is encompassed. A minimum number of constructs was sought in order to reach simplicity and decrease complexity.

In that reasoning, after the data analysis, an interactive process of verbalizing the factors grouped in constructs with other researchers, H&S experts, and workers was conducted to reach common sense. In this context, variables were grouped into constructs to reach a higher level of abstraction, keeping the properties of comprehensiveness. For instance, variables such as technical training, safety training, hazard analysis were grouped into the construct 'knowledge'. This is because 'knowledge' encompasses several factors associated with the necessity of knowing, for example, 'what to do', 'how to do', 'what are the risks involved, 'how to mitigate the risks'.

As an outcome of this stage, a set of constructs were defined as satisfactory based on the logic of 'good enough' (Eden and Ronen, 2007) to explore phenomena of interest (see Figure 21). This is because these four theoretical elements (knowledge, planning, behavior, and performance measure) sufficiently encompass in a form of constructs all variables identified in the classification stage.

In the next subsection, the propositions between how these constructs are connected are outlined.

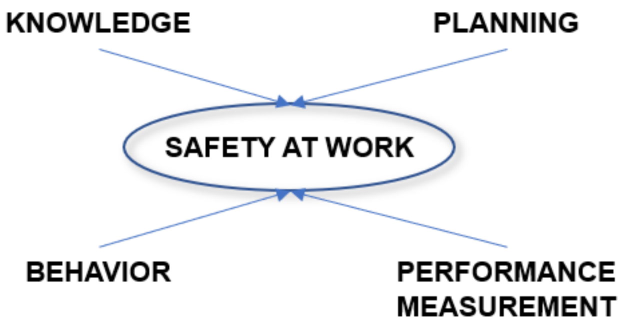


Figure 21 – Constructs associated with safety at work.

3.4.2 Correlation (building block 'how')

Once the minimum necessary constructs to explore phenomena of interest are identified, the next stage aimed to define how they are connected (co-related). Although this stage is based on careful observations and often involves a quantitative approach, the question 'why' is not asked at all. Rather the question 'how' is the center of interest (Goldratt, 1990). Based on that reasoning, the propositions were structured with the use of 194 association rules, as shown in Table 13.

Table 13 – Association rules

Rule	lhs		rhs	support	confidence	lift	count
[34]	{Inappropriate JHA}	=>	{Lost time Accident}	0.6415	1	1.1522	34
[70]	{Trained to the task}	=>	{Diminishing Risks}	0.6038	0.8889	1.1778	32
[76]	{Trained to the task}	=>	{Lost time Accident}	0.6792	1	1.1522	36
[79]	{Diminishing Risks}	=>	{Daily routine}	0.6415	0.8500	1.1551	34
[100]	{Diminishing Risks}	=>	{Lost time Accident}	0.7547	1	1.1522	40
[145]	{Trained to the task,Working in regular time}	=>	{Unappropriate JHA}	0.5283	0.8000	1.2471	28
[155]	{On-time, Trained to task}	=>	{Diminishing Risks}	0.5283	0.9333	1.2367	28

Also, the researchers sought to take benefit from the direct observation of works being performed safely. This is because the set of investigation reports analyzed is about 'how things went wrong' (unsafe work). However, seeking for broadening the research perspective, the researchers also focused to verify 'how things go safe' (work safely), to confirm some association rules and intuition. According to Whetten (1989), although the researcher may be unable to test all the links (propositions between constructs), restrictions in methods do not invalidate the inherent causal nature of theory. In this reasoning, and consistent with the understanding that most of what passes for theory in organizational studies consists of approximations (Weick, 1995), the connections and the propositions between constructs are introduced in the framework depicted in Figure 22.

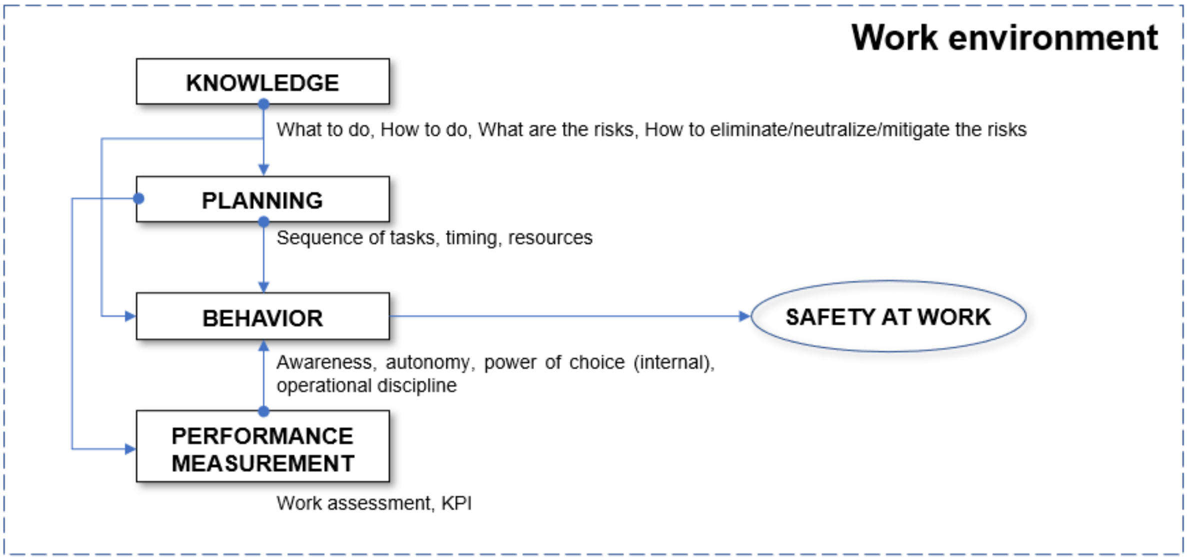


Figure 22 – Propositions that govern safety at work.

The framework is comprised of fours constructs, and it should be read as the following narrative: knowledge is the starting point. It is represented by work elements such as 'what to do', how to do', 'what are the risks' and 'how to eliminate/neutralize/mitigate the risks'. Knowledge is a construct presented in every type of work. This is consistent with the investigation reports analyzed and coherent with the direct observations conducted throughout the research. In both situations of work safely or work unsafely, knowledge (or the lack of knowledge) is present as a plausive construct that partially governs and explains phenomena of interest. In the case of safety at work, it also represents a baseline since common sense is that knowledge is critical for working safely.

However, knowledge is necessary but far away as sufficient to explain phenomena safety at work. This is consistent with the association rules, e.g. rules [70, 76, 155]. According to those rules, even workers trained to perform their tasks can get involved in lost-time accidents. This association is highly represented in rule [76], in which 36 out of 46 investigation reports analyzed, the worker was trained to the task in question (confidence =1; lift = 1.1522). Moreover, our observations confirmed that trained workers might diminish risks due to possible reasons, such as their work experience or due to the fact they never had a work-related accident before. Thus, other plausive constructs are necessary to explain what governs safety at work.

Knowledge is connected to construct planning. This reasoning is explained by conceptualizing planning as the way the work is expected to be done, in which sequence of tasks, timing, and with what resources. Following this logic, it sounds clear that ‘to plan’ depends on ‘to know’. By defining a good sequence of tasks, a standard operational procedure, or an estimation for a set of tasks to be completed, it is fundamental to know what is this activity about, how the activities are performed, and what resources are available. Planning also represents the way of performing a task. Well-defined tasks are the ones where the resources, timing, and logical sequence of each activity are established to raise productivity without taking out safety is a core aspect.

The question to be responded at this point in time is whether knowledge and planning are sufficient to defining the minimum constructs that govern safety at work. If so, an expert performing a well-planned task would be ever working safely. Our intuition indicates not, and also the association rules, e.g. rules [34, 59, 145] in which confidence and lift present a high level. Firstly (rule [34]), the lack of operational discipline in doing job hazard analysis (JHA) is associated with trained works. It means that even experts do not follow the planning. Second (rule [59]), resources such as personal protective equipment do not guarantee safety at work. Investigation reports indicated that very often accidents occur with employees equipped with PPEs. This suggests such a level of personal confidence that nothing wrong can happen, and risks are ignored. Finally (rule [145]), diminishing risk is highly associated with lost time accidents, and therefore, the behavioral aspect is another plausive construct to be considered.

In this context, behavior is a comprehensive construct. It is present in the literature in numerous studies about accident prevention, such as in the studies of Han et al. (2010) and Li et al. (2015). Also, motivation and work behavior are present in a robust body of knowledge in social sciences (Porter et al., 2002). Consistent with the existing literature, the results of the association rules put light on the effects of behavior in the work environment, verified in the consequent factor ‘diminishing risks’, and based on its high association with lost-time accidents (rule [100]). This comprehensiveness is expressed in the proposed framework through the fact that behavior is the most interconnected construct in the system. All other constructs are connected to it, and it is the only one directly connected to the work.

In that reasoning, both knowledge and planning are connected to construct behavior by one of the two directional flows presented in the framework. Both constructs impact the way a person behaves at work. This was verified through direct observations carefully conducted besides the association rules. For instance, consider a worker performing maintenance services. If he/she lacks the required knowledge about what to do and how to perform a repair, or if the worker does not know the risks associated with the task, the potential risk for an incident to occur is increased as the worker tries to perform the task. Also, if the timing defined for the service is inadequate, or if necessary resources are not available, the worker's behavior is impacted negatively, leading towards the opposite direction of safety at work.

Behavior is, therefore, a key construct in the proposed framework. In the context of this research, it is represented by four elements: awareness, autonomy, power of choice, and operational discipline. Each of these elements plays an important role in safety at work.

Awareness is the state of being conscious of something. More specifically, it is the ability to directly know and perceive, or to be aware of events. Autonomy, in turn, is a condition of self-government, and that needs to be outlined by managers. It is an important element to neutralize risks arising from externalities.

Next is the power of choice, which means the attitude of using awareness and autonomy in every decision at work. Finally, operational discipline means doing the right thing, the right way, every time. It encompasses the other constructs towards promoting safety at work.

From another direction, behavior is also impacted by another construct, represented by the way workers are measured. The performance measurement did not come up with the analysis of investigation reports. Rather, it emerged through the inductive approach and it is consistent with the theory of constraints. Goldratt (1990) pointed out that the way an organization defines its work assessment and KPIs impact how workers behave at all levels. For instance, even in the case of an expert performing a well-planned task, if the KPIs are not consistent with the timing required for the task and with the resources available, the behavior is impacted.

This is deeply explained by social cognitive theory (SCT), which explains behavior in organizations in terms of the reciprocal causation among the person, the environment, and the behavior itself (Porter et al., 2002). Because of these combined influences, under SCT organizational participants would at the same time be products and producers of their motivation, their respective environments, and their behaviors. In that reasoning, SCT and TOC justify the connection between performance measurement and work behavior.

Finally, performance measurement is also connected with planning. This is because performance assessment is intrinsically related to a comparison between what is realized versus what was planned. Moreover, KPIs and targets are typically defined based on strategic planning and organization capabilities (resources). For instance, the expected sales growth rate of a Retail store is defined in management reviews. The organization may expect more sales if more sellers are working for them, or, in the case of use of technologies to increase sales, e.g. web platforms. Both examples are resources, and resources are associated with planning.

The four constructs are interconnected in the boundaries of the work environment, as previously depicted in Figure 22. It represents a system to explain what are the constructs that govern safety at work, and how they are connected.

The work environment is characterized by both external (e.g. market regulations) and internal (organizational culture) existing factors in any work environment that might impact positively or negatively any construct. It plays a critical role for safety at work since it acts directly in promoting (dis)harmony between the connections, and therefore, affects the level of complexity as further explained in subsection 3.4.4.

Internal consistency and parsimony were sought to sustain every proposition's argument. Each construct in the system has a certain number of in-out connections. In this context, behavior represents the central construct because it is connected with all constructs and it is directly connected with phenomena safety at work. It follows the reasoning of considering 'to behave' an expression of 'acting', such as 'working'. Therefore, work behavior is positively or negatively impacted by knowledge, planning, and performance measurement, and all framed into the work environment.

The next section seeks for exploring the causation consistency.

3.4.3 Effect-cause-effect (building block 'why')

The previous sections were extremely helpful. 'What' and 'How' provide a framework for interpreting patterns in empirical observations (Whetten, 1989). However, only 'why' explains phenomena.

Existing literature in the field of safety science often lacks explaining causation, being limited to verified correlations. The inherent limitation of any correlation, e.g. findings from association rules, is the lack of understanding of the cause-and-effect relationships between the propositions (Goldratt, 1990). After identifying the constructs and exploring the reasoning of how they are connected, the next stage accounted for asking the question why?. In other words, the researchers are focused on what might be causing the existence of each proposition to explain safety at work as the effect of interest.

This stage is aimed no longer just to observe what already exists to explain phenomena, but also to use logical derivations based on existing causes to uncover underlying issues and predict the outcome of entirely new situations. Moreover, this stage accounts for fulfilling the minimum requirements of the conceptualization phase of theory building (Swanson and Chermack, 2013).

At this theory-development stage, logic replaces data as the basis for evaluation (Whetten, 1989). This is consistent with the use of common sense proposed by Goldratt (1990) to go through the effect-case-effect stage. Goldratt outlines that it represents the third stage of science, and the most important one because only at this stage there is a widely accepted recognition that the subject is, actually, a theory-building.

Therefore, the starting point of this stage is to become aware of an effect. The 'effect' of interest in this research is 'safety at work', and in the context of this study, safety at work means the action of working safely. "One effect is enough", said Dr. Goldratt, and the effect comes together with a challenging question: Is 'safety at work' a complex or exceedingly simple matter?

Once the effect and a challenging question are defined, more information is not much needed. Rather, to think and to speculate of plausive causes grounded in common sense are the next step (Goldratt, 2008, 1990; Swanson and Chermack, 2013). To do that, principles of causation consistency derived from the theory of constraints thinking processes are applied for each proposition: causality existence, causality clarity, the sufficiency of cause, and additional cause (Ermel et al., 2021). In that reasoning, the causal consistencies are presented in 5 through a narrative for each connection (Table 14), and thus, the framework is translated into confirmable propositions or knowledge claims to an explicit connection between the conceptualization phase and practice (Cohen, 1989).

Table 14 – Causation consistency

Connection	Causal consistency
Knowledge → Planning	Knowledge is presented in every type of work. In the context of safety at work, it is a baseline. Knowledge impacts planning because 'to plan' any activity requires knowledge about the nature of the work to be performed. Causal existence is evidenced by examples to sustain that this connection is always the case. For instance, to plan the construction of a house, a common sense is that a body of knowledge is necessary, e.g., what raw materials are required, the method of how to do it, the sequence of tasks, the risks involved in the work, and what other resources are needed. This reasoning is applied to construction but also any other type of work. Planning might also be impacted by the work environment, in which the proposed framework is represented by the boundary via dashed line (see Figure 22). This is because both external (e.g., macroeconomy, market regulations) and internal factors (organizational culture) existing in any work environment might positively or negatively any construct.
Knowledge and Planning → Behavior	Knowledge and planning are necessary but not sufficient to explain safety at work. Even experts performing well-planned tasks might work unsafely. A common sense to explain why knowledge and planning are not enough is to consider the behavior at work. If a worker behaves diminishing risks or if presents a lack of awareness, the knowledge and planning will not be sufficient at all. Therefore, by common sense, behavior is another necessary construct to explain the phenomena of interest. However, it is still needed to explain the causal existence of this proposition. It is assumed the way a worker behaves performing a task is impacted by his/her knowledge and how well the task was planned. This logic is explained also by examining accidents associated with knowledge in two ways: (1) the worker with the proper knowledge to perform a task and the one with a lack of knowledge to do so. In the first case, the proper knowledge can lead the worker to behave and work safely, but also an excess of confidence can lead to failures in following safety procedures. In the second case, the lack of necessary knowledge can lead the worker to unconsciously put himself/herself at risk. The same reasoning is applied to planning. If the sequence of tasks is carefully designed, proper resources are available, and timing is adequate for the task (a general harmony), the worker with autonomy and power of choice is predicted to work safely. This explanation put light on the causal existence and clarity of this proposition. However, sufficiency is not reached yet. There is speculation that people within the organizations are responsive to the way they are measured. By considering it as a plausive, relevant, and necessary construct to explain the complexity of safety at work, performance measurement (as a construct) was added to the framework.
Planning → Performance Measurement	Performance measurement is connected by planning. This connection is intrinsically observed in management reviews and strategic planning. The definition of key performance indicators (KPIs) considers the organization's planning because it takes into account capabilities, resources, timing, and the work environment influences. For instance, typical planning for the construction of vertical buildings in Brazil varies between 36 and 48 months. This general planning cascades several other sub-plannings to define all

Connection	Causal consistency
	that is needed to accomplish each phase of the project. KPIs for each phase and each task are also defined. Therefore, clarity and the existence of causation between planning and performance are verified. Another way to reach common sense that performance measurement is impacted by planning is by exploring the main KPIs of an industry. Productivity, for instance, is a performance measure that considers the ratio outputs/inputs. To increase productivity, practitioners evaluate how the activity is planned to be performed, including resources usage, quality of processes, and lead times. Following that reasoning, a KPI defined without taking into account planning sounds like no sense.
Performance Measurement → Behavior	Within organizations “people behave under influence of how they are measured”. This quote retrieved from principles of the theory of constraints (Goldratt, 1990) is consistent with the existing literature about social cognitive theory (SCT) which explains behavior in organizations in terms of the reciprocal causation among the person, the environment, and the behavior itself (Porter et al., 2002). It is important to highlight that behavior is the most interconnected construct in the proposed framework. Based on both theories it is assumed that the way a worker behaves at work is impacted by how the performance is measured, and also by his knowledge and how well is the planning of the task to be performed. Clarity of this proposition can be reached by examining productivity. For instance, consider a production line used to produce 22 elevators per day (just quantity). This level of productivity is consistent with the resources available (machinery, personnel, and tooling), and all workers are focused only on pushing forward the production line to reach the target. However, based on some organizational changes and observing that the production was also full of wastes, managers decide to consider efficiency instead of production volume as the performance measurement. Then, workers start to carefully look after the inputs to avoid any waste to maximize efficiency. This example comes up with pieces of evidence of why performance measurement impacts behavior. In this logic, the behavior characterized by a higher level of attention to avoid wastes was influenced by the changes in the performance measure.
Behavior → Safety at work	Finally, behavior is directed connected to safety at work, because in the context of this research it means the phenomena of working safely (co-existence). In more practical words, the action of working safely. Behavior is, therefore, a key construct in the proposed framework due to its high interconnection with other constructs. Moreover, besides being impacted by knowledge, planning, and performance measurement, it represents the utmost connection to the phenomena, expressed through a few elements such as worker’s awareness, autonomy, power of choice, and operational discipline. The existence of causation between behavior and safety at work is well-known in literature and also between practitioners. This is consistent with the concepts of behavior-based safety (BBS), as well as voluntary safety programs within organizations to raise safety awareness as a tentative to prevent accidents. Each of the mentioned elements of behavior at work plays a critical role in safety at work. In the instance of safety at work, they encompass the action of doing the right thing, the right way, every time.

3.4.4 The complexity of safety at work

A major outcome from the stages of classification, correlation, and causation consistency, is to underlying the issues that govern safety at work, and therefore, its complexity. Through the comprehension about what minimum constructs are sufficient to explain safety at work, how they are connected and why, this research's seed is postulated:

Proposition: The complexity of safety work is a function of the degrees of freedom and harmony between constructs that govern the work environment within an organization.

Every organization has an unique system as depicted in Figure 22, represented by the individual and collective knowledge, the work planning, and the performance measurement system. The way these constructs are connected impacts the behavior of workers, and therefore defines the complexity of safety at work.

Although each connection between constructs has generalizability, which means that it can be verified in every organization, it does not mean it is harmonious. The concept of inherent simplicity is grounded in two main beliefs: simplicity and harmony: Simplicity is expressed by the fact that there are very few elements that govern the whole system. Harmony, in turn, is expressed by considering that any conflict can be eliminated (Goldratt, 2008).

The framework and propositions depicted in Figure 22 follow the same reasoning that Figure 17 (B). It demonstrates that a system to represent safety at work might be exceedingly simple. This is possible since the system is comprised of four interconnected constructs that represent only one degree of freedom. However, this is necessary but not sufficient. The harmony between constructs is also a key factor.

Organizations usually face serious problems to properly address well-defined internal processes, and local optima is preferable instead of thinking as a whole. Moreover, problems arise from conflicts and disharmonies. As a result, organizations increase the number of system's degrees of freedom, fail in eliminating conflicts, and tend to address safety as a very complex matter.

This explains the challenges often faced by larger organizations. For instance, the disconnection between the planning department and the operations (who perform the work) or changes in the performance measurement system without taking into account the resources needed, causes disharmony and adds degrees of freedom to the system. Following the inherent simplicity concept, more points have to be touched by management in that case.

Therefore, we postulate that the complexity of safety at work is based on inherent simplicity, governed by very few constructs (knowledge, planning, performance measurement, and behavior), and simply explained as a function of the system's degrees of freedom and harmony between of constructs that govern the work environment within an organization.

3.5 Conclusion

This study was framed into the conceptualization phase of theory building to identify and to present propositions between constructs to explain the causalities that govern safety at work. By following a general method of theory building in applied sciences, and consistent with the principle of inherent simplicity from TOC, our findings indicate the existence of four constructs that govern safety at work: knowledge, planning, behavior, and performance measurement.

Moreover, each construct and its interconnections comprised a set of propositions expressed through a conceptual framework that explains the underlying issues in safety at work and put behavior as a key element. Furthermore, as a result of our analysis based on the stages in which every science has gone through (classification, correlation, and causal consistency), the phenomenon of safety at work was represented as a system in which the level of complexity depends on the interdependencies between constructs and harmony.

A major theoretical outcome from this research is a conceptualization narrative that defines the complexity of safety at work as a consequence of degrees of freedom (interdependencies between constructs) and harmony (absence of conflicts between constructs). We postulate that as much interdependent and harmonious is the system the less complex is safety at work. In that reasoning, both circumstances affect safety at work and determine whether safety at work is a complex or exceedingly simple matter. Although foster future research is highly encouraged to cover other phases of this theoretical model, this study presents generalizability regarding temporal and contextual factors discussed.

Finally, from the practitioner's perspective, our findings contribute to the improvement of safety practices at the organizational level by redefining their structures, connections and focusing on behavior-based safety under a broader perspective.

APPENDICES CHAPTER 3

A 1. Structure of the investigation report

Category	Data field (required information)
Time-horizon (n=3)	Fiscal year Month Sequence
Location (n=5)	Business Unit Operation Unit Country Branch Geographic region
Individual (n=6)	Age Scholar level Technical background Job function Years of experience Years working for the company
Accident data (n=9)	Type of accident, e.g., Elevator. Equipment Lost days Level of severity Body's part affected Nature of illness/injury Weekday Shift Location where the accident occurred
Process planning (n=7)	Task condition, e.g., routine, non-routine Job site (OTD status), e.g., on-time, delayed Worked hours in the circumstances of the event PPE: Was appropriate PPE being used? (Y/N) Tools: Were there appropriate tools available? (Y/N) JHA: Was it performed (Y/N) JHA: Was it performed according to the task? (Y/N)
Previous accidents/santions/audits (n=3)	Previous accident reported? (Y/N) Previous sanctions in the last 12 months? (Y/N) Audited in the last 12 months? (Y/N)
Training (n=3)	Hours of training (last 12 months) 10 rules training up to date? (Y/N) Has been trained for the task being performed (Y/N)

Behavior	Behavioral assessment in the last 12 months? (Y/N) Behavioral change observed recently? Psychological test performed during onboarding?
Violated rules	Technical rule violated? e.g., PPE usage, fall protection etc. Behavioral trap associated with the accident? e.g., Diminishing risks, lack of concentration etc.

A 2. Script R for association rules

```
RStudio v. 4.0.5

# Require packages
if(!require(readxl)) install.packages("readxl")
if(!require(arules)) install.packages("arules")
if(!require(arulesViz)) install.packages("arulesViz")
if(!require(tidyr)) install.packages("tidyr")

# Load packages
library(readxl); library(arules), library(arulesViz), library(tidyr)

# Load dataset
data <- read_excel("Lost-time accidents Report.xlsx", sheet='DATA')
View(data)

# Adjust dataset
data_aj <- dados [, c(-2,-3,-4,-5,-6,-7,-8,-9)]
View(data_aj)

# Convert dataset into file .csv
write.csv(dados_aj,"AR.csv", quote=FALSE, row.names=FALSE)

# Convert dataset into transaction format
tr <- read.transactions('AR.csv', format = 'basket', sep=',')
tr
summary(tr)

# Create association rules
rules = apriori(tr, parameter=list(suppor = 0.5, conf = 0.8, minlen = 1, maxlen = 3))
rules
inspect(head(rules))

# Remove redundant rules
rules = rules[!is.redundant(rules)]
rules
inspect(rules)
result = inspect(rules)

# Print association rules
write.csv2(result, "Association rules.csv")
```

4 PERFORMANCE MEASUREMENT IN OCCUPATIONAL HEALTH AND SAFETY MANAGEMENT SYSTEM: AN ANALYSIS OF CRITICAL ELEMENTS COMPRISING THE ASSESSMENT INSTRUMENTS[5]

Abstract

Occupational Health and Safety Management Systems (OHSMS) establish mechanisms to deploy strategies to control workplace risks and compliance with regulations. However, despite greater attention given to safety management in the past years, the burden concerning work-related accidents remains significant. Thus, an appropriate evaluation of safety performance arises as a fundamental instrument for supporting organizations in this context. This study aims at identifying the existing OHSMS assessment instruments and their critical elements. Moreover, we investigate how these elements are interconnected, as well as critically analyze the conceptual foundations of the literature examined. As a result, 13 assessment instruments and 39 critical elements were identified through thematic analysis. Additionally, the critical elements were presented in a form of a mixed coding scheme and had their association rules revealed from data-driven structural analysis. Finally, a critical interpretative synthesis was undertaken discussing the conceptual foundations. From the managerial perspective, this research exploits the critical elements that constitute OHSMS assessment instruments, drawing the attention of practitioners to the relevance of using rigorous performance measures concepts to avoid misinterpretations of the results. From the theoretical perspective, new conceptual directions are given for further studies of OHSMS evaluation, based on recognized literature grounded outside of safety science.

Keywords: OHSMS. Occupational Health and Safety Management System. Assessment instrument.

4.1 Introduction

Occupational health and safety (OHS) has become a global issue, and solutions to enhance its performance have been urgently required in modern industrial environments (Wang et al., 2020). According to International Labour Organization (ILO) (2020), more than 2.8 million deaths per year result from occupational accidents or work-related diseases. When considering non-fatal work-related injuries, this number increases to approximately 376.8 million a year. This impact is associated with poor occupational health and safety management (OHSM) and accounts for economic losses estimated at 3.94% of the global gross domestic product (Brocal et al., 2018; ILO, 2020; Wang et al., 2020).

[5] Article submitted to Journal of Occupational Safety and Ergonomics. Submission ID # 222706089.

The research on OHSM grew rapidly since 1990 (Wang et al., 2020), and throughout this decade, international standardization organizations published the first guidelines and standards with concepts related to occupational health and safety management systems (OHSMS) (BSI, 1999, 1996). In the same period, the subject gained academic notoriety with the seminal works conducted at the University of Michigan (Dyjack et al., 1998; Dyjack and Levine, 1996b; Redinger and Levine, 1998).

Despite the growing importance of OHSMS evidenced by standards and scientific publications, work-related injuries remain a challenge. In 2018, aiming at helping organizations to improve employee safety, reduce workplace risks, and create better working conditions (Brocal et al., 2018), the standard ISO 45001 was released (ISO, 2018). ISO 45001 suggests that OHSMS should be based on the PDCA[6] reasoning. Nevertheless, it does not deepen the concept of performance assessment (check), being limited to superficial implementation guidance.

The evaluation of OHSMS is as important as other issues in organizational management, such as production or service delivery rate (Arezes and Sérgio Miguel, 2003). It represents a critical process to drive consistent improvements in the field of safety at work. When compared to other OHSMS elements, like initiatives or processes, assessment has received less attention from scholars and practitioners (Mohammadfam et al., 2017). As a result, a wide range of misunderstood concepts about OHSMS performance measures has spread in the literature (Bianchini et al., 2017). The undesirable effect is that a lack of clarity about the meaning of these measures may result in wasted investments, increased costs, and negligible reductions in work-related injuries.

In this circumstance, the assessment instruments play a fundamental role in ensuring that the evaluation of OHSMS is consistent with the performance measure concept. For this reason, they must be clear as to their context of use, the elements that constitute them, as well as their measuring object. For example, instruments to measure efficacy (e.g., the percentage of achieving a target accident rate) consider the number of accidents and the number of hours worked (OSHA, 2012). However, when viewed from the perspective of efficiency, these instruments should also consider the resources employed to achieve the desired performance (Piran et al., 2020). From a managerial standpoint, this comprehensive understanding is relevant for two

[6] Plan-do-check-act cycle (Deming, 1986).

reasons: (i) to clarify the meaning of results; and (ii) to rise assertiveness in decision-making.

Given the aforementioned, this study aims at identifying the OHSMS assessment instruments and the critical elements that compose them. Moreover, it seeks to investigate how these elements are interconnected, as well as to critically analyze the conceptual foundations used in the assessment instruments examined.

To reach these objectives, a systematic literature review of 26 peer-reviewed articles (1998-2020) was performed following a robust research design. As a result, 13 assessment instruments and 39 critical elements were identified through a thematic analysis. Additionally, the critical elements were classified into 5 categories (i.e. context, performance measure, inputs, mechanisms, and outputs) and had their association rules revealed from data-driven structural analysis. Finally, a critical interpretative synthesis discussing the conceptual foundations underlying this research pointed out the need of positioning each piece of research in an appropriate disciplinary context, which means using concepts consistent with literature outside of safety science (Rae et al., 2020).

This work comes up with contributions in the field of occupational health and safety. From the managerial perspective, it puts light on the critical elements that constitute OHSMS assessment instruments, drawing the attention of practitioners to the relevance of using rigorous performance measures concepts to avoid misinterpretations of the results. Also, this study focused on unhiding how the elements that constitute the OHSMS are connected, offering a comprehensive view of through what mechanisms the initiatives generate the outcomes, and in which context. Finally, from the theoretical perspective, directions are given for further studies of OHSMS evaluation, based on recognized literature grounded in operations management.

Besides this section, this paper is structured as follows. Section 4.2 outlines the theoretical background. The research approach and methodological procedures are presented in section 4.3. Section 4.4 highlights the results of the content analysis. Section 4.5 critically analyses the research findings. Finally, section 4.6 draws the conclusions, contributions, and limitations along with its future research directions.

4.2 Theoretical background

The positive impact of introducing OHSMS at the organizational level has been recognized by the public and private sectors (ILO, 2001). Such increased attention was initially attributed to the activities of the International Organization for Standardization (ISO), notably the publications of the quality assurance model (ISO, 1994) and the environmental management system model (ISO, 1996). Nevertheless, OHSMS has its origins in the early twentieth century, based on the tools and techniques of process statistical control as reasoned by Redinger et al. (2011) and depicted in Figure 23.

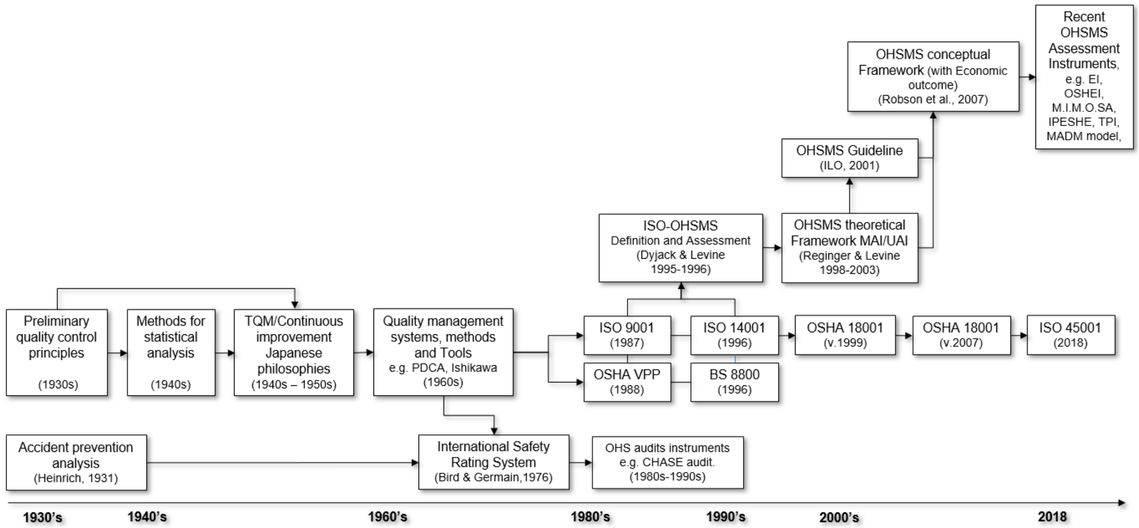


Figure 23 – Assessment of OHSMS: origins and timeline.

These techniques got wide acceptance after their integration with the Japanese improvement philosophies. This integration resulted in the work of Bird and Germain (1976), which was considered the earliest assessment instrument focused on accident causation analysis. In like manner, Dyjack and Levine at the University of Michigan (Dyjack and Levine, 1995) pioneered in developing an ISO standard intended to set the foundation for OHSMS assessment. From this point in time, and boosted by the works of the Occupational Safety and Health Administration (OSHA), more attention to safety management systems was given in both academic and practical fields.

Despite its recognized importance, a lack of consensus about the definition and scope of OHSMS emerged (Robson et al., 2007). For example, Dyjack and Levine (Dyjack and Levine, 1996b) defined OHSMS as an orderly arrangement of interdependent activities and related procedures driving the organization's occupational health performance. ILO, in turn, refers to OHSMS as a set of interrelated or interacting elements to establish OHS policy and achieve its objectives. In an attempt to clarify the definition and scope, Redinger and Levine (Redinger and Levine, 1998) introduced basic constructs to describe an OHSMS with its elements and boundaries, and to propose a universal assessment instrument.

Additionally, other frameworks found elsewhere (Arezes and Sérgio Miguel, 2003; Ramli et al., 2011; Redinger et al., 2011; Robson et al., 2007) converge to the fact that, like any other system, the OHSMS is composed of inputs, mechanisms, and outputs. However, not much attention is paid to the evaluation piece, and at least two questions frequently remain open concerning OHSMS: (i) has the organization been effective? (ii) has the organization been efficient considering the resources used to generate the outcomes?

This gap is related to a lack of clarity on which performance measure is used in the assessment instrument since different perspectives might take place. On one hand, it can be focused on efficacy (e.g., reducing accidents at a certain level). In this case, the assessment instrument is target-driven and does not consider the resources used. On the other hand, the assessment could be focused on efficiency, which is a relative measure in which it is sought the optimization of the relation between inputs and outputs.

By following any of these perspectives, the assessment instruments represent a critical tool to support safety practitioners in their daily challenges. Moreover, depending on the criteria established by the performance measure, the analysis of results follows distinct approaches. For example, an organization performing at a 95% efficacy level means it roughly reaches its target. Yet, this result is insufficient to guide decision-makers in what should be done to close such a 5% gap. In this case, benchmarks are rarely possible to be identified because the resources used to achieve this level of efficacy are out of scope. In turn, an organization performing at a 95% efficiency level means there are opportunities for optimizing the relationship between the inputs and outputs of the system. Also, it is possible to identify internal and external benchmarks, and directions on how to improve efficiency are given.

Despite these concepts being widely known outside of safety science (Coelli et al., 2005), not much attention has been paid by researchers and practitioners to the use of adequate performance measurement concepts in the field of OHSMS evaluation. As a result, conceptual divergences are frequently observed in the assessment instruments found in the literature, compromising the analysis of results. Also, as a side effect, the ability of OHSMS and its evaluation process to become a lever for safety improvements have already been questioned in numerous papers (Skład, 2019).

To combat this problem and based on the reasoning of positioning each piece of research in an appropriate disciplinary context (Rae et al., 2020), the assessment instruments should take into consideration the concepts synthesized in specialized literature (Piran et al., 2020) (see Figure 24). This is necessary because as it occurs in the field of operations management, clarity about the meaning of the results is critical for decision-makers, and thus for improving safety performance.

These measurements are well-defined in the specialized literature and they are different in concept (Coelli et al., 2005). Yet, very often they are erroneously considered equals. Productivity, for instance, is the ratio between outputs and inputs, corresponding to a performance indicator that organizations wish to maximize. Efficacy relates to the achievement of a defined objective, without considering the resources used. Efficiency, in turn, is a relative measure of performance. A system is considered 100% efficient if performance observed in others does not show that is possible to improve some of the inputs or outputs without worsening the other input and outputs (William W. Cooper et al., 2011a). Finally, effectiveness represents the capacity of a system to achieve the objectives under the most optimized input/output relation.

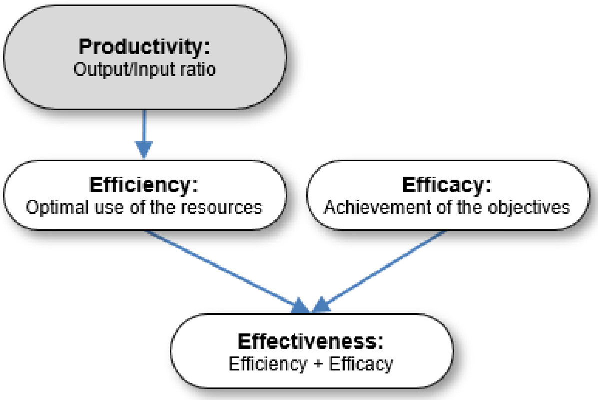


Figure 24 – Synthesis of performance measurement concepts (Piran et al., 2020).

A misunderstanding of this body of knowledge has led practitioners to take insufficient and costly decisions in the field of operations management. From the academic perspective, such a lack of conceptual consistency has implications for literature cohesion, resulting in different directions in the same disciplinary context.

4.3 Research design

This study has the evaluation of OHSMS as the core subject. The research design follows the Literature Grounded Theory as proposed by Ermel et al. (2021), as depicted in Figure 25. It aims at identifying the existing assessment instruments and their critical elements. Also, it seeks to investigate how these elements are interconnected, as well as critically analyze the conceptual foundations of the assessment instruments examined.

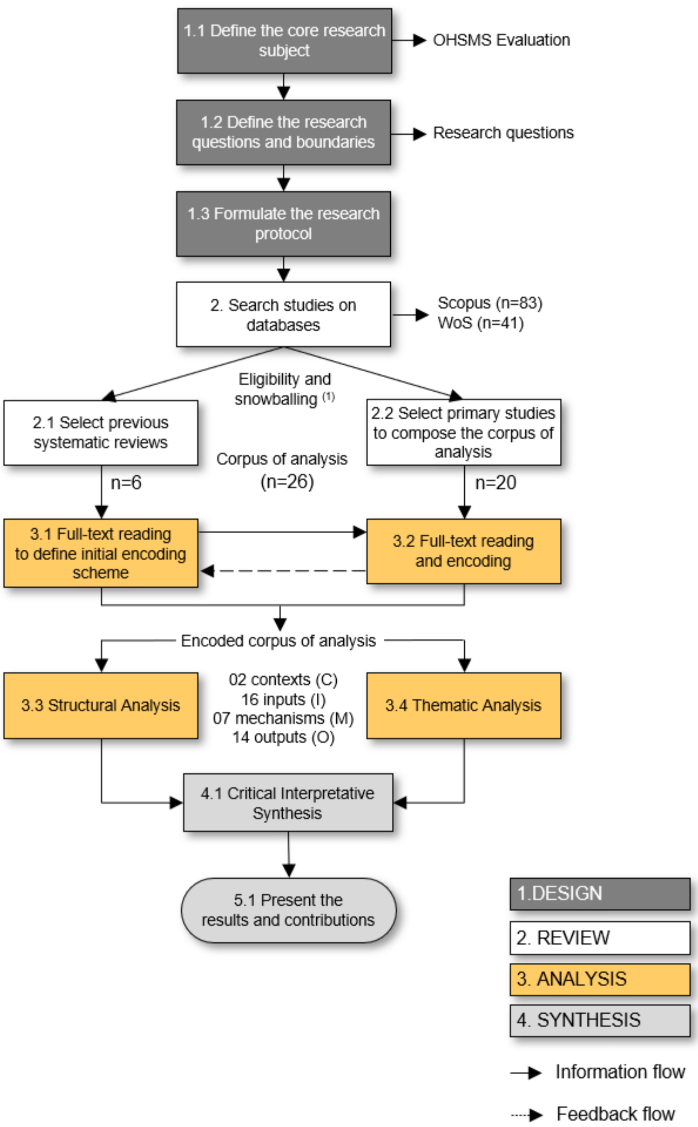
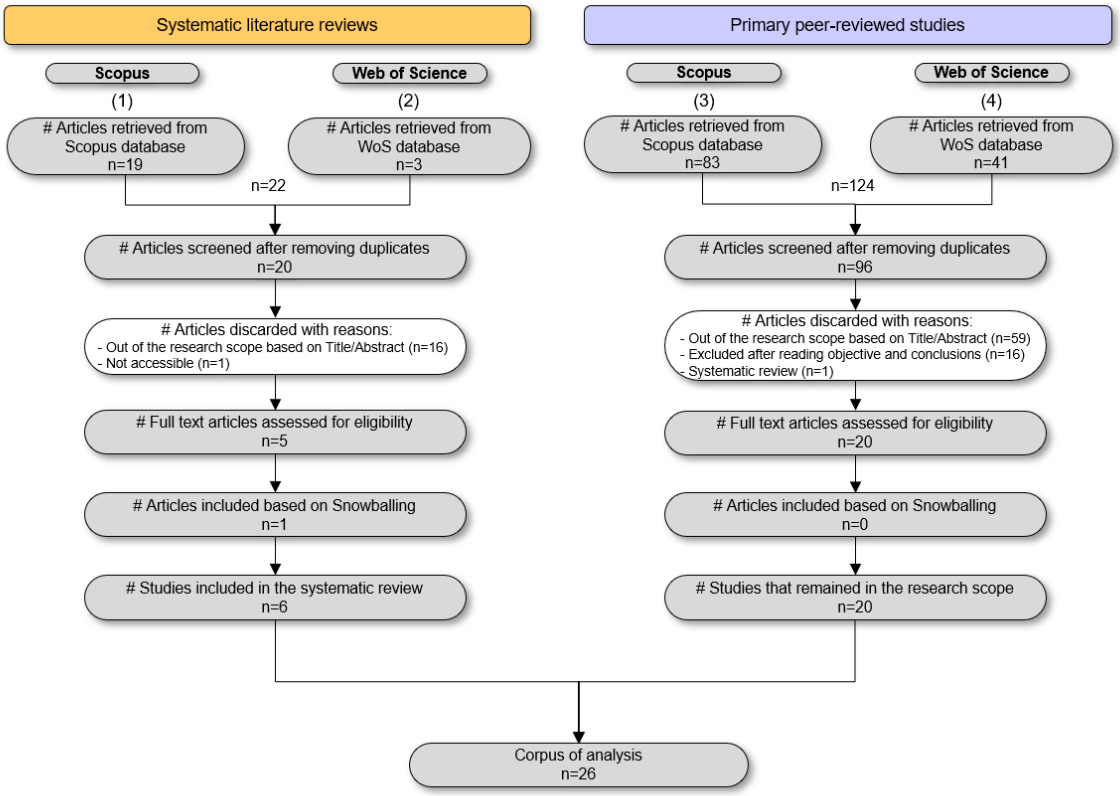


Figure 25 – Research design for diagram

The process started by defining the OHSMS evaluation as the core research subject. Then, in step 1.2 the research questions were defined based on a narrower field under interest: ($RQ1$) which instruments have been used to assess the OHSMS performance? ($RQ2$) what are their critical elements ($RQ3$) how these critical elements are interconnected? Then, the protocol for systematic literature review was developed based on Ermel et al. (2021) (see appendix A 3).

The selection of eligible studies followed methodological steps 2.1 and 2.2 (Figure 26). The search was conducted in Scopus and Web of Science databases, wherein only peer-review articles were consulted. The eligibility was split into two axes: on the one side, systematic literature reviews were searched as a basis for the theoretical framework, and for the elaboration of the a priori encoding scheme to be used in the content analysis. On the other side, peer-reviewed studies were retrieved in compliance with the research scope.



(1) Search string: TITLE-ABS-KEY ("safety management system" AND "systematic review")
(2) Search string: TOPIC ("safety management system" AND ("systematic review")
(3) Search string: TITLE-ABS-KEY (("occupational health and safety management system" OR "OHSMS") AND ("performance" OR "efficiency" OR "efficacy" OR "productivity" OR "effectiveness")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))
(4) Search string: TOPIC: (("occupational health and safety management system" OR "OHSMS") AND ("performance" OR "efficiency" OR "efficacy" OR "productivity" OR "effectiveness")). Refined by: DOCUMENT TYPES: (ARTICLE)

Figure 26 – Flowchart of searching, eligibility, and snowballing to compose the corpus of analysis.

Limited to the article title, abstract, and keywords, the search resulted in 22 systematic literature reviews and 124 articles (see Figure 26). The duplicates were discarded, and titles and abstracts of the remaining studies were inspected to discard articles either identified as out of scope or not accessible online, as recommended elsewhere (Brunton et al., 2012). Then, 5 systematic literature reviews and 20 empirical studies were analyzed in-depth and considered in the snowballing search. The snowballing was conducted backward and forward as reasoned by Wholin (2014). In the backward search, twelve candidates were considered for inclusion based on the inspection of titles. However, after an in-depth analysis, only one article followed the eligibility criteria and was included in the scope. The same procedure was carried out in the forward search but no additional study was included. As a result, 6 systematic literature reviews (SLR_i) and 20 primary studies (R_i) composed the corpus of analysis as presented in Table 15 and Table 16.

Table 15 – List of systematic literature reviews composing the corpus of analysis

Id	Title	Source	Authors (Year)
SLR_1	The effectiveness of occupational health and safety management system interventions: A systematic review	Safety Science	Robson et al. (2007)
SLR_2	Standardized risk assessment techniques: A review in the framework of occupational safety	Safety and Reliability	Brocal et al. (2018)
SLR_3	Critical factors of success and barriers to the implementation of occupational health and safety management systems: A systematic review of literature	Safety Science	Da Silva and Amaral (2019)
SLR_4	Making occupational health and safety management systems 'work': A realist review of the OHSAS 18001 standard	Safety Science	Uhrenholdt Madsen et al. (2020)
SLR_5	Examining Organizational, Cultural, and Individual-Level Factors Related to Workplace Safety and Health: A Systematic Review and Metric Analysis	Health Communication	Lee et al. (2020)
SLR_6	A systematic review on the research progress and evolving trends of occupational health and safety management: A bibliometric analysis of mapping knowledge domains	Frontiers in Public Health	Wang et al. (2020)

Table 16 – List of primary studies composing the corpus of analysis

Id	Title	Source	Authors (Year)
R_1	Development and evaluation of the Michigan Occupational Health and Safety Management System Assessment Instrument: A universal OHSMS performance measurement tool	American Ind. Hygiene Association Journal	Redinger and Levine (1998)
R_2	Evaluation of an occupational health and safety management system performance measurement tool—iii: Measurement of initiation elements	American Ind. Hygiene Association Journal	Redinger et al. (2002)
R_3	The role of safety culture in safety performance measurement	Measuring Business Excellence	Arezes and Miguel (2003)
R_4	An integrated approach for improving occupational health and safety management: The Voluntary Protection Program in Taiwan	Journal of Occupational Health	Su et al. (2005)
R_5	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	Chen et al. (2009)
R_6	An empirical analysis of the effectiveness of occupational health and safety management systems in SMEs	International Small Business Journal	Arocena and Núñez (2010)
R_7	Making work environment auditable - A “critical case” study of certified occupational health and safety management systems in Denmark	Safety Science	Hohnen and Hasle (2011)
R_8	Possibilistic regression analysis of influential factors for occupational health and safety management systems	Safety Science	Ramli et al. (2011)
R_9	Occupational health and safety management systems	Patty’s Industrial Hygiene (Book chapter)	Redinger et al. (2011)
R_{10}	Effect of occupational health and safety management system on work-related accident rate and differences of occupational health and safety management system awareness between managers in South Korea’s construction industry	Safety and Health at Work	Yoon et al. (2013)
R_{11}	The relationship between the implementation of voluntary Five-Star occupational health and	Safety Science	Hedlund (2014)

Id	Title	Source	Authors (Year)
	safety management system and the incidence of fatal and permanently disabling injury		
R_{12}	Framework for continuous assessment and improvement of occupational health and safety issues in construction companies	Safety and Health at Work	Mahmoudi et al. (2014)
R_{13}	Quantitative assessment of occupational safety and health: Application of a general methodology to an Italian multi-utility company	Safety Science	Saracino et al. (2015)
R_{14}	Measuring operational performance of OSH management system - A demonstration of AHP-based selection of leading key performance indicators.	Safety Science	Podgórski (2015)
R_{15}	An innovative methodology for measuring the effective implementation of an Occupational Health and Safety Management System in the European Union	Safety Science	Bianchini et al. (2017)
R_{16}	An assessment of occupational health and safety measures and performance of SMEs: An empirical investigation	Safety Science	Gopang et al. (2017)
R_{17}	Evaluation of the Quality of Occupational Health and Safety Management Systems Based on Key Performance Indicators in Certified Organizations	Safety and Health at Work	Mohammadfam et al. (2017)
R_{18}	A multiple attribute decision model to compare the firms' occupational health and safety management perspectives	Safety Science	İnan and Yılmaz (2017)
R_{19}	A cross-sectional study of factors influencing occupational health and safety management practices in companies	Safety Science	Nordlöf et al. (2017)
R_{20}	Assessing the impact of processes on the Occupational Safety and Health Management System's effectiveness using the fuzzy cognitive maps approach	Safety Science	Skład (2019)

Steps 3.1 and 3.2 consisted of encoding the corpus of analysis, which was undertaken by iteratively defining and aggregating codes into categories, and then assigning codes and categories to the full texts. It adopted a mixed coding scheme composed of categorical and open codes, as well as a priori and a posteriori categories (Dresch et al., 2015a).

The categorical codes and a priori categories were defined based on the full-text reading of the selected systematic literature reviews and followed the CIMO logic (Denyer et al., 2008). It represents a combination of a problematic Context, for which the design proposition suggests a certain Intervention type, to produce, through specified generative Mechanisms, the intended Outcomes, and definitions are given in Table 17

Table 17 – CIMO-Logic used to define a priori categories.

Component	Scope
C – Context	Contextual layers in which the research is interested.
I – Intervention	Initiatives and resources used on the influence factors (inputs) to achieve the expected outcomes.
M – Mechanisms	Combination of the resources and initiatives offered in the intervention and the reasoning that leads actors to change their behavior when confronted with these resources.
O – Outcomes	Results identified (outputs)

Following this reasoning, the encoding scheme was structured to capture from the examined literature the following aspects: the organizational contexts under the research's interest, the performance measures used in the OHSMS assessment instruments, initiatives employed, the mechanisms used to bring about a change in workplaces, and outcomes achieved.

The open codes and a posteriori categories, in turn, emerged during the reading and in-depth analysis of the primary studies, as reasoned by Strauss and Corbin (Strauss and Corbin, 1990). The main aspects and findings of each study are summarized in appendix (see A 4, A 5). Concerning the rationale used for grouping codes into categories, this research adopted the thematic criterion and followed the principles of mutual exclusion, homogeneity, pertinence, objectivity, and productivity, as posed by Bardin (1993).

The outcome from steps 3.1 and 3.2 was the mixed coding scheme of the corpus of analysis (Table 18). composed of 5 a priori categories, 6 a posteriori sub-categories, 6 categorical codes, and 36 open codes, as further detailed in the next session.

Table 18. Mixed encoding scheme

Structure	Elements
A priori Categories (n=5)	Context, Performance Measure, Mechanism, Input, Output

A posteriori Sub-Categories (n=6)	Sub-Input (3), Sub-Output (3)
Categorical codes (n=6)	Types of Context (2), Types of Performance Measures (4)
Open codes (n=36)	Mechanisms (7), All inputs (16), All outputs (14)

In step 3.3 a thematic analysis was conducted. It consisted of calculating and comparing the frequencies of elements under interest in the research (codes) previously grouped into significant categories (A. Bardin, 1993). It follows the reasoning that the more frequently cited the more important the code/category is. Also, the relationship between the frequencies of codes and categories follows the reasoning that if an item set is frequent, then all of its subsets must also be frequent (Tan et al., 2019), and details are found elsewhere (Ermel et al., 2021). As a result, both absolute and relative frequencies are presented in an occurrence matrix (see appendices A 6, A 7, and A 8).

Also, seeking to assess how the OHSMS elements are arranged and to reveal underlying aspects of those relationships, a structural analysis was conducted in step 3.4. This type of analysis assumes that the co-occurrence of two or more codes/categories, in the same context unit, provides information about mental and ideological structures as well as latent concerns (A. Bardin, 1993).

To do that, based on the occurrence matrix, *Apriori* was used to identify critical relationships between variables, i.e., how an initiative or a set of initiatives (antecedent factors, named lhs) impact a particular outcome (consequent factor, named rhs). It consists of a data mining algorithm that systematically controls the exponential growth of candidate itemsets (Zhang and Zhang, 2002). For this analysis, the parameters support (supp=0.05), and confidence (conf=0.5) were set up as thresholds in R (R core team, 2018). The support determines how often a rule applies to a given database. Besides, it aims to identify the most relevant rules in the database (Alves, 2020; Gauss et al., 2020). Confidence, in turn, determines how frequently consequent factors [rhs] appear in relationships that contain antecedent ones [lhs]. It is used to measure the strength of an association rule, expressed as the times a specific itemset is found together with a specific item out of the total times this specific itemset is found in the entire database (Kouzis-loukas, 2014). The thresholds were defined to expand data mining with minimal restrictions and based on adopted criteria from previous studies (Isa et al., 2018; Kouzis-loukas, 2014).

Finally, by offering an argumentative narrative in step 4.1, a critical interpretative synthesis was conducted to question the nature of assumptions on which the existing literature has influenced the way OHSMS has been assessed. The results of this process are highlighted (step 5.1) and discussed in the next sections.

4.4 Results

The results of this study are organized into two subsections. Subsection 4.4.1 introduces the assessment instruments identified in the literature and outlines their critical elements. Based on thematic and structural analyses, the prevalence and the patterns of co-occurrence of these elements are shown in subsection 4.4.2.

4.4.1 OHSMS assessment instruments and their critical elements

Assessment instruments might be structured in a form of a method, modeling, or any other way of performance checking. Moreover, apart from their intrinsic elements (context, input, mechanism, and output) another critical factor is the performance measurement (e.g., efficiency). It comes from outside of safety science and defines the necessary elements, types of data, and techniques in which the analysis of results is conducted.

In total, 13 OHSMS assessment instruments were identified in the primary studies examined (see Table 19). On the quantitative side, four performance indexes were found. Ramli et al. (Ramli et al., 2011) proposed the *OHS Efficiency Implementation* (OHSEI) based on the ranked order of influential factors. Mahmoudi et al. (Mahmoudi et al., 2014) developed the *Total Performance Index* (TPi) to classify different construction companies at levels of performance. Saracino et al. (Saracino et al., 2015) proposed the *Index of PErformance for Safety and Health* (IPESHE) based on the *Methodology for the Implementation and Monitoring of Occupational Safety* (M.I.M.O.SA.). Bianchini et al. (Bianchini et al., 2017) introduced the *Efficacy Index* (EI) by adding economic factors. A commonplace of these indexes is the need of ranking organizations based on defined criteria, such aforementioned TPi.

From the modeling perspective, İnan et al. (2017) developed a multiple attribute decision-making model (MADM) for determining and comparing the firm's OHSMS performance in defined aspects. These quantitative studies comprise a significant body of knowledge in the field of OHSMS evaluation, even though the realistic application might be challenging for practitioners due to the lack of resources and reliable data.

Table 19 – List OHSMS assessment instruments examined in the literature

Id	Assessment instrument	Performance measure considered by authors	Type of approach
R_1, R_2, R_9	MAI/UAI	Effectiveness	Criteria rating/scoring
R_6	Statistics and Clustering		Statistics
R_{20}	FCM (Fuzzy Cognitive Maps)		Simulation
R_4	Statistics	Efficacy	% of the target
R_{15}	EI		Quantitative Index
R_8	OHSEI	Efficiency	Quantitative Index
R_{13}	M.I.M.O.SA./IPESHE		Quantitative Index
R_{11}	5-star system	'OHSMS performance'	Criteria rating/scoring
R_{12}	TPi		Quantitative ranked Index
R_{16}, R_{17}, R_{19}	Statistics and/or Regression		Criteria rating/scoring
R_{18}	MADM model		Modeling for Criteria rating/scoring
$R_3, R_5, R_7, R_{10}, R_{14}$	* Studies do not present any OHSMS assessment instrument.		

Comprising the qualitative studies, Redinger et al. (Redinger et al., 2002a; Redinger and Levine, 1998) pioneered when developed the *Michigan Assessment Instrument* (MAI), lately called *Universal Assessment Instrument* (UAI), a seminal proposal for OHSMS evaluation. Hedlund (Hedlund, 2014) introduced the *5-Star System* to rank South African companies based on OHS requirements by defining criteria and categories concerning the organization's safety level. On one side, the assessment instruments based on rating/scoring of critical factors seem to be more feasible for the use of safety practitioners, including in small and mid-size organizations. On the other side, the lack of quantitative measurement techniques might compromise the evaluation in terms of the economic dimension, or be too superficial in identifying internal or external benchmarks.

Overall the examined literature has neglected the use of conceptual rigor concerning performance measurement. Six out of twenty articles lack a precise definition of what the generic term 'OHSMS performance' means when confronted with well-defined literature elsewhere (Coelli et al., 2005). This is inadequate because it does not refer to any specific pattern of measurement. As a result, an avenue of vague conclusions is brought to management reviews due to the lack of clarity about what has been measured, and in which context. With regards to measurement concepts derived from productivity, effectiveness was the terminology more frequently cited in the corpus of analysis (5 out of 20), followed by efficacy and efficiency. Nevertheless, five studies do not even refer to an OHSMS assessment instrument.

The use of an appropriate conceptual base is critical for OHSMS evaluation. This is because different measures require distinguished elements for the assessment. For example, efficacy is adequate for tracking targeted outcomes but not sufficient to identify internal benchmarks. In turn, efficiency is a relative measure that considers the impact of the resources used to generate the results, therefore indicated for benchmark analysis. In this reasoning, a high level of efficacy is not equal to efficiency. This comprehension is fundamental for management reviews, as organizations should be able to understand in-depth the reasoning composing the OHSMS assessment.

In this context, critical elements were identified and structured in a form of a mixed coding scheme (Table 20). It resulted from the in-depth analysis of the examined literature and represents a practical guide for developing and analyzing assessment instruments. In total, 39 elements were mapped as critical in OHSMS assessment instruments: 2 contexts, 16 inputs, 7 mechanisms, and 14 outputs. The inputs and outputs were sub-classified based on their characteristics, for example, inputs were categorized into voluntary, mandatory, or hybrid. The outputs, in turn, were classified into intermediate, final, and economic following previous conceptual fr ameworks (Redinger and Levine, 1998; Robson et al., 2007).

Table 20 – A coding scheme for the content analysis (based on Gauss et al., 2020)

Category/sub-category (a priori)	CIMO-Logic components	Sub-Code
Contextual layer (CL_i)	C	(CL_1) Certified organization (CL_2) Non-certified organization
Conceptual performance measure (PM_i)	C-I-M-O	(PM_1) Productivity (PM_2) Efficiency (PM_3) Efficacy (PM_4) Effectiveness
Mechanisms employed (M_i)	M	(M_1) Integration (M_2) Learning (M_3) Awareness (M_4) Motivation (M_5) Leadership commitment (M_6) Error-proof (Poka-Yoke) (M_6) Standardization
Inputs (I_{ij}) (I_{1j})Voluntary (I_{2j}) Mandatory (I_{3j}) Hybrid	I	(I_{11}) OHS Policy (I_{12}) Management participation (I_{13}) Employees participation (I_{14}) Hazard/Risk assessment (I_{15}) Incident investigation and root-cause analysis (I_{16}) Proced., routines, and resources (PPE, tools)

Category/sub-category (a priori)	CIMO-Logic components	Sub-Code
		(I_{17}) Fire protection and emergency preparedness (I_{18}) Accountability (I_{19}) Auditing and self-inspection (I_{1a}) System to support performance management (I_{1b}) Behavioral-based initiatives (I_{1c}) System to support performance management (I_{21}) Regulatory compliance and system conformity (I_{31}) Training and personnel development (I_{32}) Safeguarding process design (I_{33}) Occupational health services
Outputs (O_i) (O_{1j}) Intermediate (O_{2j}) Final (O_{3j}) Economic	O	(O_{11}) Better safety climate (O_{12}) Increased safety awareness (O_{13}) Increased hazard reporting (O_{21}) Reduction of #accidents or occup. Illness (O_{22}) Reduction on the hazard exposure (O_{23}) Reduction on lost-time injury rates (O_{24}) Completion rate of corrective and preventive measures (O_{31}) Disability-related cost reduction (O_{32}) Fines/sanctions cost reduction (O_{33}) Management cost reduction (O_{34}) Increase of economic KPIs (e.g., Sales, GM) (O_{35}) Economic results (i.e. Sales, GM)

In the field of OHSMS evaluation, the inputs are the resources (i.e. initiatives) that organizations use to generate results. The outputs, in turn, are the outcomes organizations want to achieve in three possible dimensions: intermediate (e.g. safety climate), final (e.g. number of accidents), and economic.

Each dimension plays an important role in safety management. For example, the reduction of accidents is a final outcome targeted by any organization. This dimension is relevant because it objectively demonstrates a focus on prevention. Intermediate outcomes (e.g., safety awareness) are important to promote engagement in safety while the economic ones reveal the impact of OHSMS on the organization's overall result.

The connection between the initiatives and outcomes occurs through the mechanisms. They are the reasoning that explains the change in the workplace's practices when confronted with a set of initiatives. In other words, the mechanisms explain how initiatives are expected to generate results in a particular context.

In the mixed coding scheme presented in Table 20, both certified and non-certified organizations were considered as contextual layers based on previous systematic reviews. Besides, these configurations represent common sense concerning the contexts of OHSMS. On one side, certified organizations are more related to larger companies based on voluntary initiatives or due to market requirements, such as ISO certification (2018). On the other side, the majority of small and mid-size companies do not follow this context. This is consistent with the results identified in the examined literature, further presented in the next subsection based on the thematic analysis.

Concerning inputs and outputs, the majority of primary studies encompass these critical elements in their OHSMS assessment instruments. However, the relationships among inputs, mechanisms, and outputs by using a realist review were outlined only in the study of Uhrenholdt Madsen et al. (Uhrenholdt Madsen et al., 2020). In general terms, the literature lacks studies focusing on the connections among these critical elements. Moreover, none of those studies employed techniques to unhide underlying issues, such as to investigate how a set of initiatives (input) is associated with a particular result (output), or which mechanism explains a particular relation input-output.

Finally, the results concerning the relative frequencies and the patterns of co-occurrence of the critical elements composing OSHMS assessment instruments are outlined in the next subsection.

4.4.2 Thematic and structural analysis

The most frequent contexts, inputs, mechanisms, and outputs were identified based on the counting principles and the occurrence/frequency matrixes as a result of the thematic analysis. A summary of the top-3 findings is given in Table 21 and additional data are shown in appendix (A 6, A 7, and A 8).

Table 21 – Most frequent critical elements based on thematic analysis

CIMO-Logic	TOP 3 listed codes	Relative Frequency
Context	(CL_2) Non-certified organizations	50%
	---- Indefinite	30%
	(CL_1) Certified organization	20%
Input	(I_{11}) OHS policy	55%
	(I_{12}) Management engagement	55%
	(I_{16}) Proc., routines and resources	55%
Mechanism	(M_5) Engagement	30%

	(M_3) Awareness ($M_{1,2}$) Integration, Learning	20% 15% (each)
Output	(O_{21}) Reduction of the number of accidents (O_{23}) Reduction of lost-time injuries (O_{34}) Increase of economic KPI	50% 20% 20%

Non-certified organizations were prevalent as a contextual layer in the revised literature (50%). For some primary studies, the authors did not make clear in which context their studies were conducted, i.e., whether the assessment instrument applies to both contexts or a particular one, and what are its limitations. This lack of clarity raises a critical problem for OHSMS assessment since the comprehension of the context is a fundamental matter to a realist evaluation based on CIMO-Logic. Also, it represents the boundary in which the assessment makes sense, and should draw the attention of researchers and practitioners.

Concerning the inputs, on the one side results indicate that OHS policy ($I11$), management engagement ($I12$), and procedures, routines, and resources ($I16$) are critical inputs. On the other hand, behavioral-based initiatives ($I1$b) and compliance and regulatory conformity ($I21$) were less mentioned. The analysis from the primary studies also indicates that inputs are still very traditional and not much attention has been paid to topics such as behavior and mental health. This lack of attention was observed in several selected studies in this research and it represents a necessary and vast research field for scholars and practitioners.

With regards to the mechanisms, only 8 out of 20 (40%) of the primary studies approached the inputs in correlation with the mechanisms and outcomes. This is critical to understand the underlying issues of how the inputs are expected to generate outcomes. Engagement ($M5$), for example, was the most prevalent mechanism (30%) followed by awareness ($M3$), integration ($M1$), and learning ($M2$). The only mechanism not referenced in the explored literature was error-proof ($M6$), as a result of the absence of inputs related to poka-yoke technologies.

Outputs are classified into intermediate ($O1j$), final ($O2j$), and economic ($O3j$) (2007) and represent the utmost interest of OHSMS. Our findings indicate that the reduction of the number of accidents ($O21$), the reduction of lost-time injuries rates ($O23$), and the contribution of OHSMS to the increase of economic KPIs ($O34$) are the prevalent outcomes. The focus on final outcomes has driven the attention of most scholars. However, the contribution of OHSMS to economic results also appeared as a relevant outcome in the revised literature. This result draws attention to the necessity of studies associated with OHSMS assessment in which multiple outcomes are evaluated, e.g. accident rate and costs. The existing literature is still more focused on safety efficacy (achievement of objectives) rather than evaluating relative measures (e.g. efficiency).

Finally, through the structural analysis, association rules were retrieved from the use of the algorithm *Apriori*, revealing how a set of initiatives and mechanisms (antecedent factors, named *lhs*) impact a particular outcome (consequent factor, named *rhs*) in a given context. As a result, 489 rules were retrieved for rhs = $\{O_{21}\}$, 281 for rhs = $\{O_{21}\}$, and 112 rules for rhs = $\{O_{34}\}$. These outputs were selected once they represent the most frequently cited in the revised literature as previously shown in Table 21. The top set of rules sorted by the parameter support and for each selected outcome is depicted in Table 22.

Table 22 – Association rules retrieved from *Apriori* for the outcomes O_{21}, O_{23} and O_{34}

Rule	Lhs		rhs	support	confidence	lift
[1]	$\{CL_2\}$	→	$\{O_{21}\}$	0.30	1	1.5000
[2]	$\{O_{23}\}$	→	$\{O_{21}\}$	0.20	1	2.5000
[3]	$\{PM_4\}$	→	$\{O_{21}\}$	0.15	1	1.5000
[4]	$\{M_5\}$	→	$\{O_{21}\}$	0.15	1	1.2500
[5]	$\{CL_2, O_{23}\}$	→	$\{O_{21}\}$	0.15	1	2.5000
[6]	$\{O_{11}\}$	→	$\{O_{21}\}$	0.10	1	1.6667
[7]	$\{O_{21}\}$	→	$\{O_{23}\}$	0.20	0.5000	2.5000
[8]	$\{CL_2, O_{21}\}$	→	$\{O_{23}\}$	0.15	0.5000	2.5000
[9]	$\{M_5, O_{21}\}$	→	$\{O_{23}\}$	0.10	0.6667	3.3333
[10]	$\{PM_3\}$	→	$\{O_{23}\}$	0.10	0.5000	2.5000
[11]	$\{I_{1b}\}$	→	$\{O_{23}\}$	0.05	1	5.0000
[12]	$\{O_{32}\}$	→	$\{O_{23}\}$	0.05	0.5000	2.5000
[13]	$\{O_{22}\}$	→	$\{O_{34}\}$	0.05	1	10.000
[14]	$\{PM_2\}$	→	$\{O_{34}\}$	0.05	0.5000	5.0000
[15]	$\{O_{14}\}$	→	$\{O_{34}\}$	0.05	0.5000	5.0000
[16]	$\{I_{17}\}$	→	$\{O_{34}\}$	0.05	0.5000	5.0000
[17]	$\{O_{22}, PM_2\}$	→	$\{O_{34}\}$	0.05	1	10.000
[18]	$\{CL_1, O_{22}\}$	→	$\{O_{34}\}$	0.05	1	10.000

Based on the association rules, patterns are identified as schematically given in Figure 27. For example, the reduction of incidents and occupational illness (O_{21}) is highly likely to occur (100% of confidence) when it is associated with engagement (M_5) (see rule [4]). Also, rule [9] concerning the outcome $O23$ as a consequent factor, shall be read as follows: the outcome $O23$ (reduction of the lost-time injury rate) is more likely to be achieved (66.67%) when it is associated with the output $O21$ (reduction of accidents) and with the mechanism M_5 (engagement).

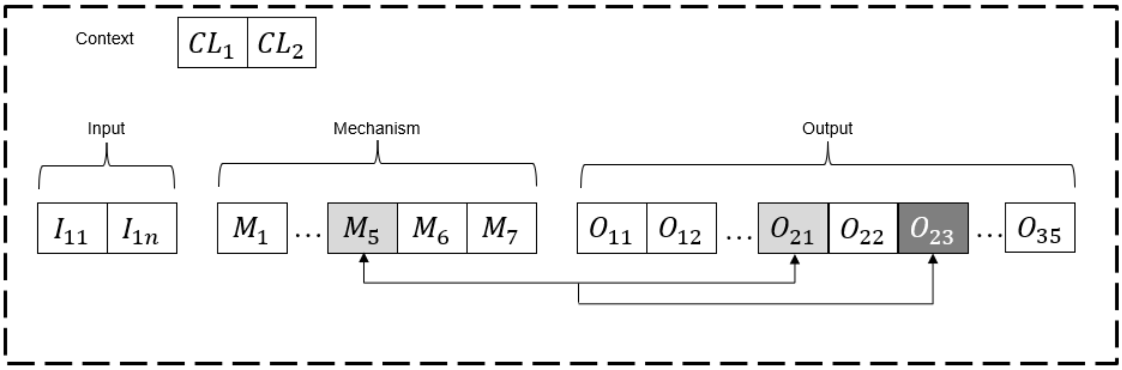


Figure 27 – Association rule $\{M_5, O_{21}\} \Rightarrow \{O_{23}\}$

Another example following the same reasoning is the rule [18] concerning the outcome $O34$: the increase of economic KPIs, such as sales and gross margin ($O34$), is more likely to be achieved (100% of confidence) when it is associated with the certified organizations ($CL1$) and the reduction on hazard exposure ($O22$).

By examining the set of rules, other inferences can be made by following the same reasoning, to understand how the critical elements composing OHSMS assessment instruments are connected. Economic outputs, for example, are more likely to occur when associated with both intermediate and final outcomes (see rules [13] and [15]). This underlying pattern means the reduction of accidents, for example, will more likely lead to economic benefits when it is also associated with safety climate and safety awareness, confirming the framework proposed by Robson et al. (Robson et al., 2007). This comprehension draws the attention of management to the importance of intermediate factors. By understanding how OHSMS can generate economic outcomes, the assumption verified in some organizations that safety incurs decreased productivity is broken, and more benefits might emerge through H&S initiatives.

Patterns related to performance measures were also identified. For example, the use of Effectiveness (PM_4) and Efficacy ($PM3$) as a performance measure in OHSMS assessment instruments is more likely associated with final outcomes (see rules [1] and [10]). In turn, the relative measure Efficiency ($PM2$) is more likely associated with economic outcomes (rule [14]). Those patterns suggest that efficacy and effectiveness are frequently used in the literature as the same measure, which deviates from concepts grounded in operations management (Coelli et al., 2005; Piran et al., 2020). Also, this finding suggests the use of performance measures to assess OHSMS has not been well addressed in the literature, and results might be impacted by this lack of understanding.

4.5 Discussion

The results of this study are discussed in two streams. First, a critical interpretative synthesis of the explored literature is presented and a conceptual base is given (subsection 4.5.1). Second, underlying issues are outlined in subsection 4.5.2.

4.5.1 Critical synthesis on conceptual rigor concerning OHSMS evaluation

Six systematic reviews of the literature and twenty primary studies comprised the corpus of analysis in this research. From the historical perspective, the discussion on the need for OHSMS evaluation emerged at the University of Michigan when the researchers raised the question of whether the international community should also consider the development of an ISO-compatible OHSMS.

Critical features in OHSMS assessment instruments were then identified, such as structure and scope, predictive validity, and implications to organizations of modest resources. Those features became seminal frameworks for creating a valid, reliable OHSMS assessment instrument.

However, these theoretical frameworks (Redinger and Levine, 1998; Robson et al., 2007) neglected the definition of performance measure in the context of OHSMS evaluation. As a result, assessment instruments were developed without considering well-established literature outside of safety science, and this is consistent with our results.

Considering the primary studies examined in this research, 9 out of 20 (45%) used somehow the concepts of efficiency, efficacy, or effectiveness as previously shown in Table 19. Others referred to ‘OHSMS performance’ generically to express either quantitative or qualitative measurement.

By taking into consideration rigorous concepts of efficacy and efficiency, only the empirical work of Su et al. (Su et al., 2005), and the efficiency indexes presented in the works of Ramli et al. (Ramli et al., 2011) and Saracino et al. (Saracino et al., 2015) used adequate concepts according to specialized literature (Piran et al., 2020).

The use of effectiveness as an OHSMS performance measure followed the same pattern (Arocena and Núñez, 2010; Redinger et al., 2002a; Redinger and Levine, 1998). Rather than assessing effectiveness, studies use criteria/rating score systems to measure the performance of categories under evaluation. As they do not clearly take into consideration the input/output ratio, the efficiency was not properly evaluated, and, as a consequence, neither was the effectiveness (as depicted in Figure 24).

In this context, our findings evidenced a lack of conceptual rigor in using performance measures in the context of OHSMS evaluation. The relevance of this finding is that, by considering OHSMS assessment instruments a critical mechanism to management reviews, and utmostly to improve safety results, the conceptual divergences might compromise the analysis of the results within organizations, leading to ineffective decisions. Therefore, based on well-recognized literature (Coelli et al., 2005; Farrel, 1957), a conceptual base is proposed to assess OHSMS performance as presented in Table 23.

Table 23 – Performance concepts in the OHSMS context

Concept	Definition
Efficacy	Efficacy is represented by the percentage of achievement of targeted outcomes, e.g., accident rate. For this particular evaluation, the resources used do not influence the level of the efficacy index.
Efficiency	Efficiency is a relative measure and has to do with the optimal use of resources. An efficient OHSMS focuses on maximizing its resources (inputs) to produce intermediate, final and economic outcomes. Thus, efficiency in OHSMS does not necessarily mean target achievements. It represents how close the system is to the efficiency frontier, which means the optimal ratio between all inputs and outputs.
Effectiveness	Effectiveness comprises both efficacy and efficiency. That means a multi-axial evaluation concerning target achievement and optimal use of resources. A highly effective OHSMS is one in which targets are reached under an optimized inputs/outputs ratio.

This conceptual base is also consistent with the specialized literature (Piran et al., 2020) and should not be neglected by researchers in the field of safety science. By using proper measures in the assessment instruments, it is argued that practitioners will be able to make better decisions. One dimension of assessment is to track the level of efficacy, e.g. in reducing 10% in accident rates year on year. In this case, organizations can verify how successful is the OHSMS concerning this particular objective. Another dimension is to identify internal and external benchmarks based on an analysis of efficiency. To do that, a technique like data envelopment analysis is required to support the analysis and to give directions on how to achieve higher performance. This is particularly important for organizations structured in business units, or to identify best-in-class within an industry. Finally, to evaluate under a broader perspective, the effectiveness (a composite measure of efficacy and efficiency) is recommendable to offer an in-depth analysis of OHSMS, as it connects all dimensions of performance measurements.

The application of those concepts in the evaluation of OHSMS should not be reduced to models and calculations. It requires specialized knowledge and a deep analysis of the elements that constitute the OHSMS, to interpret results and drive continuous improvement. Therefore, researchers and practitioners should consider this discussion when defining the strategy for designing OHSMS assessment instruments, keeping literature in the field of safety science aligned with other disciplines.

4.5.2 Underlying issues on OHSMS assessment

"Business databases reflect the uncontrolled real world, where many different causes overlap and many patterns are likely to co-exist" (Morandi and Camargo, 2015) (p.251). In this study, the co-occurrence of critical elements composing OHSMS assessment instruments was identified, and association rules mining (Tan et al., 2019; Zhang and Zhang, 2002) was used to unhide how these elements are connected.

In OHSMS assessment, it is reasoned to expect that all outcomes are a direct result of the initiatives. However, *Apriori* (Agrawal and Srikant, 1994) revealed that economic outcomes, for example, are more likely to occur when associated with intermediate and final outcomes, which means a consequent effect from other outcomes. This is important for practitioners to take into consideration when defining what has to be achieved first, and it draws attention to the importance of intermediate outcomes, such as safety awareness.

Another pattern identified was the association between a particular performance measure with a sub-category of outcome (intermediate, final, or economic). It suggests the necessity of clarity in using appropriate measures to specific outcomes. For example, efficacy is more associated with final outcomes, while efficiency is more related to economic outcomes.

Other significant patterns were identified. As previously mentioned in section 4.4, the definition of an OHS Policy, participation of managers and employees, and investment in training represented the most cited inputs and there is not much to add to that. However, our findings draw the attention of scholars and practitioners to some contemporaneous topics. More attention to mental health and behavior should be paid to the set of H&S initiatives since only a few studies focused on correlating these initiatives to the outcomes.

Consistent with the main initiatives, engagement was the most prevalent mechanism. The participation of the management and employees in topics concerning H&S is a sign of integration and engagement. However, while participation has to do with ‘being present’, the engagement is related to beliefs and actions that lead to the workplace’s change, seeking to promote safer work every time at everywhere. This explains why the engagement of top management has been considered a critical issue for H&S.

The comprehension of the mechanisms is fundamental to unhidden issues on H&S. Asking questions like how an initiative is expected to generate an outcome is fundamental for decision-makers. If the safety climate needs to be improved within an organization, involving the workforce in some safety meetings makes sense. This is because the mechanism of engagement is the key factor to lead the safety climate to a better level. All these connections have to be clear when designing OHSMS assessments.

Although engagement appeared as the most relevant mechanism, learning remains to play a fundamental role. It follows the assumption that the more qualified is the workforce, the safer their performance at work. Furthermore, learning contributes to increasing safety awareness. This is reasoned by considering the traditional approach that qualification has the potential to foresee hazardous conditions, and therefore, to avoid incidents.

Concerning the outcomes, the results were pretty much predictable. The reduction of lost-time injuries rate, the increase of economic KPIs, and a better safety climate represented the prevalent outcomes in the examined literature. Even though a significant part of the corpus of analysis is very recent, the outcomes remain closely the same as at the end of the ’90s. Also, the less cited outcomes were those sub-classified as intermediate when compared to the final and economic ones. This predictable pattern retrieved from the structural analysis might be related to the influence of certifications and standards that, somehow, influence practitioners in the way inputs and outputs are defined.

4.6 Conclusion

This study aimed at identifying the existing OHSMS assessment instruments and their critical elements. We investigate the patterns of how these elements are interconnected, as well as critically analyze the conceptual foundations of the assessment instruments examined. To do that, a robust research design was conducted based on a corpus of analysis comprised of 26 peer-reviewed studies.

As a result, 13 assessment instruments were identified, and the critical elements were presented in a form of a mixed coding scheme. The encoding scheme resulted in 39 critical elements categorized into 2 contexts, 16 inputs, 7 mechanisms, and 14 outputs.

From the thematic analysis, frequencies of occurrence indicate that certified organizations are the most cited context in OHSMS. Also, predictable initiatives (e.g. training) and outcomes (e.g. frequency rate) were prevalent when compared to contemporaneous topics, such as behavioral initiative and mental health. Concerning the mechanisms, engagement was the most significant one, followed by awareness.

Furthermore, underlying issues were revealed based on association rules mining. First, economic outcomes are more likely to occur as a consequent effect of other outcomes. It means the economic contribution from OHSMS does not result from a particular initiative. Rather, initiatives generate intermediate and final outcomes, and then economic results are verified. Second, patterns are verified between performance measures and sub-categories of outcome (intermediate, final, or economic). While efficacy is more likely associated with final outcomes, efficiency is more likely to co-occur with economic ones.

Finally, our results also evidenced divergences concerning the concepts of performance measures used in the field of OHSMS evaluation when confronted with well-defined literature outside of safety science. The relevance of this finding is that other disciplinary contexts should not be neglected by researchers in the field of health and safety since these conceptual divergences might compromise the analysis of the results within organizations, leading to ineffective decisions. To close this gap, a conceptual reference is proposed, and the definitions of efficacy, efficiency, and effectiveness in the field of OHSMS assessment are given.

This research presents some limitations. First, the study considered a relatively small database for association rules mining based on the research protocol criterion. Second, despite the structural analysis revealing how the elements that constitute an OHSMS are interconnected, causality was not investigated.

Future research is highly encouraged in the field of OHSMS assessment. Following the systemic perspective, scholars and practitioners are encouraged to employ techniques for OHSMS assessment, by taking into consideration the concepts suggested in this research.

APPENDICES CHAPTER 4

A 3. Protocol for systematic literature review

<table>
<tr><td>Research title</td><td colspan="6">Occupational Health and Safety Management System: A Systematic Literature Review of Assessment Instruments</td></tr>
<tr><td>Researcher:</td><td colspan="3">Gomes, Rodrigo Frank de Souza</td><td colspan="2">Revision: 00</td><td>Date: 23.05.2021</td></tr>
<tr><td>Stakeholders:</td><td colspan="6">OHS practitioners; senior executives; public representatives</td></tr>
<tr><td>Research questions:</td><td colspan="6">(RQ1) which performance measures, concepts, and assessment instruments have been used to assess an OHSMS?
(RQ2) what are the critical elements of OHSMS?
(RQ3) And how the critical elements are connected?</td></tr>
<tr><td>Research objectives:</td><td colspan="6">To investigate the concepts and the instruments utilized to assess occupational health and safety management systems within the organizations. In addition, to identify the underlying issues on how the OHSMS initiatives (inputs), mechanisms, and outcomes (outputs) are linked.</td></tr>
<tr><td>Review scope:</td><td>Amplitude:</td><td>Narrow</td><td>Deepness:</td><td>Deep</td><td>Review type:</td><td>Aggregative</td></tr>
<tr><td>Theoretical framework:</td><td colspan="6">Theoretical roots are presented in Figure 2. Highlights from previous reviews are outlined in appendix A 4.</td></tr>
<tr><td>Time horizon</td><td colspan="6">Not applied</td></tr>
<tr><td>Search strings:</td><td colspan="6">TITLE-ABS-KEY ("safety management system" AND "systematic review")
TITLE-ABS-KEY (("occupational health and safety management system" OR "OHSMS") AND ("performance" OR "efficiency" OR "efficacy" OR "productivity" OR "effectiveness")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English"))
TOPIC ("safety management system“ AND ("systematic review")
TOPIC: (("occupational health and safety management system" OR "OHSMS") AND ("performance" OR "efficiency" OR "efficacy" OR "productivity" OR "effectiveness"))
Refined by: DOCUMENT TYPES: (ARTICLE)</td></tr>
<tr><td>Search sources:</td><td colspan="6">Databases: Scopus and Web of Science</td></tr>
<tr><td>Searching approach:</td><td colspan="6">[X] Direct searching</td></tr>
<tr><td rowspan="2">Eligibility criteria:</td><td colspan="2">Inclusion criteria:</td><td colspan="4">Articles reviewed per peers and published in the selected databases, in English</td></tr>
<tr><td colspan="2">Exclusion criteria:</td><td colspan="4">Not accessible online; out of research scope based on Title/Abstract reading</td></tr>
<tr><td rowspan="5">Data analysis:</td><td colspan="2">Scientometry:</td><td colspan="4">Not applied</td></tr>
<tr><td colspan="2">Bibliometric analysis:</td><td colspan="4">Not applied</td></tr>
<tr><td colspan="2">Content analysis:</td><td colspan="4">[X]</td></tr>
<tr><td colspan="2">Thematic analysis:</td><td colspan="4">[X]</td></tr>
<tr><td colspan="2">Structural analysis</td><td colspan="4">[X]</td></tr>
<tr><td>Data synthesis:</td><td colspan="6">Critical interpretative synthesis</td></tr>
</table>

Source: Adapted from (Ermel et al., 2021; Morandi and Camargo, 2015)

A 4. Synthesis of systematic reviews (SR_i)

Id	Authors	# of reviewed studies	Research design	Main findings
SR_1	Robson et al. (2007)	13	Quality appraisal	Introduced a conceptual framework where the outcomes are split into intermediate, final, and economic outcomes and utilizes the concept of effectiveness as a performance measure of an OHSMS. Synthesized the best available evidence on the effects of mandatory and voluntary OHSMS initiatives on employee health and safety and associated economic outcomes.
SR_2	Brocal et al. (2018)	6	Bibliometric	Classified the main techniques included in the ISO/ IEC 31010:2009 standard applicable in the field of occupational safety and in line with the requirements for the ISO Standard 45001.
SR_3	Da Silva and Amaral (2019)	21	PRISMA	Identified the success factors, barriers, and indicators present in the process of implementing an Occupational Health and Safety Management System.
SR_4	Uhrenholdt Madsen et al. (2020)	16	Realist review	Identified in the literature three program theories (PT), i.e., institutional, operational, and compliance to deliver expected outcomes in Certified OHSMS based on context-mechanism-outcome.
SR_5	Lee et al. (2020)	51	PRISMA	Examed organizational, cultural, and individual-level factors related to workplace safety and health. Revealed six categories associated with the organizational factor (management commitment, management support, organizational safety communication, safety management systems, physical work environment, and organizational environment), two cultural (interpersonal support and organizational culture), and four individual-level (perception, motivation, attitude, and behavior).
SR_6	Wang et al. (2020)	4,852	Bibliometric	Built a time-zone distribution and a clustering map of high-frequency keywords into three dimensions based on an integrated perspective: individual, organization, and society.

A 5. Primary studies

Id	Authors	Assessment Instrument	Main findings
R_1	Redinger and Levine (1998)	Y	Introduced the Michigan Assessment Instrument (MAI) for OHSMS as a tentative of being considered a seminal OHSMS assessment tool. The MAI structure can be summarized as containing 5 organizing categories, 27 sections, and is based on 118 OHS principles and 486 measurement criteria. As a robust model, MAI contains (1) OHSMS principles; (2) measurement criteria for each principle; (3) suggested measures for each measurement criterion; (4) data collection mechanisms; (5) a scoring/ranking scheme, and 6) methods for score/rank interpretation.
R_2	Redinger et al. (2002)	Y	Reported an initial evaluation of the instrument's first four sections of their Universal Assessment Instrument (*UAI*), a wider approach from *MAI*. It is suggested that the variables/measures presented in the UAI's OHSMS initiation organizing category may contain performance measures that may serve as key leading indicators of overall OHS performance.
R_3	Arezes and Miguel (2003)	N	Outlined that safety performance measure is a crucial aspect of the OHSMS and examined qualitatively how this measurement can be done by considering safety culture, arguing that traditional safety indicators, such as accident statistics indices, may not reflect OHSMS performance.
R_4	Su et al. (2005)	Y	Collected frequency rate and severity rate data to evaluate the efficacy of voluntary protection programs (VPP) in Taiwan through statistical ANOVA method employing the software SPSS.
R_5	Chen et al. (2009)	N	The study included 11 PCB manufacturers in Taiwan aimed to be a pioneer with regards to OHSAS 18001 implementation and the consequent selection of performance indicators for health and safety performance appraisal. Key success factors to implement OHSAS 18001 such as 'top management promises and supports' were identified. Additionally, KPIs were categorized into conditional performance indicators (CPI), management performance indicators (MPI), and operational performance indicators (OPI) and rated based on a five-point Likert Scale.
R_6	Arocena and Núñez (2010)	Y	Employed descriptive statistics and regression models to analyze the characteristics and effectiveness of OHS management systems in a sample of 193 Spanish manufacturing SMEs through 12 preventive dimensions. Results of the cluster analysis indicated four different approaches and intensities of OHS management: (1) advanced, (2) technical, (3) basic, and (4) missing.
R_7	Hohnen and Hasle (2011)	N	Discussed the impact of certification on OHSMS in a Danish manufacturing company. Their findings suggest that certification and risk management are not purely technical or objective matters. In particular, results point towards significant although unintended shortcomings, especially the omission of complex work environment issues such as well-being, work intensity, and psychosocial work-related issues.
R_8	Ramli et al. (2011)	Y	Constructed an intelligent data analysis (IDA) employing a possibilistic regression model based on six inputs (OHS policy, OHSMS consultation sessions, training strategy, hazard assessment, risk control strategies, promoting of OHSMS improvements) and one output (OHSMS efficiency). Each coefficient represented a fuzzy parameter. The objective of LP was to determine the lower and upper limits of the fuzzy coefficients, maximizing or minimizing the OHSMS efficiency.
R_9	Redinger et al. (2011)	Y	A historical overview about OHSMS is presented and a detailed examination of four formal OHS management system standards is provided: (1) OHSAS 18001:2007 and ISO 14001-based approaches, (2) International Labor Office OHSMS, (3) ANSI/AIHA Z10:2005 and (4) The U.S. Occupational Safety and Health Administration's Voluntary Protection Program.
R_{10}	Yoon et al. (2013)	N	Presented the effects of OHSMS on the work-related rate in the context of top-100 construction certified companies in South Korea between 2006 and 2011. Results clearly showed decreasing figures after the OHSMS implementation. Additionally, when compared to non-certified companies the accident-rate was also lower.
R_{11}	Hedlung (2014)	Y	Examined the '5-Star System' of the National Occupational Safety Organisation (NOSA) of South Africa based on two properties: management's efforts to reach the objectives and the disabling incident incidence rate (DIIN). The 5-Star System intends to support OHSMS assessment and comprises 72 elements organized under five main sections: (1) premises and housekeeping, (2) mechanical, electrical and

			personal safeguarding, (3) fire protection and prevention, (4) incident recording and investigation, and (5) health and safety organization. Elements are subdivided into 300 components and the OHSMS assessment consists of assessing the level of compliance with each component and assigns a point score accordingly. As a combination of those two properties, a star-rating is then proposed. Although the method has been considered questionable by other researchers, results indicated that manufacturing companies, which were committed to the NOSA system, experienced fewer fatal and permanently disabling injuries than the reference group an inverse correlation between the Star-rating and the serious injury incidence rate.
R_{12}	Mahmoudi et al. (2014)	Y	Proposed a framework for OHSMS performance evaluation based on their empirical studies with construction companies. The authors identified seven main OHSMS elements based on a literature review of standards ("leadership and commitment," "policy and strategic objectives," "organization, resources, and documentation," "risk assessment and management," "planning," "implementation and monitoring," and "measuring performance, auditing, and reviewing") and A total performance index (TPI_i) was constructed to rank companies into five levels of performance. Their findings indicated the most important element at the organizational level is "leadership and commitment"
R_{13}	Saracino et al. (2015)	Y	Proposed a novel Methodology for the Implementation and Monitoring of Occupational Safety (M.I.M.O.SA.) aiming at quantifying the occupational health and safety level of a company and thus of its OHS Management System (OHSMS). The framework was built with the purpose of the self-assessment of the performance for both big and small scale enterprises, based on six key elements: (1) leadership and consistency of targets, (2) orientation to risk reduction and people protection in compliance with the law, (3) involvement, learning and development of individual culture, (4) continuous improvement and innovation, (5) formal and general compliance and (6) social responsibility. M.I.M.O.SA. counts on both planning and acting checklists for the assessment and each KPI is quantified by properly weighting its corresponding value. For an OHSMS overall assessment, the authors also introduced the Index of PErformance for Safety and HEalth (IPESHE) scaled into 0-100%. This study shows that the methodology presents some drawbacks and it needs to be improved in some points after being implemented in an Italian multi-utility company.
R_{14}	Podgórski (2015)	N	Employed Analytic Hierarchy Process (AHP) for prioritization and selection of leading KPIs for measuring an OHSMS but it did not introduce any integrated method to conduct an overall evaluation. Additionally, five previous methods for measuring the performance of the safety management systems were briefly described. (1) Safety Element Method (SEM), (2) Universal Assessment Instrument (UAI), (3) Self-Diagnostic OHS Tool, (4) Tripod Delta, and (5) Safety Climate Assessment Questionnaires. Those methods can be classified as follows: (1) and (2) are composed of a set of selected OHSMS elements (ie. principles, categories, or measurement criteria) to be evaluated. While (1) is relatively simple, (2) might be considered quite complex to be applied for some organizations. (3) is a self-assessment method; (4) is a comprehensive questionnaire-based tool that assumes that the most efficient manner of accident prevention is to control the working environment and identify its weak areas which may lead to human errors or system disturbances, and consequently to accidents; (5) explores the level of safety culture using questionnaires for the measurement of safety climate as an alternative OHSMS assessment.
R_{15}	Bianchini et al. (2015)	Y	Introduced the Efficacy Index (EI) to objectively quantify the effective implementation of an Occupational Health and Safety Management System (OHSMS). EI is a basic ratio ($EI = B/(A + B)$ where A is the consequent costs related to the accident, incident, near miss, and professional disease, and B are the prevention costs to prevent and protect from accident, incident, near miss and professional disease. The research was conducted using data from nine Italian SMEs and results of EI varied between 0.27 and 0.86, indicating a clear economic approach for the OHSMS assessment.
R_{16}	Gopang et al. (2016)	Y	Undertaken an empirical study in the industrial zones of Pakistan to identify the relationship between OHSMS and the performance of small- and medium-sized enterprises (SMEs). By evaluating 35 companies through descriptive and inferential statistics tools, their results revealed a moderate positive correlation between OHSMS and the performance of SMEs.
R_{17}	Mohammadfam et al. (2017)	Y	Based on a sample of six companies involved in large-scale industrial projects in Iran, the authors evaluated the responses of 30 OHS managers with regards to a structured survey based on five core OHSMS activities proposed by OHSAS 18001: policy, planning, implementation, checking, and management review. The findings of this study indicate that certified companies are most likely to enforce OHS rules and procedures.

R_{18}	İnan and Yılmaz (2017)	Y	Aimed to build a multiple attribute decision making (MADM) model for determining and comparing the firms' OHSMS performances of certified companies, based on the same criteria applied by Mohammadfam et al. (2017). MADM is valid for assessing performance among firms based on OHSAS requirements. However, it does not approach how to measure the performance of an OHSMS such as effectiveness or efficacy.
R_{19}	Nordlöf et al. (2019)	Y	Performed a cross-sectional study in manufacturing companies in Sweden and employed statistical and regression analysis to investigate factors associated with OHSM practices. Company size, safety culture, and creditworthiness were found associated factors with OHSM practices.
R_{20}	Skład (2019)	Y	Developed a fuzzy cognitive map (FCM) to examine the impact of OHSMS on the effectiveness of the system. Through a series of simulations, it was finally proven that safety performance increased most significantly under the influence of improvement of the leadership process, confirming the greatest positive impact on its OHSMS effectiveness.

A 6. Binary code-document matrix for Context-Input-Mechanism-Outcomes

	CL.1	CL.2	I11	I12	I13	I14	I15	I16	I17	I18	I19	I1a	I1b	I1c	I21	I22	I31	I32	I33	M1	M2	M3	M4	M5	M6	M7	O11	O12	O13	O14	O15	O21	O22	O23	O24	O25	O31	O32	O33	O34
R1		1	1	1	1	1	1	1		1	1					1	1	1	1		1			1																
R2		1	1	1	1																			1																
R3						1					1		1				1				1			1			1					1		1				1		
R4		1																														1		1						
R5		1		1	1												1					1		1			1					1			1					
R6	1				1			1									1	1	1	1												1								
R7		1					1	1		1												1				1														
R8	1		1	1		1		1									1																1							1
R9			1	1	1	1	1	1		1	1				1		1	1	1	1	1	1		1																
R10		1																				1	1					1		1		1							1	
R11		1					1	1	1	1		1						1	1	1				1								1		1						
R12		1	1	1		1		1			1					1																1								
R13			1	1	1											1	1																							
R14			1		1	1	1			1	1	1		1		1	1																							
R15																																					1	1		
R16						1		1	1								1		1								1			1										1
R17	1	1																														1		1						
R18		1	1			1	1	1			1	1		1	1	1	1																							
R19	1		1	1	1	1		1				1		1					1												1									
R20																																								
T	4	10	9	8	8	9	6	10	2	5	6	4	1	3	2	5	10	4	6	3	3	4	1	6	0	1	3	1	0	2	1	8	1	4	1	0	1	2	1	2

A 7. Inputs

	I11		I12		I13		I14		I15		I16		I17		I18		I19		I1a		I1b		I1c		I21		I22		I31		I32		I33	
	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%
R_1	1	9.1%	1	9.1%	1	11,1%	1	11.1%	1	14.3%	1	9.1%			1	20%	1	14.3%									1	16.7%	1	10%	1	25%	1	16.7%
R_2	1	9.1%	1	9.1%	1	11,1%																												
R_3							1	11.1%									1	14.3%			2	100%							1	10%				
R_4																																		
R_5			4	36.4%	1	11,1%																							1	10%				
R_6					1	11,1%					1	9.1%																	1	10%	1	25%	1	16.7%
R_7									2	28.6%	1	9.1%			1	20%																		
R_8	2	18.2%	1	9.1%			1	11.1%			1	9.1%																	1	10%				
R_9	1	9.1%	1	9.1%	1	11,1%	1	11.1%	1	14.3%	1	9.1%			1	20%	1	14.3%							1	50%			1	10%	1	25%	1	16.7%
R_{10}																																		
R_{11}									1	14.3%	1	9.1%	1	33.3%	1	20%			1	16.7%											1	25%	1	16.7%
R_{12}	1	9.1%	1	9.1%			1	11.1%			1	9.1%					1	14.3%									1	16.7%						
R_{13}	1	9.1%	1	9.1%	1	11,1%																					1	16.7%	1	10%				
R_{14}	1	9.1%			1	11,1%	1	11.1%	1	14.3%					1	20%	2	28.6%	2	33.3%			1	33.3%			2	33.3%	1	10%				
R_{15}																																		
R_{16}							1	11.1%			2	18.2%	2	66.7%															1	10%			1	16.7%
R_{17}																																		
R_{18}	1	9.1%					1	11.1%	1	14.3%	1	9.1%					1	14.3%	1	16.7%			1	33.3%	1	50%	1	16.7%	1	10%				
R_{19}	2	18.2%	1	9.1%	2	22,2%	1	11.1%			1	9.1%							2	33.3%			1	33.3%									1	16.7%
R_{20}																																		
T	11	100%	11	100%	9	100%	9	100%	7	100%	11	100%	3	100%	5	100%	7	100%	6	100%	2	100%	3	100,0%	2	100%	6	100%	10	100%	4	100%	6	100%

A 8. Outputs

	O_{11}		O_{12}		O_{13}		O_{14}		O_{15}		O_{21}		O_{22}		O_{23}		O_{24}		O_{25}		O_{31}		O_{32}		O_{33}		O_{34}	
	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%	f	%
R_1																												
R_2																												
R_3	1	33.3%									1	10%			1	25%							1	50%				
R_4											1	10%			1	25%												
R_5	1	33.3%									1	10%					1	100%										
R_6											1	10%																
R_7																												
R_8													1	100.0%													2	50%
R_9																												
R_{10}			1	100%			1	50%			3	30%													1	100%		
R_{11}											1	10%			1	25%												
R_{12}											1	10%																
R_{13}																												
R_{14}																												
R_{15}																					1	100%	1	50%				
R_{16}	1	33.3%					1	50%																			2	50%
R_{17}											1	10%			1	25%												
R_{18}																												
R_{19}									1	100%																		
R_{20}																												
T	3	100%	1	100%	-	0%	2	100%	1	100%	10	100%	1	100%	4	100%	1	100%	-	0%	1	100%	2	100%	1	100%	4	100%

5 MEASURING EFFICIENCY OF SAFE WORK ENVIRONMENT FROM THE PERSPECTIVE OF DECENT WORK[7]

Abstract

Decent Work Agenda consists of a comprehensive initiative for promoting safety at work and social protection. Over 20 years since its conceptual release, measuring the progress of its elements is still challenging even after the publication of the decent work indicators guideline by the International Labour Organization in 2012. To close this gap, we use a Directional Distance Function (DDF) to measure the efficiency of safe work environment – a substantive element of decent work. To illustrate the application of DDF in a reality-based case, we conducted a longitudinal study in a multinational organization. Data were collected from 21 branches of the company over 4 years (2018-2021). Our results indicate that 49% of the branches were efficient year on year composing an overall efficiency score of 0.82. Also, branches with higher efficiency scores were mostly associated with reduced undesirable outputs, i.e. work-related accidents and lost days equal to or close to zero. In contrast, inefficient branches presented expanded undesirable outputs or excessive use of resources relative to others. This research presents some contributions. One is the novelty approach of measuring the efficiency of safe work environment using a DDF model grounded in a real-world application. Another is the managerial benefits of identifying benchmarks based on efficiency scores, as well as revealing potential improvements as a mechanism to reduce decent work deficits. From a modeling perspective, our conclusions suggest caution in considering only efficiency to measure safe work environment due to its relative nature. Thus, further studies are recommended to explore the use of composite measures in the analysis of decent work.

Keywords: Decent work. Safe work environment. Efficiency. Directional Distance Function.

5.1 Introduction

A healthy and safe (H&S) work environment not only is desirable from the workers' perspective but also contributes considerably to labor productivity and promotes economic growth (Heuvel et al., 2017). Despite technological progress, the globalization of economies, the availability of artificial intelligence, and robotization, human capital continues to be the key factor behind the economic and sustainable development of all countries (Rantanen et al., 2020).

From the perspective of both public and private sectors, growing efforts have been verified to draw attention to the importance of H&S work. Yet, a recent report published jointly by the World Health Organization (WHO) and International Labour

[7] Article submitted to Journal World Development.Manuscript number #WD-25882.

Organization (ILO) estimates that, in 2016, more than 1.8 million deaths and approximately 90.8 million disability-adjusted life years were officially informed as a consequence of work-related injuries (WHO and ILO, 2021), accounting for economic losses estimated at 3.94% of the global Gross Domestic Product (Brocal et al., 2018; Wang et al., 2020).

In this context, comprehensive international initiatives including the Decent Work Agenda (DWA) have been proposed to deliver quality jobs along with social protection and respect for rights at work to achieve sustainable, inclusive economic growth, and eliminate poverty (ILO, 2022a). The DWA was launched in 2000 as a new strategy adopted by ILO to ensure human-oriented development and to provide an effective response to the challenges of globalization. In short, DWA sums up the aspirations of people in their working lives and plays a fundamental role in promoting safety at work initiatives and social protection (ILO, 2022a). Furthermore, in September 2015, DWA confirmed its importance and became an integral element of the new 2030 Agenda for Sustainable Development proposed by the United Nations member states, which includes 17 sustainable development goals (SDGs) and 169 indicators.

As a fundamental part of the DWA, H&S is represented by the 'safe work environment', one of the ten substantive elements considered in the decent work measurement framework (DWMF) (ILO, 2013). The framework was adopted by ILO on the occasion of the 18th International Conference of Labour Statisticians in December 2008, and it is linked with the four strategic pillars of the DWA: (i) international labor standards and fundamental principles and rights at work, (ii) employment creation, (iii) social protection and (iv) social dialogue and tripartism.

Measuring the progress towards decent work is critical and still challenging even after the first published version of the ILO manual on decent work indicators in 2012. These statistical indicators presented in the framework help to identify, for instance, which population groups (e.g. countries) within the priority areas may be experiencing decent work deficits based on the selection of a set of decent work indicators. Worth to mention that a good assessment relies on accurate data and a selection of an appropriate analysis method, conditions that require transparency, and good management practices. Thus, a similar approach is suitable for identifying internal benchmarks and supporting practitioners to prevent work-related injuries in operations facing deficits. At the organizational level, the driver of decent work is represented by efforts related to Occupational Health and Safety Management System (OHSMS). This system provides the foundation for the implementation of strategies to promote H&S in practice, and its performance represents an important mechanism for management reviews.

Distinct performance measures have been used in the literature for evaluating OHSMS, such as the use of efficacy, efficiency, and effectiveness (Bigelow and Robson, 2005b; Ghahramani, 2017; Podgórski, 2015; Redinger et al., 2002b). From the perspective of the DWMF, apart from just monitoring the efficacy (checking outcomes against goals), measuring efficiency arises as a promising solution to distinguish between efficient and inefficient decision-making units (DMUs) and to give directions for improvements in H&S. While efficacy concerns achieving goals, the efficiency is a relative measure that compares realized productivity with maximum productivity (Farrel, 1957). In other words, it takes into consideration the resources used (inputs) and outcomes obtained (outputs).

Consistent with DWA, resources should be used to promote safer work environments. Thus, this work aims to measure the efficiency of the safe work environment from the DWMF perspective. To do that, a model based on Directional Distance Function (Chambers et al., 1996; Chung et al., 1997) is proposed and tested in a real-world case study within a large organization in the elevator industry comprising 21 branches (DMUs) over 4 years (2018-2021). The model deals with undesirable outputs without requiring any data manipulation, a typical pitfall in DEA applications (Dyson et al., 2001; Halkos and Petrou, 2019; Sarkis, 2002). Also, the study explores the concept of internal benchmarking and estimates targets for each variable composing the inefficient branches (de Souza et al., 2018). Finally, efficacy was considered to verify the correlation between efficient DMUs and the achievement of the goals.

This study offers a novelty approach to measuring the efficiency of safe work environment using a DDF model in a longitudinal study. The model was tested in a real case and practical contributions were verified, e.g. managerial benefits in identifying internal benchmarks based on efficiency measurement, as well as the potential for improvements to be considered by the inefficient units as a mechanism to promote safer work environments. In addition, this work addresses an assessment instrument to be considered and implemented by practitioners. Furthermore, it is original in using the concept of internal benchmarking as a driver for advances in decent work at the organizational level. From a modeling perspective, our results suggest caution in considering only the efficiency for measuring safe work environment due to its relative nature. If a set of DMUs under assessment is characterized by poor safety management, the frontier of efficiency might be composed of DMUs in which the performance level is behind the intended managerial goals established by the organization. Therefore, internal benchmarking solely based on efficiency might be insufficient to promote significant changes in decent work deficits, and another dimension to compare the performance level against strategic goals is necessary. Following this reasoning, the use of efficacy as a complimentary and contextual measure is necessary to ensure a more effective assessment of safe work environment.

This paper is structured as follows. Section 5.2 introduces the main aspects of the literature that contextualizes this work. Section 5.3 outlines the research design. Section 5.4 presents the results based on a real-world application, and discussions are made in section 5.5. The conclusions, limitations of this work, and further research prospects are drawn in section 5.6.

5.2 Related literature

5.2.1 Measuring safe work environment

The measurement of decent work was first presented to the 18th International Conference of Labour Statisticians in December 2008 (ILO, 2022b, 2013). As a major outcome of this tripartite meeting, a framework covering ten substantive elements was introduced, including the safe work environment. All elements comprising the so-called decent work measurement framework are somehow linked to the four DWA pillars (ILO, 2022a) (rights at work, employment, social protection, and social dialogue). Also, they are connected to sustainable development goal 8 of the 2030 Agenda of the United Nations, which seeks to promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all (Nations, 2015).

According to the aforementioned framework, a set of indicators for measuring safe work environment is presented in Table 24.

Table 24 – Decent work indicators for measuring safe work environment (ILO, 2013)

Statistical of decent work indicators	Related concepts	Variables
SAFE1: Occupational injury frequency rate, fatal	Fatal work-related accident or work-related illness	F_{acc}
SAFE2: Occupational injury frequency rate, non-fatal	Non-fatal work-related accident or work-related illness	NF_{acc}
SAFE3: Time lost per occupational injury	Days lost by cases of temporary incapacity	$Lost_t$
SAFE4: Labour inspection rate	Labour inspections	N_{insp}

The four indicators proposed in DWMF represent critical elements of safety at work. The fatal occupational injury frequency rate (SAFE1), e.g., provides information on the number of fatal occupational injury cases (F_{acc}) per working hours (w_h) by the concerned population during a given period (SAFE1 = $F_{acc}\ x\ 1.000.000/w_h$). It is a measure of the risk of having a fatal work-related accident based on the duration of exposure to adverse work-related factors, normalized to 1 million hours. The same reasoning is used for the non-fatal occupational injury frequency rate (SAFE2 = $NF_{acc}\ x\ 1{,}000{,}000/w_h$), just modifying the type of injury in scope.

Time lost due to occupational injuries ($Lost_t$) measures the consequences of occupational injuries in terms of lost days (SAFE3). It may be used to design prevention mechanisms and estimate the cost of occupational injuries. Hence, it gives a quantifiable measure of the impact of the injuries which is comparable across cases. Finally, the rate of inspectors per 10,000 employed persons (SAFE4) is a “crude proxy measure” of the resources for monitoring and enforcing work conditions and standards. It can also be represented (or replaced) by the quantity of labor inspections conducted in the workplace (N_{insp}).

5.2.2 Safety performance: the efficacy and efficiency measures

Performance measurement plays an important role in safety at work. This is because “people behave under influence of how they are measured” (Goldratt, 1990). This quote retrieved from the Theory of Constraints draws attention to the fact that performance measures and indicators must be carefully defined, and properly connected to other constructs that influence safety at work (Gomes et al., 2022). For instance, if authorities increase taxes on organizations with high accident rates, it is expected that managerial actions will be focused on H&S. However, it might cause a side effect in organizations with poorer safety climate, such as the increase in underreporting cases, or an overlooking of other important metrics, not included in regulations (Probst and Estrada, 2010). Therefore, the definition of safety performance measures should consider the risk of causing an unbalanced managerial focus, where some factors are overmanaged and others overlooked.

A common approach to evaluating safety performance is to check relevant indicators (e.g., accident frequency rate) within a given period (t) against a defined goal. This is called efficacy (δ^t) (Enrique and Marta, 2020; Piran et al., 2020). It might be represented by either a binary approach $\delta^t = \{0; 1\}$ (non-achieved, achieved) or a percentage of achievement $\delta^t = \{0 - 1\}$, e.g., if $\delta^t = 0.9$, the unit under analysis achieved 90% of its goal. We argue binary approach is suitable for critical goals, e.g. zero fatality, because it makes no reasonable sense to track its percentage of achievement. In contrast, for operational indicators, efficacy might be measured in a range between 0 and 1, in which 1 means 100% of achievement. Note that the efficacy of a DMU k under assessment might be also composed of multiple components. In this case, it is represented as the multiplication of each of its elements in a period t: $\delta_k^t = \delta_1^t \,.\, \delta_2^t \ldots \delta_n^t$. To illustrate that condition, consider the composite efficacy δ_k^t based on two goals, e.g. zero fatalities and non-fatal accident rate. Thus, if the DMU k reached goal 1 and only 90% of goal 2, the composite efficacy δ_k^t is calculated as follows: $\delta_k^t = \delta_1^t \,.\, \delta_2^t = 1x0.9 = 0.9$).

Also, efficacy is commonly used in real-world managerial applications to track performance over time. In this case, a typical representation of performance improvement is $\delta_k{}^{t+1} > \delta_k^t$, which means the efficacy for a firm k in the period $(t + 1)$ is higher than the one in the period (t). Besides its intuitiveness and vast prevalence in management reviews, this assessment model presents some limitations (Karanikas, 2016). One is that it does not consider the resources used to achieve the outcomes (Enrique and Marta, 2020). It means that, even in the circumstance of a goal being achieved, resources might be poorly used, and opportunities for improvements remain hidden. Another is that the setting of management goals might be impacted by management biases, following unclear reasoning without considering comparable units to set goals.

Furthermore, this approach does not identify potential benchmarks in a group of comparable units (e.g., organizations in the same sector or business units within one organization). This is quite useful as managerial information to investigate deficiencies in some units by considering comparable references with higher performance. In short, we argue efficacy is adequate for tracking targeted indicators but not sufficient for managing health and safety systems as a whole.

In this context, efficiency appears as a promising measure to qualify the assessment process since it takes into consideration the resources used by each comparable unit. Moreover, it allows the use of benchmarking as a powerful organizational mechanism to drive continuous improvement in the field of H&S (Piran et al., 2022).

Differently from efficacy, the concept of efficiency ($\emptyset$) used in this work follows the reasoning proposed by Farrell (1957) and considers the relation between outputs y_{rj} and inputs x_{ij} from a set of n decision-making units (DMUs). Thus, the efficiency of a firm k under assessment is defined by $\emptyset_k = \frac{Output(y_{rj})}{Input\ (x_{ij})}$, i.e., it compares realized productivity with a maximum productivity of a group of units under assessment ($j = 1, \ldots, n$), seeking to optimize performance by decreasing the use of resources (inputs $r = 1, \ldots, s$) and/or increasing the outcomes (outputs $i = 1, \ldots, m$). A DMU is considered efficient if and only if the performance of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs (William W. Cooper et al., 2011b).

A major benefit of measuring performance via efficiency is the identification of efficient units that can be considered benchmarks to others. In addition, based on comparable empirical data, the use of internal benchmarking reveals opportunities for better allocating resources and supports the organization to scale appropriate efforts to increase overall efficiency. Moreover, efficiency measurement is useful for the assessment of management initiatives over time with comparable organizational units. Finally, it qualifies the process of identifying and disseminating good practices among organizational units.

Benchmarking can be considered a systematic process for evaluating products, services, and work processes in organizations recognized for their best practices in order to achieve organizational improvement (de Castro and Frazzon, 2017; Vinodh and Aravindraj, 2015). Benchmarking types include external (also called functional) or internal (Jain et al., 2008; Piran et al., 2021). External benchmark is used to compare against the best organizations operating in the same industry/sector. It is difficult to be conducted because reliable data are commonly unavailable. Internal benchmarking, in turn, compares the performance of units within one organization (Piran et al., 2022). It can be used to address improvements since it seeks to unhide opportunities for improvements through the comparison of homogeneous operations. Moreover, internal benchmarking in longitudinal applications can offer additional benefits for management, including (i) capturing the effect of initiatives that present some delay between the intervention and verifiable outcomes, (ii) identifying changes in the work environment based on the variation of efficiency scores over time, (iii) assessing the effectiveness of corporate policies and H&S initiatives in different subsidiaries within one organization, and (iv) observing side-effects from H&S initiatives, such as the impact on quality and employee satisfaction.

The evaluation of safety performance can benefit from both efficacy and efficiency measures. Both offer complementary information for practitioners to establish continuous improvement management toward preventing work-related accidents. Thus, the next subsection outlines the use of non-parametric techniques (e.g. DEA) as a means to evaluate efficiency in several areas, and different contexts (William W. Cooper et al., 2011b). Consistent with Golany and Roll (1989), we aimed for minimal repetition of concepts published elsewhere for which frequent reference is made throughout this work, and only necessary concepts are retrieved for the contextualization of this work.

5.2.3 Measuring efficiency with undesirable outputs

Occupational health and safety management systems are intended to manage occupational risks, and utmostly to establish initiatives and preventive measures to eliminate, neutralize or mitigate undesirable events, e.g., work-related accidents, occupational diseases, and lost days (Wang et al., 2020). As with any other system, OHSMS is composed of inputs and outputs (Redinger et al., 2011). Therefore, in a set of comparable units under analysis, the measurement of efficiency is applicable to define the efficient units, as well as the opportunities for improvement in the inefficient ones.

A well-recognized technique for measuring efficiency is Data Envelopment Analysis (DEA). It was originally developed by Charnes et al. (1978) for measuring the relative efficiency of a set of DMUs producing multiple outputs from multiple inputs. As a non-parametric frontier technique, it identifies the set of efficient and inefficient DMUs and defines how each inefficient DMU can improve its performance in terms of decreasing the use of resources (inputs) and/or increasing the results (outputs) compared to a set of benchmarks whose inputs and outputs were actually deserved in practice (Alves et al., 2021).

The prevalent standard modeling in DEA relies on the assumption of outputs to be increased, and inputs to be decreased (Scheel, 2001), under either constant return to scale (CRS) (Charnes et al., 1978) or variable return to scale (VRS) (Banker et al., 1984). However, in some real-world applications, undesirable outputs might exist in the variables under analysis. Recently, this limitation of the standard DEA model in dealing with undesirable outputs has received greater research attention, especially because this is particularly the case of eco-efficiency and other social studies (Halkos and Petrou, 2019; Oliveira, 2008; Oliveira et al., 2017).

Two mathematical approaches exist in the literature to deal with the issue of undesirable outputs: indirect and direct (Ramli and Munisamy, 2013). The indirect approach is largely used in the literature and it considers data transformation of the undesirable output so they can be included in the standard model. However, manipulating the data require translation of the results before its use as managerial analysis to avoid being merely mathematical modeling (Dyson et al., 2001). Examples related to this approach include, e.g., the additive inverse which incorporates the undesirable outputs as desirable outputs in a form of function (a) $f^k(y_b) = -y_b^k$, or multiplicative inverse in a form of function (b) $f^k(y_b) = 1/y_b^k$, or the use of undesirable output as input. In the functions (a) and (b), y_b are the undesirable outputs associated with a DMU k under assessment. In In contrast, the direct approach requires no data manipulation, and it considers the original undesirable output value directly in the model. Examples following this approach include, e.g., directional distance function (DDF) (Chambers et al., 1996; Chung et al., 1997).

In general terms, the concept of the DDF is to expand desirable outputs (y_g) and reduce inputs (x) and undesirable outputs (y_b) simultaneously based on a given directional vector $g(-x, -y_b, y_g)$ (Chung et al., 1997). The directional vector $g^{\rightarrow}$ indicates the direction of change for the inputs (x), desirable (y_g) and undesirable outputs (y_b). As presented in Formulations (1) and (2), the production technology (T) consists of the set of all feasible inputs and outputs for a certain production process (or any system under assessment). In formulation (2), $D^{\rightarrow}$ represents the directional distance function. Its objective is to maximize the existing 'potential improvement' (β) to scale inputs and desirable/undesirable outputs within the technology and its efficient frontier, defining the best practices (benchmarks). The optimum value for β can be interpreted as the inefficiency or distance from the frontier, in other words, the scope for improvement for a DMU k under assessment.

$$T = \{(x, y_b, y_g): x \text{ can produce } y_b \text{ and } y_g \} \quad (1)$$

$$D^{\rightarrow}\left(x,\ y_b,\ y_g,\ g_x,\ g_{y_b},\ g_{y_g}\right) = \max \beta\ : \left(x - \beta.|g_x|, y_b - \beta.|g_{y_b}|, y_g + \beta.|g_{y_g}|\right)\ \in T \quad (2)$$

As reasoned by Halkos et al. (2019), many researchers have pointed out that a DDF approach is the best solution as it allows for the simultaneous increase of desirable outputs and the reduction of undesirable outputs. Also, it solves a critical pitfall reasoned by Dyson et al. (2001) in DEA models dealing with anti-isotonic factors, in which data transformation is required. An anti-isotonic factor is characterized by the condition in which increasing inputs incur a reduction in efficiency, and in which increasing outputs incur a gain in efficiency is not confirmed. Furthermore, the results give a clear managerial interpretation since further data transformation is not needed.

A comprehensive review of these approaches, including mathematical treatment, is found mainly in the works of Golany and Roll, Scheel, Dyson et al., Sarkis, Ramli and Munisamy, and Halkos et al. (2001; 1989; 2019; 2013; 2002; 2001). Since this research is focused on the application of DDF to evaluate a real-world case, the foundations related to formulations (1) and (2) aforementioned are not in-deep discussed.

Studies using frontier techniques such as DEA or DDF to measure efficiency in the field of occupational health and safety are limited in the literature when compared to other areas. The prevalence of H&S applications is in specific areas such as construction safety (Dou and Zheng, 2011; El-Mashaleh et al., 2010; Kang et al., 2020; Nahangi et al., 2019; Nissi and Rapposelli, 2012) and manufacturing/industrial safety (Beriha et al., 2011; Said et al., 2013; Yeh, 2017). Also, researchers in accident analysis and prevention have increased publications using DEA in traffic/road safety (Alper et al., 2015; Hermans et al., 2009; Shen et al., 2020; Tešić et al., 2018).

When it comes to applications related to DWA, studies are scarce. No previous study was found on how to measure the safe work environment, one out of the ten elements of the decent work measurement framework. For instance, the works of Thore and Taverdyan (2009), Cooper, Thore and Taverdyan (2011), and Ertugrul

Karsak and Goker (2019) are focused on comparing efficiencies among selected countries in a broader perspective of DWA. Blanco, Bares and Ferasso (2022) explored SDG 8 using DEA to identify benchmarks in a sample of Latin American universities. Finally, the work of Ning et al. (2018) is the one dedicated to evaluating eco-efficiency among Chinese forestry industries.

It is clear the existence of a knowledge gap on how safe work environment should be measured from the DWA perspective. The literature examined does not explore how to connect the statistical indicators proposed in the measurement framework to any practical method or assessment instrument that could be used within organizations. We intend to better address this issue by applying the classic DEA model, explaining the necessary data transformation to identify efficient units, and using internal benchmarks to give directions on how to improve the performance of those outside the efficient frontier.

5.3 Work method

This work adopted case-based research as the methodological approach. A case study is particularly suitable when a phenomenon is broad and complex, and cannot be studied outside the context in which it occurs (Benbasat et al., 1987; Bonoma, 1985; Dubé and Paré, 2003; Yin, 1994). This is exactly the case with investigating decent work at the organizational level where context plays a fundamental role. Also, this approach is consistent with studies following ‘reality-based safety science’, where theory is grounded in rigorous observations of existing practice, and the practice is based on established theory (Rae et al., 2020).

More specifically, we conducted a longitudinal study over 4 years in a large organization in the elevator industry, structured in a set of business units. The criteria used to select the case and the units of analysis, the selection of variables, and the model configuration are presented in the following subsections.

5.3.1 Unit of analysis and context

The research was conducted in a subsidiary of a global organization. The organization operates in the elevator industry (manufacturing and services), doing business in more than 100 countries, with over 1,000 sales and service locations, and over 50,000 employees. Among its strategic principles, health and safety appear as a core value, and management is considered a key factor in ensuring safe work at all levels. To be aligned with its purpose, the organization set up a zero-fatality goal. Also, a decreasing non-fatal accident rate goal has been established year on year. To be consistent with its principles, these goals comprise a set of objectives that impact directly annual incentives for managers and operational leaders.

The subsidiary selected in this analysis is an Operating Unit with headquarters located in the south of Brazil, composed of 1 factory, 1 shared service center (SSC), and 21 branches. Each branch has its operational structure, including a manager, salesforce, decentralized workforce, and dedicated H&S professionals. The branches provide customers with a variety of services, including maintenance, repairs, retrofits, and new installations of equipment produced in the factory. SSC, in turn, is a back-office entity in charge of administrative processes. As presented in Figure 28, by considering the necessary condition of homogeneity and comparability to evaluate relative efficiency, only the 21 branches were selected as DMUs since both factory and SSC operates for different purposes.

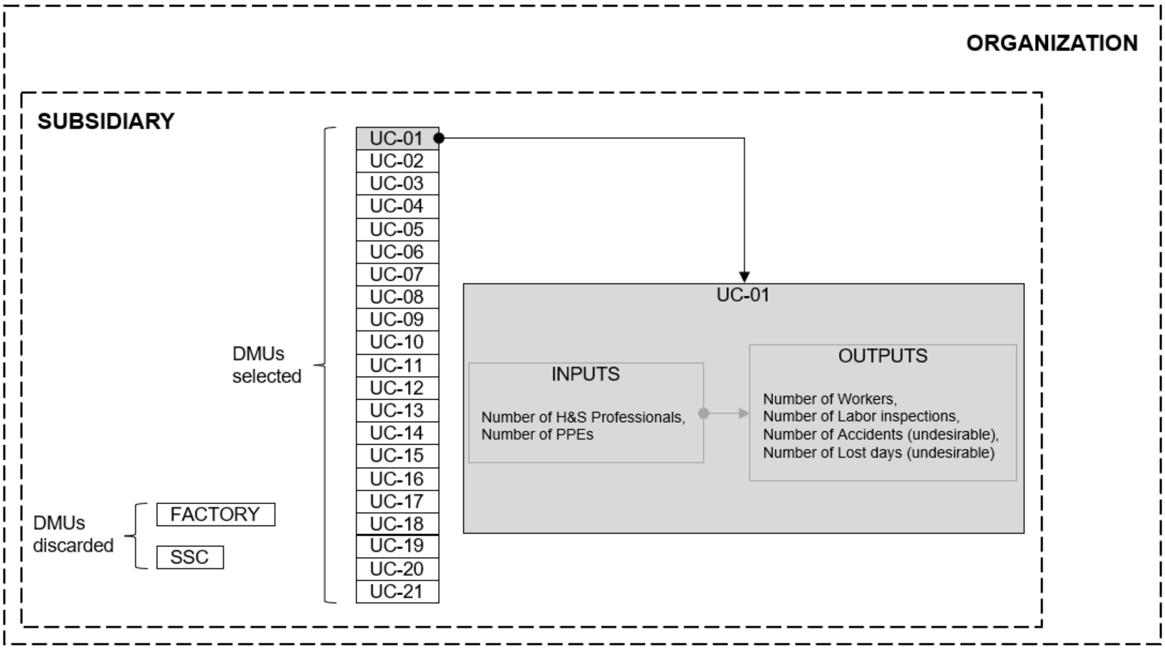


Figure 28 – DMUs in scope in the case-based research

This case-based analysis is therefore designed to evaluate the efficiency of safe work environment (unit of analysis) in the set of 21 homogeneous DMUs, also so-called units of context (UC). The selection of this case took into consideration factors that include: (i) a large organization composed of comparable business units; (ii) ISO-45001 certification to ensure data reliability; (iii) (iv) well-structured H&S database; (iv) access to data related to at least 3 consecutive periods, and homogeneity in which internal benchmarking might be applied. All requirements were met, and a period of 4 years (2018-2021) was defined as the scope of analysis. Next section outlines the proposed model to measure the efficiency scores of the DMUs, as well as explore how it connects to the Decent Work Agenda.

5.3.1.1 Selection of variables and data collection

The variables selected for the model stick to the DWFM (Table 24). All indicators required in the framework were considered outputs, and the inputs were defined by the authors considering the fundamental elements of OHSMS, which include H&S personnel and the number of personal protective equipment delivered to workers. Also, the output (Y_{g1}) was included in the model because any safety management system should consider the number of workers as a fundamental variable. Thus, Table 25 summarizes the variables used in the model, their type and definition, as well as their correspondence with DWFM.

Table 25 – Input and Outputs considered in the model

Variable	Type	Definition
X_1 X_2	Input	Quantity of H&S professionals Quantity of personal protective equipment delivered
Y_{g1} Y_{g2}	Output (desirable)	Quantity of workers in scope Quantity of labor inspections conducted
Y_{b1} Y_{b2}	Output (undesirable)	Quantity of non-fatal work-related accidents Quantity of lost days due to work-related injuries

To achieve a reasonable level of discrimination, the number of inputs and outputs is consistent with the reasoning proposed by Dyson et al. (2001). Also, note some differences in how we define the variables (X, Y) based on the DWFM original indicators. For instance, rather than using the non-fatal accidents rate (SAFE2), we use the quantity of non-fatal accidents. This is intended to avoid 'falling into some DEA pitfalls' (Dyson et al., 2001), such as mixing indices and volume measures.

We collected the database directly from the management systems used in the organization. Data directly associated with H&S (i.e. accidents, lost days, and labor inspections) are well-structured in a cloud-based system and displayed in dashboards we accessed on the organization's Power B.I. workspace. Other variables such as personal protective equipment delivered per period of analysis were extracted from SAP[8] and provided by the Logistics area. Similarly, data associated with personnel was provided by Human Resources based on SAP reports. In addition, we also collected data concerning the goals established by the organization throughout the analysis. This data is necessary for the efficacy assessment and further cross-analysis between efficiency and efficacy scores. The complete data set of all input and output variables for each DMU in the period of analysis (2018-2021) is displayed in appendix A 9.

5.3.2 Model configuration

In this work, we decided to deal with undesirable outputs in their original form. That means no data transformation was carried out, such as the multiplicative inverse or any other data manipulation data technique used in indirect approaches (see (Halkos and Petrou, 2019)). Benefits from this decision include: (i) it preserves interval scale; (ii) no further transformation is required at the phase of analysis of results; (iii) the model is easily comprehended; (iv) more suitable for real-world application.

Following this reasoning, instead of using the standard DEA with data manipulation to deal with undesirable outputs, we use Directional Distance Function, which also resorts to linear programming. Thus, formulation (3) displays the standard form of a DDF model. It aims at maximizing the potential of improvement value of β for each DMU, leading to production on the efficient frontier of the system (Oliveira et al., 2017).

$$Max\ \beta_k \quad (3)$$

$$S.t.$$

$$\sum_{j=1}^{n} \lambda_j y_{gj} \geq y_{gk} + \beta_k |g_{y_g}| \quad g = 1, \dots, r \quad (3.1)$$

$$\sum_{j=1}^{n} \lambda_j y_{bj} \leq y_{bk} - \beta_k |g_{y_b}| \quad b = 1, \dots, s \quad (3.2)$$

[8] SAP stands for Systems Applications and Products in Data Processing. SAP, by definition, is also the name of the ERP (Enterprise Resource Planning) software as well as the name of the company.

$$\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{ik} - \beta_k |g_{x_i}| \quad i = 1, \dots, m \tag{3.3}$$

$$\lambda_j \geq 0 \quad j = 1, \dots, n \tag{3.4}$$

In model (3), the factor β_k indicates the extent of the DMU's inefficiency. It corresponds to the maximal feasible contraction of inputs x_{ij} $(i = 1, \dots, m)$, and undesirable outputs y_{bj} $(b = 1, \dots, s)$ and expansion of outputs y_{gj} $(g = 1, \dots, s)$ that can be achieved simultaneously. Therefore, DMU k is radially efficient when $\beta_k = 0$. Similarly to the DEA models, λ_j are the intensity variables representing the convex combination of the peers. In addition, y_{gk}, y_{bk} and x_{ik} are, respectively, the desirable outputs, undesirable outputs, and the inputs observed in DMU k under assessment. For each DMU k, the DDF model aims at identifying the optimum value β_k^*, which is interpreted as the distance from the frontier, or the 'potential of improvement' for k (also called the inefficient score). Positive values for β_k^* are associated with inefficient DMUs. A value of β_k^* equal to zero means that DMU k is at the frontier, indicating it is relatively efficient.

The Direction Distance Function $D^{\rightarrow}$ was previously introduced in Formulation (2). The components of the direction vector $g(-x, -y_b, y_g)$ indicate the direction of the change in the outputs and inputs. The positive component in g_{y_g} indicate expansion of desirable outputs and the negative components $-g_{y_b}$ and $-g_x$ indicates the contraction of undesirable outputs and inputs.

The directional vector g can assume a variety of configurations, e.g., $g(0, -1,1)$ or $g(-1, -1,1)$. This configuration defines how an inefficient DMU can reach the frontier, and the impacts on the reduction of inputs and expansion of outputs. Details regarding the most common configurations and interpretations are found in the works of Oliveira et al. (Oliveira et al., 2020, 2017).

One well-accepted configuration in academia is to consider the direction vector composed of values equal to the current inputs and outputs, i.e. $g(-x, -y_b, y_g)$. It has the advantage of allowing the interpretation of β_k^* in terms of the proportional improvements to inputs and outputs required for DMU k to achieve the frontier of the technology. However, if the decision maker desires to prioritize efforts to enhance the company's efficiency by searching for improvements in specific input and output dimensions, other directional vectors can be specified under careful analysis.

In this work, we opted for the use of the DDF model in constant returns to scale (CRS) with an output-oriented. This configuration was used considering the assumption of homogeneity among DMUs since all are part of one organization, and because it is suitable for internal benchmarking (Piran et al., 2020). Also, the model was applied for each period of analysis, i.e. 2018, 2019, 2020, and 2021. Thus, the extent of the DMUs' inefficiency can be verified in the context of each year, and it is consistent with practical management review within the selected case. Results were obtained using a computation script in R (R core team, 2018), based on efficiency analysis functions embedded in the packages *'nonparaeff'* and *'Benchmarking'* (Bogetoft et al., 2022; Oh and Suh, 2022).

Finally, we introduce an additional analysis based on the efficacy of achieving the management goals established by the organization. Results were organized in correspondence with the efficiency scores, and different perspectives are discussed based on the combination of efficiency and efficacy.

5.4 Results

5.4.1 Efficiency scores

The so-called inefficiency score β_k^* represents the potential for improvement of one inefficient DMU k. However, the inefficiency scores also can be used to calculate the efficiency based on the relationship between the DEA efficiency score $(\emptyset)$ and the DDF inefficiency score (β) in the form of $\emptyset_k = 1/(1+\beta_k^*)$ for output-oriented models (see appendix A 16). An efficient DMU k is therefore the one in which $\beta_k^* = 0$, and thus, $\emptyset_k = 1$.

Note, e.g., that in 2018 there were 11 efficient branches (DMUs) in the organization $(\beta_k^* = 0)$. It represents that 52% of the branches were performing at the frontier of efficiency. When considering the portion of efficient DMUs throughout 4 years of analysis, 49% of the branches performed efficiently on average. For those operations, the model considers them as benchmarking relative to others in the sample, with no scope for improvement. Table 26 consolidates the results year on year of the efficiency assessment considering the set of DMUs analyzed in the period covering 2018-2021, obtained based on formulation (3) as reasoned in the previous section.

Table 26 – Potential for improvements (β_k^*) and efficiency ($\emptyset_k$) year on year

DMU	2018		2019		2020		2021		*Mean*	*Median*
	β_k^*	$\emptyset_k$	β_k^*	$\emptyset_k$	β_k^*	$\emptyset_k$	β_k^*	$\emptyset_k$	$\overline{\emptyset_k}$	$\widetilde{\emptyset_k}$
UC-01	0.015	0.984	0	1	0.241	0.806	0	1	0.948	0.992
UC-02	0	1	0	1	0	1	0	1	1	1
UC-03	0.396	0.604	0.738	0.575	1	0.500	0.296	0.788	0.617	0.590
UC-04	0.264	0.736	0.821	0.549	0.907	0.524	0.293	0.773	0.646	0.643
UC-05	0.067	0.932	0.292	0.774	1	0.500	0.311	0.762	0.742	0.768
UC-06	0.752	0.248	0.567	0.638	1.735	0.365	0.821	0.549	0.450	0.457
UC-07	0	1	0	1	0.352	0.739	0	1	0.935	1
UC-08	0.183	0.816	0.291	0.774	1.933	0.341	0.433	0.698	0.657	0.736
UC-09	0.883	0.116	1.915	0.343	1.281	0.438	0.489	0.671	0.392	0.391
UC-10	0.105	0.895	0.439	0.695	1.141	0.467	1	0.500	0.639	0.598
UC-11	0	1	0.105	0.905	0	1	0	1	0.976	1
UC-12	0	1	0	1	0.882	0.531	0.501	0.666	0.799	0.833
UC-13	0.339	0.661	0.489	0.671	0.882	0.531	0	1	0.716	0.666
UC-14	0	1	0	1	0	1	0	1	1	1
UC-15	0	1	0	1	0	1	0	1	1	1
UC-16	0	1	0	1	0	1	0	1	1	1
UC-17	0	1	0	1	0	1	0	1	1	1
UC-18	0.138	0.878	0.403	0.712	0	1	0	1	0.898	0.939
UC-19	0	1	0	1	0	1	0	0.861	0.965	1
UC-20	0	1	0.342	0.745	0.718	0.582	0	1	0.832	0.873
UC-21	0	1	0	1	0.096	0.912	0	1	0.978	1
Mean	-	0.851	-	0.828	-	0.726	-	0.870	**0.819**	**-**
Median	-	1	-	0.905	-	0.739	-	1	**-**	**0.958** [b]

[a, b] calculated based on all values included in the data set.

Along with the analyzed period (year on year), some variations in the heterogeneity of the efficiency scores are verified, as illustrated in Figure 29. The boxplot demonstrates a significant variability concerning the efficiency scores in 2020, but a relatively small variation in 2018 and 2021. Note the median of the efficiency scores decreased in 2019 and 2020, but recovered in 2021 to a similar pattern as in 2018. A possible reason to explain this scenario is that the organization was classified as an essential service enterprise during the COVID-19 pandemic started at the beginning of 2020, and it was necessary to expand the use of resources to keep the operations running even during the most critical periods of lockdowns (Gomes et al., 2021). However, different scenarios were faced among the states in Brazil concerning the level of lockdowns. This explains the increase in heterogeneity in 2020. In addition, except in 2020, the organization performed with a quite high third quartile, with more than 75% of its branches being close to the efficiency frontier.

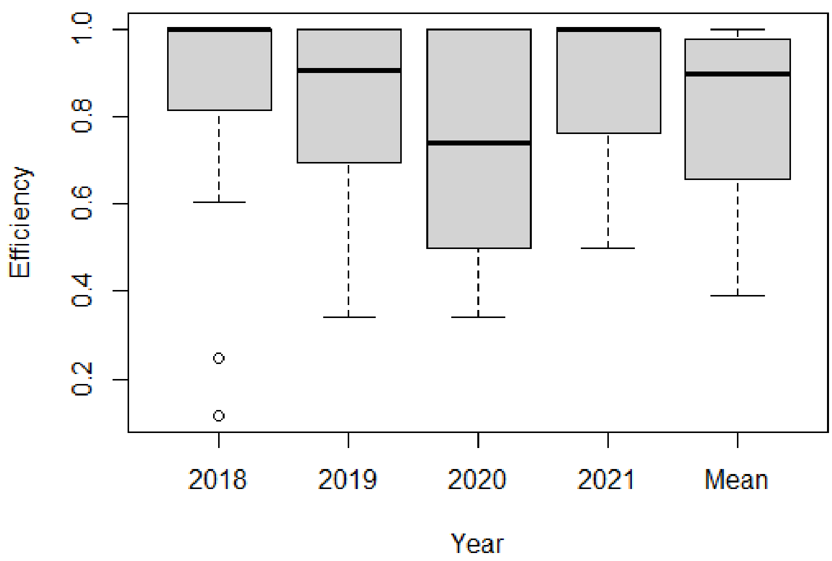


Figure 29 – Patterns of variability in efficiency scores

The overall efficiency of the organization is represented in this model by the arithmetic mean of the efficiency scores of its DMUs ($\overline{\emptyset_j}$). Thus, the organization performed 82% efficiently on average in the period of analysis, i.e. 85.1% in 2018, 82.8% in 2019, 72.6% in 2020, and 87% in 2021. This is a result of a significant portion of DMUs performing close to or at the efficiency frontier (see appendix A 10). A DMU k at the efficiency frontier is one that $\emptyset_k = 100\%$. In other words, it is the DMU in which a maximal feasible contraction of inputs and undesirable outputs, and expansion of outputs is reached, and the performance of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs.

From the perspective of the median, the analysis of the data set (2018-2021) also confirms the impact of the pandemic on efficiency in 2020, which is consistent with the heterogeneity observed in Figure 29. In general terms, the overall median equal to 95.8% indicates the organization is operating either very close or at the efficiency frontier for over 50% of observed data, therefore, with opportunities to improve safety in its less efficient work environment.

Our results also show that DMUs with higher efficiency scores were mostly associated with reduced undesirable outputs, i.e. non-fatal work-related accidents and lost days equal to or close to zero (see appendices A 9 and A 10). Thus, on the one side, based on the variables defined in the DWMF and considered in this model, the efficient branches are supposed to be the ones with safer work environments. On the other side, since weight restrictions were not applied in our model, the results should be carefully analyzed because the variables, including undesirable outputs (e.g. fatal and non-fatal accidents), were considered equally relevant.

In contrast with the efficient DMUs, the set of inefficient ones are those in which the opportunities for improvements are represented by the score β_k^* (see Table 26). Therefore based on β_k^*, the targets for each variable are defined to give direction on how a DMU might become efficient. In general terms, the model indicates how to achieve the efficiency frontier by expanding desirable outputs, and reducing undesirable outputs (see data for 2021 in Table 27).

Table 27 – Targets for outputs of inefficient DMUs (Period 2021)

DMU	Desirable outputs				Undesirable outputs			
	Y_{g1}		Y_{g2}		Y_{b1}		Y_{b2}	
	Observed Y_{g1k}	Target $Y_{g1k}.t$	Observed Y_{g2k}	Target $Y_{g2k}.t$	Observed Y_{b1k}	Target $Y_{b1k}.t$	Observed Y_{b2k}	Target $Y_{b2k}.t$
UC-03	193	245	604	767	0	0	0	0
UC-04	184	238	499	645	1	0.70	14	9.89
UC-05	76	99	190	249	0	0	0	0
UC-06	61	111	179	326	0	0	0	0
UC-08	101	145	193	277	0	0	0	0
UC-09	47	70	254	378	0	0	0	0
UC-10	91	182	356	712	1	0	0	0
UC-12	162	243	514	772	1	0.50	15	7.48
UC-19	926	1,075	2,348	2,725	3	2.5	249	209

The targets for the outputs presented in Table 27 are explained as follows: for DMU UC-10 to become efficient, it should: (i) expand the quantity of workers in scope from 926 to 1,075, (ii) double the quantity of labor inspections from 356 to 712, (iii) achieve zero non-fatal work-related accidents, and finally (iv) maintain its zero lost days. The same reasoning is applied to the inputs contraction, and the results for all periods are shown in appendices A 11 - A 14. Note that DDF is beneficial for reality-based applications in the field of H&S since no data manipulation is required to deal with desirable and undesirable outputs simultaneously. Also, by using original data, the interpretation of the results is clear and does not require any further manipulation. This is particularly important for managerial application, and it also preserves the discrimination power of the technique.

5.4.2 Efficacy scores

As introduced in section 5.2, efficacy is a commonly used performance measure associated with setting and achieving goals (see Table 28). Thus, let's consider δ_k^t the efficacy score of a DMU k in a given period t. The efficacy might be composed of one or more indicators. In this real-world case study, two indicators and their respective goals were defined by the top management of the organization: (1) zero fatalities ($F_{acc} = 0$) and (2) non-fatal accident rate lower or equal to a given reference ($NF_{accR} \leq REF.$).

Table 28 – Management goals

Indicator	Efficacy score	range	*REF.* Goal per period (t)			
			2018	2019	2020	2021
1 - Fatal accident rate	$\delta_{k_{(1)}}^t$	{0;1}	0	0	0	0
2 - Non-fatal accident rate	$\delta_{k_{(2)}}^t$	{0-1}	≤4.00	≤ 3.85	≤ 3.00	≤ 2.14

The given reference for 'goal 2' is defined in management reviews and based on a global benchmark the organization is interested to achieve year on year. Table 29 summarizes the efficacy scores for each management goal in the period of analysis.

Results related to 'goal 1' are binary (0 – non achieved or 1 – achieved). Thus, in case of an occurrence of a fatal accident within period t, the efficacy score $\delta_{k_{(1)}}^t = 0$. In turn, whether a DMU does not report any fatal accidents within period t, $\delta_{k_{(1)}}^t = 1$.

Concerning 'goal 2', efficacy scores can assume any value between 0 and 1, in which 1 means non-fatal accident rate is lower or equal to the managerial goal. In the circumstance of a DMU achieving non-fatal accident rate (NF_{accR}) higher than the given reference, its efficacy score $\delta_{k_{(2)}}^t$ is always lower than 1, and calculated in the form of $\delta_{k_{(2)}}^t = \frac{NF_{accR} - REF.}{REF}$.

Table 29 – Efficacy scores

DMU k	$t = 2018$		$t+1 = 2019$		$t+2 = 2020$		$t+3 = 2021$	
	$\delta_{k_{(1)}}^t$	$\delta_{k_{(2)}}^t$	$\delta_{k_{(1)}}^{t+1}$	$\delta_{k_{(2)}}^{t+1}$	$\delta_{k_{(1)}}^{t+2}$	$\delta_{k_{(2)}}^{t+2}$	$\delta_{k_{(1)}}^{t+3}$	$\delta_{k_{(2)}}^{t+3}$
UC-01	1	1	1	1	1	1	1	1
UC-02	1	1	0	0.18	1	1	1	1
UC-03	1	1	1	1	1	1	1	1
UC-04	1	1	1	1	1	0.15	1	0.68
UC-05	1	1	1	1	1	0.10	1	1
UC-06	1	0.05	1	1	1	1	1	1
UC-07	1	1	1	1	1	1	1	1
UC-08	1	1	1	1	1	1	1	1

UC-09	1	1	1	1	1	1	1	1
UC-10	1	1	1	1	1	1	1	1
UC-11	1	1	1	0.96	1	1	1	1
UC-12	1	0	1	0.44	1	1	1	0.67
UC-13	1	0	1	0.70	1	1	1	1
UC-14	1	0.75	1	1	1	1	1	1
UC-15	1	1	1	1	1	1	1	1
UC-16	1	1	1	0.05	1	1	1	1
UC-17	1	1	1	1	1	0.72	1	1
UC-18	1	1	1	0.96	1	0.75	1	0.25
UC-19	0	0.50	1	1	1	1	1	1
UC-20	1	1	1	0.05	1	1	1	1
UC-21	1	0.75	1	0.18	1	1	1	0.05

The selected indicators and goals demonstrate to what extent the top management is committed to safety. This represents an advantage of measuring efficacy since it pushes managers to achieve the goals defined by the organization. For instance, by setting 'goal 1', the organization displays its not acceptance of fatal accidents. This is verified through the goal of zero fatalities throughout all periods under assessment. Note that DMU UC-19 in 2018, and DMU UC-02 in 2019 reported a fatal accident, and therefore, its efficacy score δ_1^t is equal to zero. Also, by setting a decreased non-fatal accident rate year on year ('goal 2'), the organization demonstrates to seek continuous safety improvements.

In this reasoning, efficacy represents a fundamental mechanism to ensure the vision of the organization concerning safety, and the aspiration of their managers to reach significant improvements in reducing accidents and eliminating fatalities. When used in a combined assessment with efficiency, benefits are added to the decision-making process. Figure 30 summarizes graphically how efficacy and efficiency can jointly offer a better assessment of safe work environment.

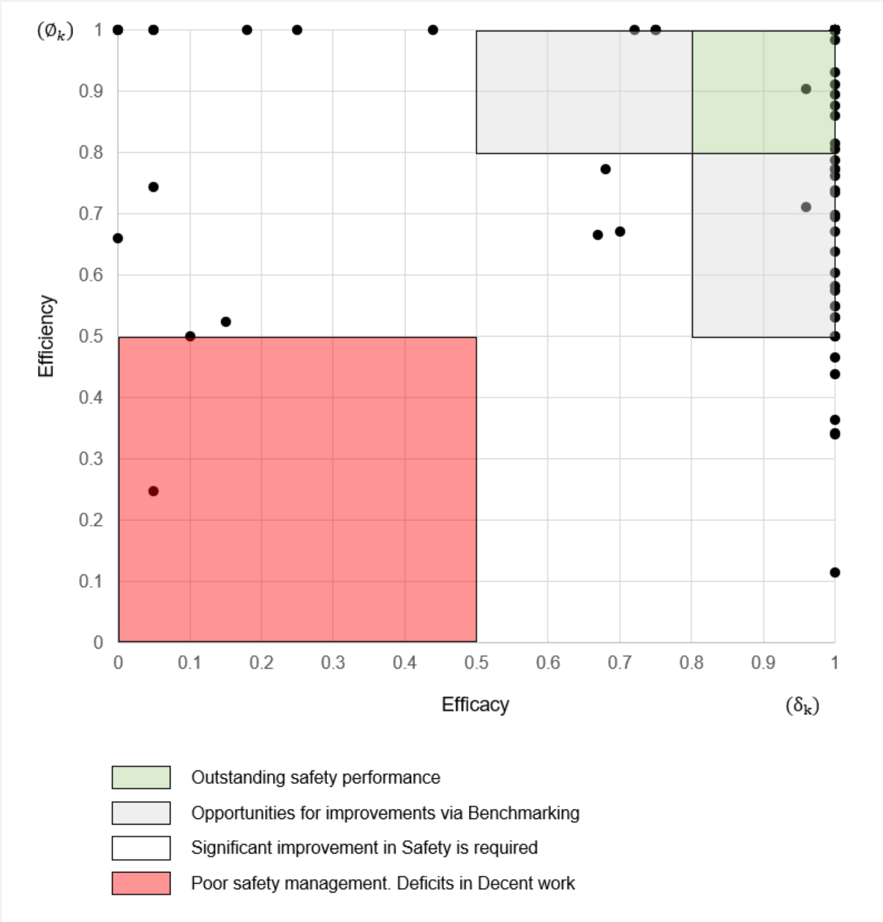


Figure 30 – Safe work environment assessment based on Efficiency and Efficacy

Different perspectives are verified in Figure 30. It is composed of data retrieved from 21 DMUs over 4 years and results are grouped into four distinct levels of performance. One is represented by a light green area, in which both efficacy and efficiency scores are higher than 80%. It represents that DMUs are operating in higher safety performance, ensuring both goals and resources are well-managed and desirable outcomes are expanded. Second, in light grey, there are DMUs with good opportunities for benchmarking. This is because they are either efficient but slightly behind their goals, or the other way around, i.e. DMUs are reaching goals but not efficiently. Third is represented in white, in which a significant improvement is required. For those DMUs, it is argued that weak signals should be deeply investigated to identify key factors that explain lower performance relative to others. Finally, poor safety management and deficits in decent work are represented in the light red, as a result of combined efficacy and efficiency scores lower than 50%. This scenario requires urgent management attention since it offers fertile ground for work-related accidents and illnesses.

5.5 Discussion

Since the first publication of the decent work indicators manual in 2012, researchers and practitioners were given a guideline of statistical and legal indicators for measuring decent work (ILO, 2013). Initially, studies explored decent work from a wider perspective, e.g. to quantify ILO objectives and identify policies conducive to decent work in a globalized environment (W W Cooper et al., 2011; Ertugrul Karsak and Goker, 2019; Thore and Tarverdyan, 2009). In these works, DEA was used to identify the efficient and inefficient countries based on major factors such as employment, rights, and social protection. However, studies focused on measuring decent work grounded in the observations of the existing practice within organizations are scarce.

To close this gap, we explored a reality-based case to measure the efficiency of a safe work environment using DDF. Differently from other standard DEA applications dealing with undesirable outputs, the proposed model sought simultaneously to maximize the use of resources, expand desirable outcomes and reduce undesirable ones (e.g. accidents) with no data manipulation. As a result, the model identified inefficiencies in a set of DMUs under assessment, as well as estimated targets for improving the efficiency of each DMU. The sort of DMUs into efficient and inefficient ones is an important outcome of the use of this method. However, as pointed out by Oliveira, Camanho, and Zanella (2017), this classification requires caution since it is only a relative measure, which depends on the sample underlying the performance comparison.

In this research, the score $\emptyset_k$ represents how efficient one DMU k is to promote a safe work environment relative to the set of DMUs in scope. Thus, the measurement of decent work of an organization is represented by a set of its J DMUs in a time series, and its overall efficiency score was calculated based on the arithmetic mean of the efficiency scores of its DMUs.

For the organization under assessment, in 2018, 11 out of 21 DMUs were considered efficient (52.38%), and the overall efficiency score $\overline{\emptyset_J}$ was 85.1%. In 2019-2020, the number of efficient DMUs decreased to 10 and 8, respectively. However, in 2021, the organization recovered performance and achieved its highest overall efficiency score with 12 DMUs at the frontier ($\overline{\emptyset_J} = 87\%$).

Some factors composing the DDF model explain the variation in these efficiency scores. One is the reduction of desirable outputs, e.g. labor inspections (Y_{g2}). This is well illustrated in 2021 when it verified a significant reduction in labor inspections due to restrictions imposed by the COVID-19 pandemic. Another is either the increase in the use of resources such as PPE (X_2) or the increase of undesirable outputs, e.g. lost days (Y_{b2}) (see aapendix A 9). These variables correspond with the indicators proposed by ILO in the decent work measurement framework and their variation influence the efficiency scores obtained through the model formulation (3).

From the perspective of the inefficient DMUs, the potential for improvements is represented by their inefficiency scores (β_k^*). It represents the distance from a DMU k to the efficiency frontier. Thus, the proposed DDF model considers β_k^* to estimate targets for each variable and it gives directions for benchmarking. The estimated targets are references of each variable to lead a DMU to the efficiency frontier. As DDF can deal with desirable and undesirable outputs simultaneously, the estimated targets seek to expand desirable outputs and reduce inputs and undesirable outputs.

However, since efficiency is a relative measure, a careful analysis is required. If a set of DMUs under assessment is characterized by poor safety management, the frontier will be composed of underperformance references (*"the best among the worst"*). Therefore, for this scenario, the efficiency frontier might be useful only for incremental improvements, but not enough to promote significant changes in decent work. Another scenario to be carefully considered is the one in which a DMU with the occurrence of a critical undesirable output (e.g. fatal accident) is considered efficient. This is the case for UC-19 in the period 2018 and UC-02 in 2019 (see appendix A 10). For these circumstances, the comprehension of efficiency as a relative measure is even more critical.

As previously mentioned, an efficient DMU is not one in which undesirable outputs do not take place (e.g. zero accidents, zero lost days, etc). Rather, it is one in which the resources used and undesirable outputs are minimized, and desirable outcomes obtained for promoting a safe work environment have been maximized.

Note, by definition, that the use of efficiency as a performance measure of decent work is very useful for studies interested in examining performance comparison. However, we argue it is necessary but not sufficient when it comes to organizational levels. Our results indicate that measuring efficacy is also critical for promoting decent work. This argument is reasoned by the assumption that managerial initiatives are goal-driven. In addition, it considers that organizations set up goals as a mechanism to engage managers toward their values and strategic objectives.

In this reality-based case, the organization set two goals: (1) zero fatalities and (2) a non-fatal accident rate lower or equal to a given reference year on year (Table 28). To illustrate the discussion regarding efficacy measurement, consider the efficient DMU $k =$ UC-02 in period $t = 2019$. It resulted in efficacy scores as follows: $\delta^t_{k_{(1)}} = 0$ (due to one fatal accident, i.e. 0% of 'goal 1'), and $\delta^t_{k_{(2)}} = 0.18$, which means it achieved 18% of its second goal. Therefore, even being efficient ($\emptyset^t_k = 1$), the efficacy of UC-02 was zero in t (i.e. $\delta^t_k = \delta^t_{k_{(1)}} . \delta^t_{k_{(2)}} = 0$).

This scenario illustrates a wider aspect of two different performance measurements: DMUs might be efficient (as a relative measure) without achieving the goals defined by the organization. From the perspective of decent work, it might represent an undesirable scenario as previously reasoned. A possible solution to address this issue is to consider the effectiveness (ϵ_k) as a composite measure in which efficiency and efficacy are combined in the suggested formulation:

$$\epsilon_k = \emptyset_k \, . \, \delta_k \qquad (4)$$

Thus, efficiency is sought to maximize productivity and management goals are given to drive management attention to critical factors (e.g. achieving zero fatal accidents). As a combined result of these metrics with no weight restriction, the more effective one DMU k is (higher ϵ_k), the safer the work environment.

Figure 31 illustrates how effectiveness could be used as a wider measure to evaluate safe work environment from the perspective of decent work. The effectiveness graphically indicates how a DMU k is positioned as a result of its relative efficiency and its efficacy. As a composite measure, we argue that the higher the effectiveness, the safer the work environment. For instance, let's take the top performers ($\epsilon_k \geq 80\%$): DMUs UC-01, UC-07, UC-11, UC-14, UC-15, and UC-17. These operations presented high-efficiency scores close to 100%, being benchmarks to others. In addition, they mostly reached management goals as presented in the appendix A 15. The same reasoning is considered to explain other layers in Figure 31, e.g. the DMUs UC-06, UC-09, and UC-12 facing poor effectiveness ($\epsilon_k \leq 40\%$). Note that poor effectiveness might be associated with either very low efficiency or efficacy. This is well-represented by UC-06 in which efficiency and efficacy scores varied throughout 2018-2021 in low performance (see A 15).

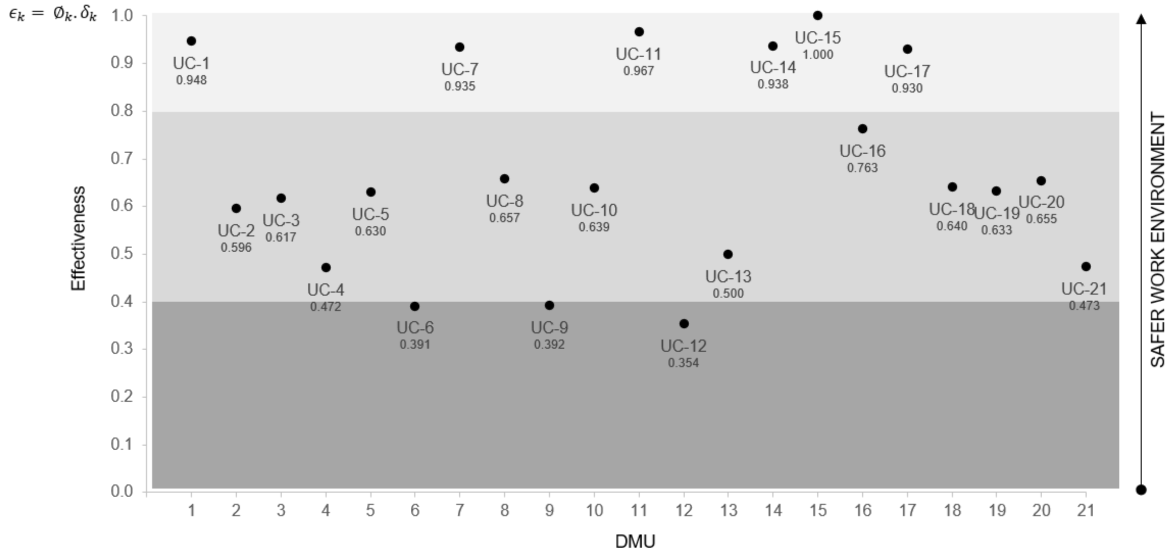


Figure 31 – Effectiveness (mean): 2018-2021

From the perspective of practitioners, DMUs are supposed to be efficient and reach management goals. In this context, effectiveness can fix typical distortion in performance evaluation since it considers both dimensions. Appendix A 15 shows how to use formulation (4) was used to calculate the effectiveness of the set of DMUs presented in Figure 31.

Results from this modeling are easily interpreted. This is relevant to ensure the practicability of the proposed model in reality-based cases, notably within organizations. In addition, the discussion about the composite measure combining efficiency and efficacy opens an avenue for managerial applications and future research in socio-economic sciences, including modeling advances.

5.6 Conclusion

This study proposed a model based on the directional distance function (DDF) for measuring the efficiency of a safe work environment from the perspective of decent work agenda. To illustrate the application of the model in a reality-based case, we conducted a longitudinal study in a multinational organization over 4 years (2018-2021).

This study proposed a model based on the directional distance function (DDF) for measuring the efficiency of a safe work environment from the perspective of decent work agenda. To illustrate the application of the model in a reality-based case, we conducted a longitudinal study in a multinational organization over 4 years (2018-2021).

Our results indicate that DMUs with higher efficiency scores were mostly associated with reduced undesirable outputs, i.e. work-related accidents and lost days equal to or close to zero. In contrast, inefficient branches presented expanded undesirable outputs or excessive use of resources relative to others. However, our results also suggest caution in considering only the efficiency for measuring safe work environment. Based on its relative nature, the results might seem inconsistent from a safety management perspective when DMUs in scope are homogeneous in their undesirable outputs.

The research presented contributions in both theoretical and managerial fields. One is the novelty approach to using the DDF model based on the statistical and legal indicators proposed in the framework on the measurement of decent work published by ILO in 2012. This approach is especially suitable for social-economic and environmental applications because it does not require any data transformation to deal with undesirable factors, and also preserves the discrimination power of the model.

Second, we measured efficiency in a real case composed of 21 DMUs over 4 years. Managerial benefits were verified in identifying internal benchmarks, as well as the potential for improvements to be considered by the inefficient units. In addition, this is the first empirical application of measuring the efficiency of safe work environment (a substantive element of decent work) using DDF in a time series.

Finally, we postulate that a composite measure combining efficiency and efficacy herewith called effectiveness is necessary for a more precise evaluation of decent work. Thus, we argue that the higher the effectiveness, the safer the work environment.

Future research is indicated to refine the indicators used in the model, rather than only stick to those proposed in the DWMF. For instance, variables may include leadership engagement, workforce qualification, and economic investment in H&S. In addition, further theoretical and empirical studies in applying composite measures are needed for evaluating safe work environment at an organizational level, expanding the discussions presented in this research.

APPENDICES CHAPTER 5

A 9. Data collected: Inputs (X) and outputs (Y)

DMU	Period																							
	2018						2019						2020						2021					
	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}
UC-01	1	458	32	310	0	0	1	447	27	387	0	0	1	447	49	401	0	0	1	269	43	186	0	0
UC-02	2	2,293	143	1,096	1	4	2	2,059	154	1,043	3	6,196[b]	2	1,731	226	1,230	1	209	2	1,715	226	885	1	0
UC-03	3	3,214	150	892	1	9	3	2,910	129	849	1	7	2	2,783	186	624	2	3	2	5,035	193	604	0	0
UC-04	2	3,031	123	782	1	83	2	3,626	94	790	1	53	2	2,191	175	681	2	21	2	1,773	184	499	1	14
UC-05	1	1,137	59	313	0	0	1	1,085	58	298	0	0	1	1,452	79	267	1	9	1	973	76	190	0	0
UC-06	1	1,212	47	247	2	40	1	1,442	53	201	0	0	1	1,067	59	196	0	0	1	1,136	61	179	0	0
UC-07	2	3,607	188	1,407	2	15	2	3,366	177	1,417	0	0	2	2,736	276	1,407	1	13	2	2,968	273	916	2	9
UC-08	1	1,836	78	579	1	68	1	1,454	64	502	0	0	2	1,830	100	200	0	0	2	1,165	101	193	0	0
UC-09	1	734	28	241	2	114	1	1,246	20	204	0	0	1	777	46	305	0	0	1	1,041	47	254	0	0
UC-10	1	926	57	204	0	0	1	698	46	303	0	0	2	1,138	94	533	0	0	2	2,194	91	356	1	0
UC-11	4	1,274	163	975	3	40	4	1,205	149	928	2	90	3	1,260	205	1,366	1	8	3	1,811	204	1,054	1	7
UC-12	2	2,308	118	842	3	23	2	2,137	120	800	2	6	2	1,815	161	637	1	15	2	2,111	162	514	1	15
UC-13	1	1,659	76	488	2	48	1	2,044	70	464	1	79	1	1,726	95	626	0	0	1	1,393	97	466	0	0
UC-14	3	2,504	280	1,320	4	50	3	1,985	270	1,257	3	35	2	1,716	377	1,517	1	15	2	2,651	393	959	1	185
UC-15	1	512	63	352	0	0	1	291	57	340	0	0	1	308	88	414	0	0	2	325	86	233	0	0
UC-16	1	2,243	119	1,281	2	31	1	2,622	128	1,220	4	35	1	1,580	211	1,178	0	0	1	2,440	210	820	1	0
UC-17	2	2,269	156	1,502	2	59	2	2,139	169	1,430	1	3	2	2,130	249	1,441	2	139	1	2,194	232	1,483	1	290
UC-18	2	3,360	153	1,318	2	39	2	4,220	162	1,255	6	110	2	4,116	255	1,308	2	8	2	3,917	496	1,053	2	19
UC-19	5	8,251	669	2,857	11	6,287[a]	5	7,598	625	3,029	4	79	5	9,743	821	2,504	5	3,068	6	12,677	926	2,348	3	249
UC-20	2	2,870	171	1,120	1	15	2	2,515	143	1,065	4	128	2	2,619	224	1,183	1	29	2	2,599	245	849	0	0
UC-21	3	1,195	136	1,035	2	178	2	1384	156	985	3	58	2	1,771	198	633	1	3	2	1,919	190	458	2	15

[a, b] according to standard ABNT NBR 14280:2001 (ABNT, 2001), 6.000 lost days associated with a fatal accident should be considered for statistical purposes.

A 10. Efficiency scores

DMU	2018 $\emptyset_k$ (DDF)	2019 $\emptyset_k$ (DDF)	2020 $\emptyset_k$ (DDF)	2021 $\emptyset_k$ (DDF)
UC-01	0.984	1	0.806	1
UC-02	1	1	1	1
UC-03	0.604	0.575	0.500	0.788
UC-04	0.736	0.549	0.524	0.773
UC-05	0.932	0.774	0.500	0.762
UC-06	0.248	0.638	0.365	0.549
UC-07	1	1	0.739	1
UC-08	0.816	0.774	0.341	0.698
UC-09	0.116	0.343	0.438	0.671
UC-10	0.895	0.695	0.467	0.500
UC-11	1	0.905	1	1
UC-12	1	1	0.531	0.666
UC-13	0.661	0.671	0.531	1
UC-14	1	1	1	1
UC-15	1	1	1	1
UC-16	1	1	1	1
UC-17	1	1	1	1
UC-18	0.878	0.712	1	1
UC-19	1	1	1	0.861
UC-20	1	0.745	0.582	1
UC-21	1	1	0.912	1
Mean	0.851	0.828	0.726	0.870

A 11. Estimated targets for inefficient DMUs (Period 2018)

DMU	Desirable outputs				Undesirable outputs				Inputs			
	Y_{g1}		Y_{g2}		Y_{b1}		Y_{b2}		X_1		X_2	
	Observed Y_{g1k}	Target $Y_{g1k}.t$	Observed Y_{g2k}	Target $Y_{g2k}.t$	Observed Y_{b1k}	Target $Y_{b1k}.t$	Observed Y_{b2k}	Target $Y_{b2k}.t$	Observed X_{1k}	Target $X_{1k}.t$	Observed X_{2k}	Target $X_{2k}.t$
UC-01	32	33	310	315	0	0	0	0	1	0.98	458	451
UC-03	150	209	892	1,245	1	0.60	9	5.44	3	1.81	3,214	1,941
UC-04	123	156	782	989	1	0.74	83	61.07	2	1.47	3,031	2,230
UC-05	59	63	313	334	0	0	0	0	1	0.93	1,137	1,060
UC-06	47	82	247	433	2	0.50	40	9.91	1	0.25	1,212	300
UC-08	78	92	579	685	1	0.82	68	55.51	1	0.82	1,836	1,499
UC-09	28	53	241	454	2	0.23	114	13.28	1	0.12	734	86
UC-10	57	63	204	225	0	0	0	0	1	0.89	926	829
UC-13	76	102	488	654	2	1.32	48	31.72	1	0.66	1,659	1,096
UC-18	153	174	1,318	1,500	2	1.72	39	33.61	2	1.72	3,360	2,896

A 12. Estimated targets for inefficient DMUs (Period 2019)

DMU	Desirable outputs				Undesirable outputs				Inputs			
	Y_{g1}		Y_{g2}		Y_{b1}		Y_{b2}		X_1		X_2	
	Observed Y_{g1k}	Target $Y_{g1k}.t$	Observed Y_{g2k}	Target $Y_{g2k}.t$	Observed Y_{b1k}	Target $Y_{b1k}.t$	Observed Y_{b2k}	Target $Y_{b2k}.t$	Observed X_{1k}	Target $X_{1k}.t$	Observed X_{2k}	Target $X_{2k}.t$
UC-03	129	224	849	1,475	1	0.26	7	1.83	3	0.79	2,910	763
UC-04	94	171	790	1,439	1	0.18	53	9.47	2	0.36	3,626	648
UC-05	58	75	298	385	0	0	0	0	1	0.71	1,085	768
UC-06	53	83	201	315	0	0	0	0	1	0.43	1,442	625
UC-08	64	83	502	648	0	0	0	0	1	0.71	1,454	1,031
UC-09	20	58	204	595	0	0	0	0	1	-0.92[a]	1,246	-1,141[a]
UC-10	46	66	303	436	0	0	0	0	1	0.56	698	391
UC-11	149	165	928	1,025	2	1.79	90	80.57	4	3.58	1,205	1,079
UC-13	70	104	464	691	1	0.51	79	40.34	1	0.51	2,044	1,044
UC-18	162	227	1,255	1,761	6	3.58	110	65.61	2	1.19	4,220	2,517
UC-20	143	192	1,065	1,429	4	2.63	128	84.22	2	1.32	2,515	1,655

[a] Negative values to be discarded.

A 13. Estimated targets for inefficient DMUs (Period 2020)

DMU	Desirable outputs				Undesirable outputs				Inputs			
	Y_{g1}		Y_{g2}		Y_{b1}		Y_{b2}		X_1		X_2	
	Observed Y_{g1k}	Target $Y_{g1k}.t$	Observed Y_{g2k}	Target $Y_{g2k}.t$	Observed Y_{b1k}	Target $Y_{b1k}.t$	Observed Y_{b2k}	Target $Y_{b2k}.t$	Observed X_{1k}	Target $X_{1k}.t$	Observed X_{2k}	Target $X_{2k}.t$
UC-01	49	61	401	497	0	0	0	0	1	0.76	447	339
UC-03	186	372	624	1248	2	0	3	0	2	0	2,783	0
UC-04	175	334	681	1299	2	0.18	21	1.94	2	0.18	2,191	203
UC-05	79	158	267	534	1	0	9	0	1	0	1,452	0
UC-06	59	161	196	536	0	0	0	0	1	-0.74[a]	1,067	-785 [a]
Uc-07	276	373	1407	1903	1	0.64	13	8.42	2	1.30	2,736	1,772
UC-08	100	293	200	587	0	0	0	0	2	-1.87 [a]	1,830	-1,709 [a]
UC-09	46	105	305	696	0	0	0	0	1	-0.28 [a]	777	-218 [a]
UC-10	94	201	533	1142	0	0	0	0	2	-0.28 [a]	1,138	-161 [a]
UC-12	161	303	637	1199	1	0.12	15	1.76	2	0.24	1,815	213
UC-13	95	179	626	1178	0	0	0	0	1	0.12	1,726	204
UC-20	224	385	1183	2033	1	0.28	29	8.16	2	0.56	2,619	737
UC-21	198	217	633	694	1	0.90	3	2.71	2	1.81	1,771	1,601

[a] Negative values to be discarded.

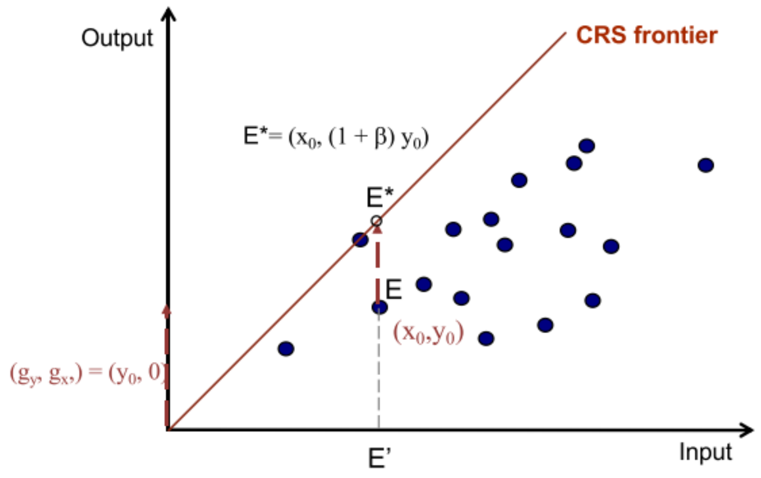
A 14. Estimated targets for inefficient DMUs (Period 2021)

DMU	Desirable outputs				Undesirable outputs				Inputs			
	Y_{g1}		Y_{g2}		Y_{b1}		Y_{b2}		X_1		X_2	
	Observed Y_{g1k}	Target $Y_{g1k}.t$	Observed Y_{g2k}	Target $Y_{g2k}.t$	Observed Y_{b1k}	Target $Y_{b1k}.t$	Observed Y_{b2k}	Target $Y_{b2k}.t$	Observed X_{1k}	Target $X_{1k}.t$	Observed X_{2k}	Target $X_{2k}.t$
UC-03	193	245	604	767	0	0	0	0	3	1.5	5,035	3,678
UC-04	184	238	499	645	1	0.70	14	9.89	2	1.4	1,773	1,253
UC-05	76	99	190	249	0	0	0	0	1	0.7	973	670
UC-06	61	111	179	326	0	0	0	0	1	0.2	1,136	204
UC-08	101	145	193	277	0	0	0	0	2	1.1	1,165	661
UC-09	47	70	254	378	0	0	0	0	1	0.5	1,041	532
UC-10	91	182	356	712	1	0	0	0	2	0	2,194	0
UC-12	162	243	514	772	1	0.50	15	7.48	2	1	2,111	1,053
UC-19	926	1,075	2,348	2,725	3	2.5	249	209	6	5	12,677	10,640

A 15. Effectiveness (mean) for a set of DMUs (period: 2018 - 2021)

DMU k	$t = 2018 - 2021)$				
	Efficacy per goal		Total Efficacy	Efficiency	Effectiveness
	$\overline{\delta_{k_{(1)}}}$	$\overline{\delta_{k_{(2)}}}$	$\overline{\delta_k} = \overline{\delta_{k_{(1)}}}.\overline{\delta_{k_{(2)}}}$	$\overline{\emptyset_k}$	$\overline{\epsilon_k} = \overline{\emptyset_k}.\delta_k$
UC-01	1	1	1	0.948	0.948
UC-02	0.750	0.795	0.596	1	0.596
UC-03	1	1	1	0.617	0.617
UC-04	1	0.708	0.708	0.646	0.472
UC-05	1	0.775	0.775	0.742	0.630
UC-06	1	0.763	0.763	0.450	0.391
UC-07	1	1	1	0.935	0.935
UC-08	1	1	1	0.657	0.657
UC-09	1	1	1	0.392	0.392
UC-10	1	1	1	0.639	0.639
UC-11	1	0.990	0.990	0.976	0.967
UC-12	1	0.528	0.528	0.799	0.354
UC-13	1	0.675	0.675	0.716	0.500
UC-14	1	0.938	0.938	1	0.938
UC-15	1	1	1	1	1.000
UC-16	1	0.763	0.763	1	0.763
UC-17	1	0.930	0.930	1	0.930
UC-18	1	0.740	0.740	0.898	0.640
UC-19	0.750	0.875	0.656	0.965	0.633
UC-20	1	0.763	0.763	0.832	0.655
UC-21	1	0.495	0.495	0.978	0.473

A 16. Relationship between ∅ and β (output-oriented)

Formula	Geometric representation
$\emptyset = \frac{1}{1+\beta}$ $\emptyset = \frac{1}{1+\left(\frac{\overline{EE^{*}}}{\overline{EE^{'}}}\right)}$ $\emptyset = \frac{1}{\left(\frac{\overline{EE^{'}}+\overline{EE^{*}}}{\overline{EE^{'}}}\right)}$ $\emptyset = \frac{1}{\left(\frac{\overline{E^{'}E^{*}}}{\overline{EE^{'}}}\right)}$ $\emptyset = \frac{\overline{E'E}}{\overline{E'E^{*}}}$	

Source: Camanho, A.S. Performance assessment using frontier techniques. Lecture notes. 2018.

6 APPLICATION OF ASSOCIATION RULES TO IDENTIFY PATTERNS OF OCCURRENCE IN WORK-RELATED ACCIDENTS: AN ANALYSIS OF THE EFFECT ON EFFICIENCY[9]

Abstract

Problem definition: Despite greater managerial attention and significant efforts to prevent work-related accidents, the occurrence of work-related injuries is still a challenge for the organization.

Problem analysis: One of the measures typically used to reduce work-related accidents is root causes analysis. This type of analysis encourages continuous improvement cycles, in which, based on the learning generated in the investigation process, measures are established to eliminate or neutralize the critical factors that were identified in the analysis. Several methods are available for this purpose (e.g. Ishikawa, 8D). However, the organization's historical data indicate that this type of analysis is necessary, but not sufficient. This is because potentially hidden factors (or the combined effect of mapped factors) are difficult to identify using traditional techniques. In this context, the use of more robust complementary techniques was considered to solve the problem.

Problem solution: To unhide elements that contribute to the occurrence of accidents, the association rules technique was used through the *Apriori* algorithm proposed by Agrawal et al. (1993; 1994). An association rule is represented by a pattern of type $R_m: \{X\} \rightarrow \{Y\}$, where X represents the antecedent factors and Y is the consequent factor, in this case, the work-related accident. The solution was initially implemented using the RStudio software to analyze the preliminary results and validate the model. Subsequently, a functional architecture was developed to integrate the proposed solution into the existing organization's management system.

Results: From the revealed co-occurrence patterns, it was possible to identify the association of critical factors with the occurrence of accidents. The results confirmed already-expected associations and also revealed counterintuitive aspects that had not been considered until then, such as the significant occurrence of accidents involving more experienced professionals and using Personal Protective Equipment (PPE). Another counterintuitive example revealed is the occurrence of accidents with employees trained in the organization's accident prevention processes, suggesting the need to review the methods applied in internal training. The results also confirm the relevance of technical qualification for accident prevention and behavioral factors to be worked on, such as risk minimization and lack of concentration. Moreover, the use of association rules also resulted in a positive effect on the efficiency even further studies are recommended to confirm causality.

Evaluation and lessons learned: From the results found, some lessons can be highlighted, such as i. technical inspections on the shop floor should expand the field of analysis, not being restricted to the verification of basic factors such as the use of PPE; ii. the technical onboarding process should

[9] A substrate of this article limited to 4.000 words was published as Conference paper in the occasion of the XLII National Conference of Industrial Engineering (ENEGEP 2022). DOI: 10.14488/enegep2022_tce_389_1931_43125. Co-authored with Leandro Gauss and awarded as the Best Business Case article 2022.

intensify the accident prevention module, reassessing its content and teaching-learning techniques; iii. organizational leaders must consider counterintuitive aspects in accident analysis; iv. Conventional techniques for identifying root causes are necessary, but not sufficient, to reveal hidden factors that, individually or in combination, contribute to the occurrence of accidents.

Organization/Case: TK Elevator is one of the world's leading companies in the manufacturing, installation, and maintenance industry of elevators, escalators, and urban mobility equipment. The object of the case used in this article was the Business Unit Latin America, present in 12 countries.

CNAE (National Classification of Economic Activity.): 28.22-4

Keywords: Decent work. Safe work environment. Efficiency. Directional Distance Function.

6.1 Introduction

Occupational health and safety management (OHSM) has become a global issue and solutions to improve its performance have been requested in modern work environments (Wang et al., 2020). This is because, despite increased managerial attention and the progressive improvement of OHSM practices over decades, many challenges remain unresolved, particularly work-related accidents and illnesses (Nicolaidou et al., 2021).

According to International Labour Organization (ILO, 2020), more than 2.8 million deaths and approximately 376.8 million non-fatal accidents occur every year. This uncomfortable scenario suggests that the guidelines proposed in the international agenda (ILO, 1999; Nations, 2015) have not been sufficient for significant progress in the promotion of workers' health and safety and that the OHSM practices adopted in organizations have not been effective in preventing accidents.

Among the measures typically used by organizations to reduce accidents is the analysis of root causes (ISO, 2018). This type of analysis stimulates cycles of continuous improvement, in which, from the learning generated in the investigation process, measures are established for the elimination or neutralization of the determining factors that were identified in the analysis.

Diverse methods are available for this purpose (e.g. Ishikawa, 8D). However, the isolated analysis of accidents does not allow the identification of statistically relevant occurrence patterns. This is because such methods do not consider the historical data set and, therefore, the co-occurrence of antecedent and consequent factors are difficult to identify.

This study considers the use of more robust techniques to identify valid occurrence patterns that enable management actions for accident prevention. Thus, the technique of data mining by association rules proposed by Agrawal *et al.* (1993; 1994) is applied in a practical case where antecedent factors considered likely to co-occur with an accident (consequent factor) are revealed.

The article is structured as follows: section 6.2 addresses the theoretical aspects of the technique applied in this research. Section 6.3 addresses the work method and the functional architecture is presented. In section 6.4, the results are presented and discussed. Finally, the conclusions and limitations of the study are set out in section 6.5.

6.2 Data mining using Association Rules

The use of association rules has the objective to identify significant patterns of co-occurrence between variables that compose an object of study, such as the association between individual factors and the occurrence of accidents (Barker, 2021).

The technique was initially proposed by Agrawal *et al.* when they formulated a rapid processing algorithm to identify co-occurrence in product consumption in a retail organization. From the registration of commercial transactions, the algorithm was able to identify occurrence patterns of the type: *"when the consumer buys the products A and B (antecedent factors), also buys product C (consequent factor)"*. Later, considering its high potential for application in several other areas, the algorithm then called *Apriori* was modified so it was possible to identify the association rules with multiple antecedent factors and multiple consequent factors (Agrawal and Srikant, 1994)

In a data set D containing transactions of the type $T = \{t_n | n = 1, \dots, N\}$, an association rule is expressed in the form $R_m: \{X\} \rightarrow \{Y\}$, in which $X = \{x_i | i = 1, \dots, I\}$ and $Y = \{y_k | k = 1, \dots, K\}$ represent the antecedent and consequent factors, respectively.

The significance of an association rule in D is defined through statistical thresholds called support $(supp)$ and confidence $(conf)$. These thresholds are used to capture a certain level of relationship between items present in a data set (Liao and Perng, 2008).

Support is defined by the probability of X and Y coexisting in the data set, or the fraction of the number of transactions in the data set that contains all items in a specific rule, i.e., $supp\ (X \rightarrow Y) = \sigma(X \cap Y)/N$. In applications using association rules, a minimum support ($minsupp$) is the threshold used to select combinations of frequent (and hopefully relevant) factors (Hahsler and Hornik, 2008).

Confidence, in turn, represents a measure of the rule's reliability, i.e., the higher the confidence $X \rightarrow Y,$ the higher the probability of Y being present in the transactions containing X, i.e. $conf\ (X \rightarrow Y) = \sigma(X \cap Y)/\sigma(X)$.

An association rule R_m: $\{X\} \rightarrow \{Y\}$ is considered relevant if at least it satisfies the thresholds defined in the model, i.e., minimum support $(supp \geq minsupp)$ and confidence $(conf\ \geq minconf)$ (Liao and Perng, 2008). Such parameters set up can be found in studies available in the literature or defined under the interest of the analyst (Alves, 2020; Baralis and Psaila, 1997; Isa et al., 2018; Kouris et al., 2005; Zhang and Zhang, 2002).

However, depending on the data set, the volume of association rules generated can be very significant despite meeting the minimum criteria established by the thresholds $supp$ and $conf$ used. One way to qualify the analysis is to consider a measure of interest called lift (l). This measure compares the frequency in which X and Y cooccur, against the frequency expected if the variables are statistically independent: $lift = supp(X \rightarrow Y)/supp(X).supp(Y)$ (Hahsler, 2015). When the lift of a rule presents values greater than 1, there is an indication that the factors are positively related (the higher the lift the greater the strength of the rule).) For lift values less than 1, the factors have a weak relationship with each other. A lift value of 1 indicates independence between X and Y.

Thus, the association rules are represented according to the elements present in Table 30.

Table 30 – Structure of association rules adapted from Gomes *et al.* (2022)

Association rule	Antecedent factor (X)		Consequent factor (Y)	$supp$	$conf$	$lift$
$R1$	{A, B}	=>	{C}	[0 – 1]	[0 – 1]	[0 - ∞]
$R2$	{A, D, E}	=>	{F, G}	[0 – 1]	[0 – 1]	[0 - ∞]

The results obtained by using association rules to identify hidden patterns of co-occurrence of factors have called the attention of academics and OSH professionals. This is evidenced by the growing volume of studies on the application of this technique in accidents prevention, particularly in the construction industry, railway and road safety (Liao and Perng, 2008; Mirabadi and Sharifian, 2010; Montella, 2011). However, the literature is still scarce on applications that integrate this technique into the architecture of the organizational health and safety management system. In the next section, we discuss the work method used in the case study.

6.3 Steps of implementation of the proposed solution

This study used the method Knowledge Discovery in Databases (KDD) proposed by Fayyad (1996). KDD means the discovery of knowledge in databases. It is, therefore, a non-trivial process of identifying valid, new, potentially useful, and ultimately understandable patterns in data sets (Fayyad et al., 1996).

In the initial stage, the data used in the research were collected from accident investigation reports that occurred between October 2020 and February 2022. In total, 74 investigation reports were mapped in the TK Elevator's Latin American operations, which include 12 countries, dozens of operating units, and one plant.

In the next step, the data were compiled into a table composed of factors and attributes considered for understanding root causes, such as characterization of the accident site, working conditions, individual and organizational factors. The list of factors and attributes is presented in Table 31 and was defined based on previous studies (Cheng et al., 2010; Verma et al., 2014) and complemented by experts from the organization.

Table 31 – Factors and attributes used in investigation reports

Seq.	Factor	Attribute
1	Type of accident	Lost time; without lost time; fatal
2	Line of business	Maintenance; new installation; modernization
3	Type of contract	Employee; contractor
4	Schooling	Fundamental; medium; college
5	Technical training	No training; mechanics; electric; Other
6	Job function	Maintenance technician; installer; and so on
7	Years of experience	<1 year; between 1 and 5 years; more than 5 years; more than 10 years
8	Equipment	Elevator; escalator; accessibility; other
9	Day of the week/shift	Monday, Tuesday, ... / morning, afternoon, night
10	Task condition	Routine; non-routine
11	Project status	On-time; late
12	Accident location	Machine room, shaft, pit, and other

13	Working time at the accident	Regular; extra
14	Risk analysis (1)	JHA performed; JHA not performed
15	Risk analysis (2)	JHA adequate to task; JHA not adequate no task
16	Use of tools	Suitable; not suitable
17	Use of PPEs	PPEs in use; Non-compliant PPEs
18	Standard working procedure	Available; not available
19	Adherence to the standard procedure	Compliant; non-compliant
20	Training for the task	Yes; No
21	Previous history of accidents	Yes; No
22	Prior history of warnings	Yes; No
23	Audits carried out in the last 12m	Yes; No
24	Observation comport. in the last 12m	Yes; No
25	Recent behavioral change	Yes; No
26	Psychological test performed before the work	Yes; No
27	Primary cause of accident	Process failure; behavior deviation
28	Type associated with primary cause	Unsafe condition; unsafe act; unsafe act of others
29	Association with violated safety rule	Fall protection, electrical circuits, barricades, and others
30	Association with behavioral trap	Lack of concentration, lack of awareness, others

Based on the structured data (Table 31), the *Apriori* algorithm was used to generate the most relevant association rules. In this step, the RStudio software (2020) was used for code processing (see appendix A 17) from pre-established support (and trust) parameters. The definition of parameters took into account $supp = 0.2)$ $(conf = 0.5)$references in the literature and the organization's interest in not excessively restricting the number of rules generated as a way to ensure a broad research in the data.

After analysis and validation of the model with the support of OSH specialists within the organization, the next step was the definition and configuration of a functional architecture to integrate the data mining process using the *Apriori* algorithm with the existing systems in the organization (Figure 32).

In simple terms, the proposed architecture has two streams. The first connects a data source composed of primary files (e.g. accident reports from users located across Latin America operations using the software ProcessMAP). Then it transforms the data into a secure and structured standard based on the analysis and accident investigation. Next, the data is stored in the *data lake* (a term commonly used in IT to express the place where data already processed is stored). From the *data lake*, a second stream is triggered for processing the *Apriori* algorithm through a platform that unifies databases and uses multiple computational languages (*databricks*). Finally, the results retrieved from the processing of the association rules return to *the data lake* and are displayed in Power B.I. for managerial analysis.

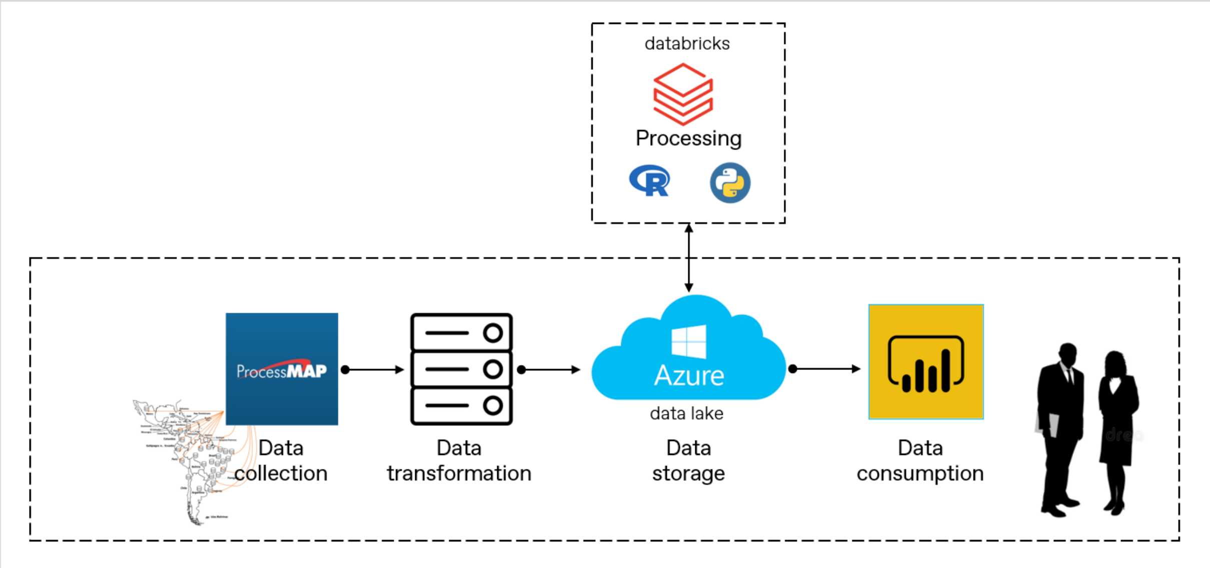


Figure 32 – Functional architecture for processing association rules

Since the architecture was developed, all accident investigation reports are structured and stored in the same database, in which the relevant association rules are updated and results are displayed in a form of a business intelligence dashboard for the use of managers and OSH specialists. The results of this architecture are presented and discussed in the following section.

6.4 Results

6.4.1 Functional architecture

The definition of an integrated functional architecture resulted in rapid adherence to data consumption by OSH specialists and operational managers. Some advantages of this architecture include: (i) the visualization of the data mining process (Figure 33); (ii) easy and quick access to the association rules generated by each of the selected parameters; and (iii) the possibility of navigating between different analyses from the selection of consequent factors of interest (RHS field).

Figure 32 and Figure 33 present the results of the developed architecture. The attributes that compose the accident investigations are processed in such a way that the strength of their associations is identified, expressed in the dashboard through the thickness and color of the line, as well as through the size of the circle. For example, the larger the circle, the higher the frequency of co-occurrence between 2 attributes, as well as the thicker the connection line between attributes, the greater the confidence or lift.

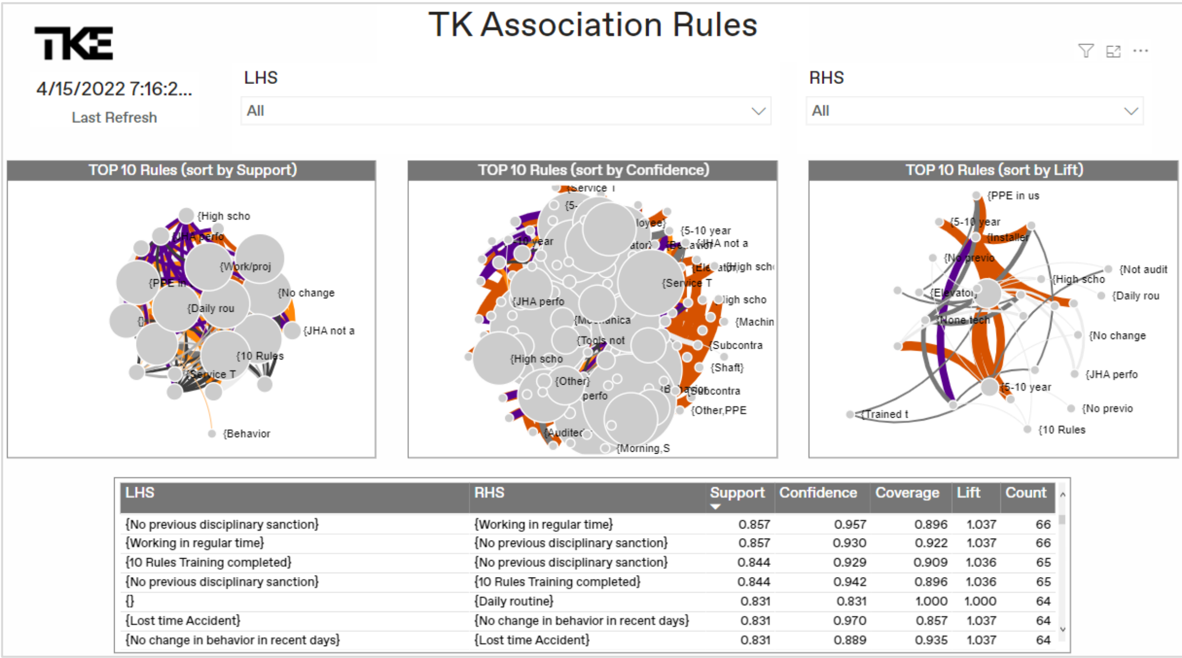


Figure 33 – Data mining in processing (screenshot)

6.4.2 Association rules and patterns of occurrence

In total, 39 association rules were generated based on the thresholds (support and confidence) established for the model. As a result, some expected patterns of occurrence have been confirmed, as well as other counterintuitive scans have been revealed. The set of association rules in which the impact on accident prevention is considered significant is presented as follows:

a) Accidents occur predominantly during normal working hours and routine activities $(supp = 0.727)$;
b) Approximately 70% of accidents occur with unaudited workers in the last 3 months $(supp = 0{,}662)$;
c) A considerable amount of injured workers had no previous history of accidents at work $(supp = 0{,}818)$, nor even disciplinary sanctions $(supp = 0{,}79)$;

d) More than 80% of injured workers were trained in the safety rules ($supp = 0{,}805$);
e) A poor job hazard analysis is associated with more than 60% of accidents ($supp = 0{,}623$);
f) The lack of technical training is associated with accidents at high confidence ($supp = 0{,}390;\ conf = 0{,}968$);
g) Accidents involving workers with more than 10 years of experience and using PPEs compose a relevant and counterintuitive pattern of occurrence of accidents ($conf = 0{,}944;\ l = 1{,}102$);
h) The co-occurrence between a poor job hazard analysis conducted by a subcontractor and a work-related accident represents a relevant association rule ($conf = 0{,}944;\ l = 1{,}102$);
i) Even though adequate work tools are available, non-audited workers in the last 3 months constitute a potential risk of accidents ($conf = 0{,}950;\ l = 1{,}108$);
j) The condition itself of unaudited subcontractors workers in the last 3 months configure a relevant association with accidents ($conf = 0{,}995;\ l = 1{,}114$);
k) Co-occurrence between a lack of technical training and the absence of psychological evaluation represents a risk to the organization ($conf = 1{,}000;\ l = 1{,}167$);
l) Accidents occur even after recent behavioral observation audits ($conf = 1{,}000;\ l = 1{,}167$).

All the association rules selected in the analysis were considered valid occurrence patterns to drive more effective plans for the reduction of accidents. Additionally, the rules revealed some counterintuitive patterns of accidents that occur with trained workers or using PPEs. The following section discusses the results from the perspective that accident prevention can be achieved by interrupting a pattern of occurrence that results in an unwanted factor.

6.4.3 Accident prevention based on patterns of occurrence

The study is usually associated with the co-occurrence of multiple factors, including deviations from procedures, inadequate risk analysis, and lack of use of personal protective equipment.

These contributing factors for the occurrence of accidents are usually identified in organizations through root cause analysis. However, despite its importance for the prevention process, isolated analysis of accidents does not allow the identification of patterns of occurrence since it does not consider the historical data set.

In this context, the combination of factors (antecedent factors) that co-occur with a certain unwanted consequent factor (e.g., accident) can be identified through advanced techniques, such as data mining using association rules.

In this study, the association rules revealed patterns of occurrence of factors that co-occur with work-related accidents (e.g., process deviation combined with the lack of use of PPEs). It is important to highlight that we should not assume the existence of causality (if "A" occurs then "B" occurs"). Instead, we want to highlight the existence of a strong association between attributes that coexist in the records of the analyzed data (if "A" occurs then it is likely that "B" also occurs).)

As previously highlighted, the set of association rules resulting from the analysis performed in the 74 accidents investigated confirms the conditions already expected by the OSH experts of the organization, as well as reveal other counterintuitive conditions.

On the one hand, some expected contributing factors were confirmed by association rules, such as the strong association between work-related accidents and lack of qualification (item f), or the co-occurrence of accidents with poor job hazard analysis (item e).

On the other hand, some counterintuitive patterns have aroused the attention of the organization's experts and managers. These hidden patterns, revealed by the association rules, bring significant benefits to the accident prevention process. For example, the occurrence of accidents with workers without a previous history of accidents or disciplinary sanctions demonstrates that accidents do not occur solely with a work profile characterized by recurrent indiscipline, but also with any worker (item c).

Another counterintuitive pattern is the occurrence of accidents with workers using PPEs. This demonstrates that the mere fact of using EPIs does not ensure injury-free work. The study draws attention to aspects such as overconfidence, typically related to experience and use of PPEs (item g).

Other rules revealed, for example, the ineffectiveness of safety rules training, the importance of psychological testing in subcontracted workers, the relevance of risk analysis, and the importance of the frequency of audits in accident prevention.

All occurrence patterns mentioned in the previous section are statistically relevant, and allow management actions to be established for accident prevention.

This process can be represented by eliminating or neutralizing the co-occurrence of antecedent factors. For example, once the strong association between accidents and unaudited subcontractor workers in the last 3 months has been identified, the organization is encouraged to establish processes that ensure more frequent audits, to eliminate or neutralize this pattern of occurrence in which a likely result is unwanted.

In this context, once a work-related accident is considered an unwanted consequent factor, managerial efforts should be driven to avoid the co-occurrence of antecedent factors identified in the relevant association rules.

6.4.4 Impact on efficiency

The results from the application of association rules not only supported managers to improve efficacy in reducing accidents, but also efficiency. In this research, each fiscal year (FY) corresponds to a period from October to September in the year after, e.g., FY18 corresponds to October 2017 to September 2018 as depicted in Figure 34.

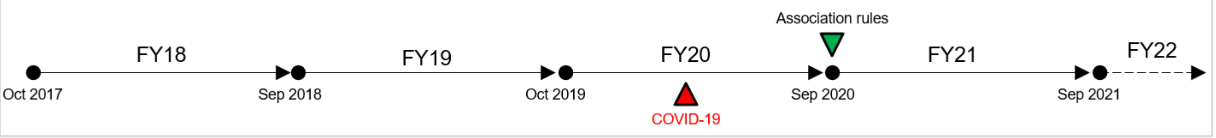


Figure 34 – Critical incidents in the period of analysis

Two critical incidents (CIs) are highlighted in the period of analysis. CIs are fundamental events to understand the effect of the intervention over time on a variable under interest (Piran et al., 2021). In this case study, one critical incident was COVID-19 which was considered a pandemic in March 2020 with a direct impact on society, including organizations (Gomes et al., 2021). The other is the fact that the organization implemented the artifact presented in section 6.3 at the beginning of FY21, i.e., in

October 2020. The implementation of a data mining technique was part of a set of initiatives undertaken by the organization to improve its results in safety, including efficacy and efficiency.

Table 32 summarizes the efficiency scores year on year before and after the implementation of the functional architecture presented in subsection 6.4.1. The efficiency scores were obtained using a standard DEA model proposed by Charnes et al.(1978) with orientation to output and constant returns to scale. This configuration is suitable for internal benchmark applications since all DMUs comprise one organization with comparable business units. The variables used as inputs and outputs are also presented in section 5. Unlike the DDF model used in the previous section, data transformation was needed to deal with undesirable outputs (Y_{b_1}, Y_{b_2}) in the DEA model, as presented in the appendices A 18 and A 19.

Table 32 – Efficiency year on year (before and after the use of association rules)

DMU	BEFORE THE USE OF ASSOCIATION RULES					AFTER	
	FY 2018 $\emptyset_k$ (DEA)	Δ	FY 2019 $\emptyset_k$ (DEA)	Δ	FY 2020 $\emptyset_k$ (DEA)	Δ	FY 2021 $\emptyset_k$ (DEA)
UC-01	1	→	1	→	1	→	1
UC-02	0.8073	↗	0.8382	→	0.8082	↗	0.9691
UC-03	0.5355	↘	0.4942	↘	0.4534	↗	0.5841
UC-04	0.6118	↘	0.5311	↘	0.4467	↗	0.6866
UC-05	1	→	1	↘	0.7500	↗	1
UC-06	0.5886	↗	1	→	1	→	1
UC-07	0.7892	↗	0.9481	↘	0.6752	↗	0.6914
UC-08	0.9257	↗	1	↘	0.5000	↗	0.6180
UC-09	0.5294	↗	1	→	1	→	1
UC-10	1	→	1	↘	0.5000	↗	1
UC-11	1	↘	0.6726	↗	0.9405	↘	0.8502
UC-12	0.6331	↗	0.6514	↘	0.4912	↗	0.5630
UC-13	0.7562	↗	0.8973	↗	1	→	1
UC-14	1	→	1	→	1	→	1
UC-15	1	→	1	→	1	→	1
UC-16	1	→	1	→	1	→	1
UC-17	1	→	1	↘	0.8199	↗	1
UC-18	0.7033	↘	0.6538	↘	0.6043	↗	1
UC-19	1	→	1	↘	0.7782	↘	0.6265
UC-20	0.8046	↘	0.7429	↘	0.5794	↗	0.8215
UC-21	1	→	0.9927	↘	0.5802	↗	0.6411
Mean	0.8421	↗	0.8773	↘	0.7584	↗	0.8596
Mean			**0.8259**			↗	**0.8596**

Results show an increase in efficiency in FY 2021 ($\emptyset_k = 86\%$) when compared to the average of antecedent periods to the implementation of the association rules, i.e. FY 2018 – FY 2020 ($\emptyset_k = 82.6\%$). Also, a significant improvement was observed between the periods 2020 and 2021. In total, 12 out of 14 inefficient DMUs improved in efficiency, including 4 cases (UC-05, UC-10, UC-17, and UC-18) that moved to the efficiency frontier.

However, since the analysis was first conducted year on year, the frontier of efficiency might have changed, so conclusions concerning efficiency improvements should be carefully considered. To avoid this typical pitfall when using frontier techniques such as DEA, a second analysis was conducted with panel data. By following this method, the period of 2018 – 2021 is evaluated as one data set, and results are presented in Table 33.

Table 33 – Efficiency in panel (2018-2021)

DMU	$\emptyset_k$						
	FY 2018	Δ	FY 2019	Δ	FY 2020	Δ	FY 2021
UC-01	1	→	1	→	1	→	1
UC-02	0.5918	↗	0.6077	↗	0.8081	↘	0.7232
UC-03	0.3384	↗	0.3460	↗	0.4399	↗	0.5
UC-04	0.3750	→	0.3750	↗	0.4456	↗	0.5481
UC-05	1	→	1	→	0.7500	↗	1
UC-06	0.500	↗	1	→	1	→	1
UC-07	0.5457	↗	0.5805	↗	0.6738	↘	0.6226
UC-08	0.7500	↗	1	↘	0.5000	→	0.5000
UC-09	0.500	↗	1	→	1	→	1
UC-10	1	→	1	↘	0.5000	→	0.5000
UC-11	0.5795	→	0.5729	↗	0.9404	↘	0.5863
UC-12	0.4524	→	0.4541	↗	0.4912	↘	0.4635
UC-13	0.5	↗	0.7500	↗	1	→	1
UC-14	0.5914	↗	0.6592	↗	1	↘	0.9275
UC-15	1	→	1	→	1	↘	0.9423
UC-16	0.8698	↘	0.8279	↗	1	→	1
UC-17	0.8170	→	0.8112	→	0.8198	↗	1
UC-18	0.5367	↘	0.4353	↗	0.5458	↗	1
UC-19	0.6047	↘	0.5881	↗	0.6798	↘	0.6264
UC-20	0.5119	↗	0.5375	↗	0.5794	↗	0.6312
UC-21	0.7331	↗	0.7524	↘	0.5802	↘	0.5025
Mean	**0.6570**	↗	**0.7284**	↗	**0.7502**	↗	**0.7654**
Mean			**0.7119**			↗	**0.7654**

As a result of the analysis using panel data, the overall efficiency increased from 71.2% (average in the antecedent period 2018-2020) to 76.54% (2021) after the implementation of association rules. This positive variation in overall efficiency is a result of the reduction of undesirable outputs, i.e., contraction of work-related accidents and lost days, as well as an increase in labor inspections, a desirable output. The effect of the variation in the inputs throughout the analysis was not relevant, since its variability was marginal as depicted in Figure 35.

Thus, apart from the use of association rules has shown benefits to drive improvements in safety management concerning the reduction of work-related accidents, it also contributes to increasing efficiency in the Occupational Health and Safety Management System.

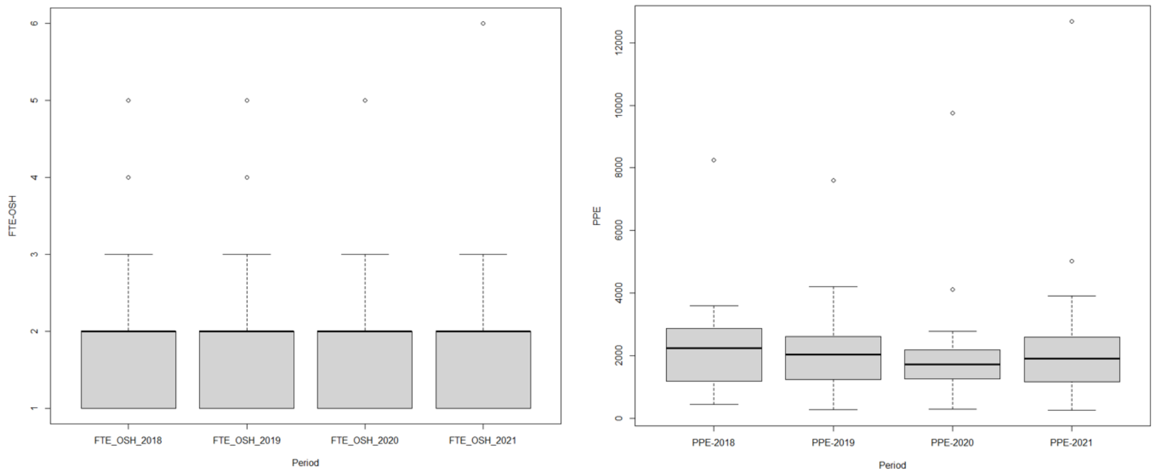


Figure 35 – Variability for inputs (FTE OSH and PPE).

In addition, we use the Wilcoxon test to evaluate the significance of the differences between the medians of the efficiency scores year on year. The purpose of the Wilcoxon test is a non-parametric method to compare the performances of each subject (or pairs of subjects) to verify if there are significant differences between their results in the two situations or periods (Wilcoxon, 1945).

Thus, based on the data set presented in Table 32, we considered the null hypothesis H_0 the condition in which no significant differences exist between the efficiency before and after the intervention. The alternative hypothesis is, therefore, that there is a significant difference in the efficiency scores after the intervention. Table 34 summarizes the results.

Table 34 – Wilcoxon test

Paired period of analysis	Wilcoxon test		Hypothesis test result
	V	$p-value$	
2018 → 2019	35	0.4846	Accept H_0
2018 → 2020	112	0.0975	Accept H_0
2018 → 2021	47	0.7536	Accept H_0
2019 → 2020	104	0.0134	Reject H_0
2019 → 2021	43	0.7837	Accept H_0
2020 → 2021	11	0.0100	Reject H_0

Results from the Wilcoxon test show the significant effect of both critical incidents presented in Figure 34. On the one side, COVID-19 impacted on efficiency as previously discussed, notably due to restriction imposed to labor inspections (a desirable output). On the other side, the use of association rules implemented at the beginning of FY 2021 impacted positively the efficiency even under the effects of the ongoing pandemic. Although the observations concerning other further periods are recommended to confirm the causality related to the gain in efficiency since the application of data mining techniques in the field of safety management, our work opens a new avenue for researchers interested in connecting safety data science and efficiency analysis.

6.5 Conclusion

This study used the association rules to identify useful and relevant patterns of occurrence for factors related to work-related accidents.

The model was applied in a data set composed of investigation reports of 74 accidents that occurred between October 2020 and February 2022 in the selected case study.

As a result, a set of 39 association rules were generated from the thresholds established for the model, in which some expected patterns of occurrence were confirmed, and other counterintuitive ones were revealed.

The identification of relevant patterns of occurrence is beneficial to drive management actions toward accident prevention. This process consists of eliminating or neutralizing the co-occurrence of the antecedent factors when the consequent factor is the work-related accident.

Furthermore, an analysis of efficiency was applied before and after the implementation of association rules, and results also shown a positive effect of the use of association rules in the efficiency, even though further studies are recommended to confirm causality.

The study has theoretical and practical contributions. First, a functional architecture is presented as a mechanism for integrating the *Apriori* algorithm into existing organizational systems, closing a gap in the literature. Second, the study offers an empirical real-world case to drive, grounded in practicability, how to use association rules in accident prevention, even though further studies are recommended.

APPENDICES CHAPTER 6

A 17. Script R for association rules

```
#R Studio v. 4.0.5
# Require packages
if(!require(readxl)) install.packages("readxl")
if(!require(arules)) install.packages("arules")
if(!require(arulesViz)) install.packages("arulesViz")
if(!require(tidyr)) install.packages("tidyr")

# Load packages
library(readxl); library(arules), library(arulesViz), library(tidyr)
# Load dataset
data <- read_excel("Lost-time accidents Report.xlsx", sheet='DATA')
View(data)

# Adjust dataset
data_aj <- dados [, c(-2,-3,-4,-5,-6,-7,-8,-9)]
View(data_aj)

# Convert dataset into file .csv
write.csv(dados_aj,"AR.csv", quote=FALSE, row.names=FALSE)

# Convert dataset into transaction format
tr <- read.transactions('AR.csv', format = 'basket', sep=',')
tr
summary(tr)
```

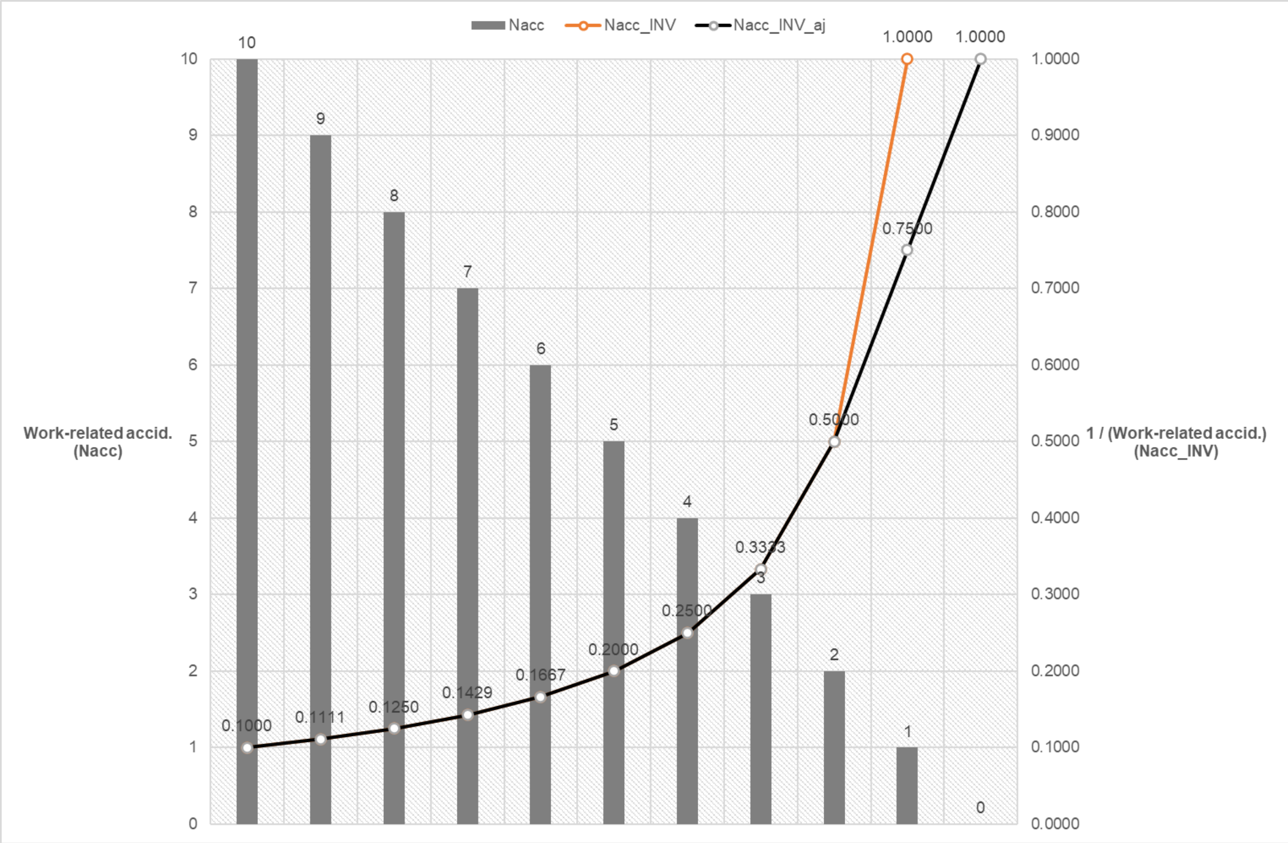
```
# Create association rules
rules = apriori(tr, parameter=list(suppor = 0.5, conf = 0.8, minlen = 1,
maxlen = 3))
rules
inspect(head(rules))

# Remove redundant rules
rules = rules[!is.redundant(rules)]
rules
inspect(rules)
result = inspect(rules)
# Print association rules
write.csv2(result, "Association rules.csv")
```

A 18. Inputs X_1, X_2, desirable Outputs Y_{g1}, Y_{g2}, and undesirable outputs[a] Y_{b1}, Y_{b2}

DMU	Period 2018						2019						2020						2021					
	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}	X_1	X_2	Y_{g1}	Y_{g2}	Y_{b1}	Y_{b2}
UC-01	1	458	32	310	1	1	1	447	27	387	1	1	1	447	49	401	1	1	1	269	43	186	1	1
UC-02	2	2,293	143	1,096	0.75	0.25	2	2,059	154	1,043	0.333	0.000	2	1,731	226	1,230	0.75	0.004	2	1,715	226	885	0.75	1
UC-03	3	3,214	150	892	0.75	0.111	3	2,910	129	849	0.75	0.142	2	2,783	186	624	0.5	0.333	2	5,035	193	604	1	1
UC-04	2	3,031	123	782	0.75	0.012	2	3,626	94	790	0.75	0.018	2	2,191	175	681	0.5	0.047	2	1,773	184	499	0.75	0.071
UC-05	1	1,137	59	313	1	1	1	1,085	58	298	1	1	1	1,452	79	267	0.75	0.111	1	973	76	190	1	1
UC-06	1	1,212	47	247	0.5	0.025	1	1,442	53	201	1	1	1	1,067	59	196	1	1	1	1,136	61	179	1	1
UC-07	2	3,607	188	1,407	0.5	0.066	2	3,366	177	1,417	1	1	2	2,736	276	1,407	0.75	0.076	2	2,968	273	916	0.5	0.111
UC-08	1	1,836	78	579	0.75	0.014	1	1,454	64	502	1	1	2	1,830	100	200	1	1	2	1,165	101	193	1	1
UC-09	1	734	28	241	0.5	0.008	1	1,246	20	204	1	1	1	777	46	305	1	1	1	1,041	47	254	1	1
UC-10	1	926	57	204	1	1	1	698	46	303	1	1	2	1,138	94	533	1	1	2	2,194	91	356	0.75	1
UC-11	4	1,274	163	975	0.333	0.025	4	1,205	149	928	0.5	0.011	3	1,260	205	1,366	0.75	0.125	3	1,811	204	1,054	0.75	0.142
UC-12	2	2,308	118	842	0.333	0.043	2	2,137	120	800	0.5	0.166	2	1,815	161	637	0.75	0.066	2	2,111	162	514	0.75	0.066
UC-13	1	1,659	76	488	0.5	0.020	1	2,044	70	464	0.75	0.012	1	1,726	95	626	1	1	1	1,393	97	466	1	1
UC-14	3	2,504	280	1,320	0.25	0.02	3	1,985	270	1,257	0.333	0.028	2	1,716	377	1,517	0.75	0.066	2	2,651	393	959	0.75	0.005
UC-15	1	512	63	352	1	1	1	291	57	340	1	1	1	308	88	414	1	1	2	325	86	233	1	1
UC-16	1	2,243	119	1,281	0.5	0.032	1	2,622	128	1,220	0.25	0.028	1	1,580	211	1,178	1	1	1	2,440	210	820	0.75	1
UC-17	2	2,269	156	1,502	0.5	0.016	2	2,139	169	1,430	0.75	0.333	2	2,130	249	1,441	0.5	0.007	1	2,194	232	1,483	0.75	0.003
UC-18	2	3,360	153	1,318	0.5	0.025	2	4,220	162	1,255	0.166	0.009	2	4,116	255	1,308	0.5	0.125	2	3,917	496	1,053	0.5	0.052
UC-19	5	8,251	669	2,857	0.090	0.000	5	7,598	625	3,029	0.25	0.012	5	9,743	821	2,504	0.2	0.000	6	12,677	926	2,348	0.333	0.004
UC-20	2	2,870	171	1,120	0.75	0.066	2	2,515	143	1,065	0.25	0.007	2	2,619	224	1,183	0.75	0.034	2	2,599	245	849	1	1
UC-21	3	1,195	136	1,035	0.5	0.005	2	1384	156	985	0.333	0.017	2	1,771	198	633	0.75	0.333	2	1,919	190	458	0.5	0.066

[a] Multiplicative inverse applied to undesirable outputs: $f_k = \frac{1}{k}$.

A 19. Data transformation: multiplicative inverse for undesirable outputs [a,b,c]



[a] Zero work-related accident assumed as 1.000
[b] 1 work-related accident assumed as 0.75
[c] N work-related accident assumed as $\frac{1}{N}$

7 DISCUSSION AND CONCLUSIONS

This section discusses the main conclusions derived from this thesis. Subsection 7.1 discusses the core argument of this thesis, the research objectives as well as the main contributions of this doctoral research. Subsection 7.2 acknowledges the limitations of this research and directions for future research are given in subsection 7.3.

7.1 Thesis, objectives, and contributions

Consistent with the well-known context in which work-related accidents and work-related illnesses remain a global challenge for governments, firms, and ultimately the society, and that the burden resulting from poor safety management impacts directly the individual, collective and economic dimensions, this thesis offers a wider view on the evaluation of OHSMS beyond efficacy.

Although setting up goals for safety is a useful and common practice to guide directions for practitioners and decision-makers, this narrow view on the evaluation of OHSMS has not been sufficient to change the current scenario of more than 90 million injured workers and almost 2 million deaths every year (WHO and ILO, 2021).

Therefore, we defend the thesis that OHSMS should be evaluated from the perspective of efficiency, and that the DEA-based model can contribute to this objective similarly to any other system. This does not mean efficacy is useless as a safety performance measurement. Instead, we consider it a necessary but not sufficient metric.

Differently from the efficacy, the analysis of efficiency offers a wider perspective of the variables impacting safety, as well as the inputs and outputs related to the performance of the OHSMS. By configuring safety as a system, a clear view is given to managers on the necessary resources applied to generate the desired outcomes and eliminate the undesirable ones. Also, since efficiency is a relative measure, it allows external and internal benchmarks to drive safety improvements based on the comparison between homogeneous entities, e.g. companies in the same industry, and branches composing one organization.

The relevance of performance measures in the field of safety at work was discussed in section 3. Based on the Theory of Constraints and following common strategies of theory building to reach common sense, the constructs that govern safety at work, i.e., knowledge, planning, behavior, and performance measurement were identified and logical propositions between them were explained. As a result, we confirmed safety is governed by very few constructs as reasoned in the concept of inherent simplicity, and postulate the complexity of safety work is a function of the degrees of freedom and harmony between constructs that govern the work environment.

By considering science as an evolutionary process, a systematic literature review was conducted to scan the existing OHSMS assessment instrumens and analyze their critical elements. In section 4, we exploit the critical elements that constitute OHSMS assessment instruments and draw the attention of practitioners to the relevance of using rigorous performance measures concepts to avoid misinterpretations of the results. For example, productivity metrics such as efficacy, efficiency, and effectiveness were verified as performance measures applied to assess safety performance. However, those measures were not consistent with recognized literature grounded outside of safety science. Thus, as a major contribution to this systematic review, conceptual directions are given for further studies of OHSMS evaluation.

As the primary objective of this research was to analyze the efficiency of OHSMS, a DEA-based model was applied in a contemporaneous context, i.e. the Agenda 2030 and its sustainable goals, including safe work environment. Thus, also seeking managerial outcomes for this thesis, in section 5 we use the directional distance function (DDF) to evaluate the efficiency (and potential for improvements) in a real-world case conducted in a large organization. DDF is a frontier technique such as DEA, but with benefits to deal with undesirable outputs with no data transformation, which makes it suitable for managerial applications. In addition, we analyze the efficacy for the same sample of DMUs, and results showed that a composite measure combining efficiency and efficacy herewith called effectiveness is necessary for a more precise evaluation of decent work. Thus, we argue that the higher the effectiveness, the safer the work environment.

Finally, since the use of data science is increasing within organizations, in section 6 we applied association rules mining to identify patterns of occurrence in work-related accidents to unhide combined factors with the potential to contribute to the occurrence of accidents. This case-based revealed important intuitive and counterintuitive associations to support managers and safety practitioners in the accident prevention process. Moreover, we investigated the effect of the use of association rules on efficiency, and the results showed a positive impact even though further studies are recommended.

The studies conducted and presented in sections 3, 4, 5, and 6 are connected, cohesive, and aligned with the objectives proposed for this thesis, and contributions are found in both theoretical and managerial fields.

In short, theoretical contributions are given to expand the understanding of safety at work as *phenomena*, in which its complexity depends on how organizations are capable to harmonize the relationships between the constructs, decreasing degrees of freedom. Also, rigorous concepts related to performance measures, and grounded in recognized literature, are given for guiding further studies in OHSMS evaluation.

This thesis also offers relevant managerial contributions. First is the analysis of efficiency in OHSMS as a mechanism to improve safety performance at the organizational level. Second is the combination of efficacy and efficiency as a composite metric so-called effectiveness to expand the quality of the analysis. The third is the artifact using association rules mining to identify patterns of occurrence in work-related accidents, as an advanced management tool to reduce undesirable outputs and improve efficiency. Table 35 summarizes the main contributions of this doctoral research:

Table 35 – Theoretical and managerial contributions of the thesis

Main contributions	
Theoretical	Managerial
1- A theoretical framework composed of a few constructs that govern safety at work is presented, and logical propositions between its elements are explored.	1- Efficiency analysis of OHSMS is applied as a complementary performance measure to drive safety improvements based on benchmarking.
2- Complexity in safety at work is discussed and postulated based on the concept of inherent simplicity stemming from the Theory of Constraints.	2- A replicable DDF model dealing with undesirable outputs is presented with the power for generalization to be applied in different economic sectors.

3- A OHSMS theoretical framework is redesigned based on a robust systematic literature review.

4- Critical elements composing OHSMS assessment instruments are identified.

5- Performance measurement concepts, i.e., efficacy, efficiency, and effectiveness applied to OHSMS are proposed to drive future research based on a rigorous conceptual basis.

6- A DDF-based model is proposed to measure OHSMS efficiency using variables defined in the Decent Work Measurement Framework published by the United Nations.

7- A combination of efficacy and efficiency is proposed to compose a wider perspective of performance analysis, so-called effectiveness.

8- An initial mathematics model for effectiveness is proposed without weight restrictions.

3- A combined graphical representation of efficacy and efficiency for OHSMS performance evaluation and management reviews.

4- An artifact based on association rules mining to reveal patterns of occurrence in work-related accidents.

5- A reality-based architecture to integrate safety data science techniques with existing management systems.

7.2 Research limitation

Although the outcomes of this thesis reached its objectives, there are some limitations. One is the fact that economic efficiency was not approached in our empirical applications. As explored in section 4 and shown in Figure 5, the outputs of OHSMS are composed of final, intermediate and economic. This work focused only on the final outcomes, and variables related to intermediate and economic were not considered in the modeling.

Another limitation is the fact that empirical applications were conducted in only one organization, so generalization should be carefully considered.

From the modeling perspective, this research did not consider weight restrictions or any other variation from classic DEA and DDF models. Thus, inputs and outputs were considered equivalent. Finally, no further evaluation was conducted to verify the power of discrimination when using DEA with data transformation since some variables vary slightly from zero to 1.

7.3 Directions for future research

This thesis opens an avenue to at least four relevant directions for future research. The first is related to the theoretical framework presented in section 3 (Figure 21). Researchers are encouraged to explore the constructs and their logical propositions to advance in reality-based safety science grounded in rigorous observations of existing practice, as reasoned by Rae et al. (2020).

The second is to expand the efficiency analysis to touch on the intermediate and economic dimensions. Although investments in safety have increased in recent years, results are far from reaching the expected outcomes. A model based on economic efficiency can put light on the resources (inputs) used, and how it is converted into desirable outputs. Also, modeling using inputs associated with intermediate outcomes, such as engagement, are important to be considered in the analysis of efficiency.

The third is to advance the studies concerning composite measures to evaluate OHSMS, such as combining efficacy and efficiency. Although this thesis is seminal in offering a mathematical formulation of the so-called effectiveness, further studies are highly recommended.

Finally, safety data science (SDS) is still a neglected body of knowledge by academics and practitioners, and critical barriers to its massive use need to be overcome as reasoned by Gomes (2022). Opportunities in this field are open to innovative models using machine learning, analytics, and artificial intelligence applications to overcome some existing challenges to prevent accidents from happening.

REFERENCES

ABNT, A.B. de N.T.-, 2001. Cadastro de acidente do trabalho - Procedimento e classificação.

Agrawal, R., Imielinski, T., Swami, A., 1993. Mining Association in Large Databases. Proc. 1993 ACM SIGMOD Int. Conf. Manag. data 207–216.

Agrawal, R., Srikant, R., 1994. Fast Algorithms for Mining Association. Proc. 20th VLDB Conf. 487–499.

Alper, D., Sinuany-Stern, Z., Shinar, D., 2015. Evaluating the efficiency of local municipalities in providing traffic safety using the Data Envelopment Analysis. Accid. Anal. Prev. 78, 39–50. https://doi.org/10.1016/j.aap.2015.02.014

Alves, F.L.D., 2020. Data mining como suporte a tomada de decisão na priorização de estoques. Univ. do Val. do Rio dos Sinos.

Alves, M., Santos, A., Rui, J., Cunha, R., 2021. Incorporating preference information in a range directional composite indicator : The case of Portuguese public hospitals R. Eur. J. Oper. Res. 294, 633–650. https://doi.org/10.1016/j.ejor.2021.01.045

Amini, M., Dabbagh, R., Omrani, H., 2019. A fuzzy data envelopment analysis based on credibility theory for estimating road safety. Decis. Sci. Lett. 8, 275–284. https://doi.org/10.5267/j.dsl.2019.1.001

Antić, B., Grdinić, M., Pešić, D., Pajković, V., 2020. Benchmarking of the road safety performance among the regions by using DEA, in: Ignaccolo M., T.M. (Ed.), Transportation Research Procedia. Elsevier B.V., pp. 78–86. https://doi.org/10.1016/j.trpro.2020.02.065

Arezes, P.M., Sérgio Miguel, A., 2003. The role of safety culture in safety performance measurement. Meas. Bus. Excell. 7, 20–28. https://doi.org/10.1108/13683040310509287

Arocena, P., Núñez, I., 2010. An empirical analysis of the effectiveness of occupational health and safety management systems in SMEs. Int. Small Bus. J. 28, 398–419. https://doi.org/10.1177/0266242610363521

ATLAS.ti (v.8.4.24), 2020. ATLAS.ti. A qualitative data analysis.

Babajani, R., Abbasi, M., Azar, A.T., Bastan, M., Yazdanparast, R., Hamid, M., 2019. Integrated safety and economic factors in a sand mine industry: A multivariate algorithm. Int. J. Comput. Appl. Technol. 60, 351–359. https://doi.org/10.1504/IJCAT.2019.101180

Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Manage. Sci. 30, 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Baralis, E., Psaila, G., 1997. Designing Templates for Mining Association. J. Intell. Inf. Syst. 32, 7–32.

Bardin, A., 1993. Content Analysis. Presses Universitaires de France Le Psychologue, Paris.

Bardin, L., 1993. L'analyse de contenu [Content Analysis]. Presses Universitaires de France Le Psychologue, Paris.

Barker, T.T., 2021. Finding Pluto: An Analytics-Based Approach to Safety Data Ecosystems. Saf. Health Work 12, 1–9. https://doi.org/10.1016/j.shaw.2020.09.010

Barratt, M., Choi, T.Y., Li, M., 2011. Qualitative case studies in operations management : Trends , research outcomes , and future research implications. J. Oper. Manag. 29, 329–342. https://doi.org/10.1016/j.jom.2010.06.002

Belu, C., Manescu, C., 2013. Strategic corporate social responsibility and economic performance. Appl. Econ. 45, 2751–2764. https://doi.org/10.1080/00036846.2012.676734

Benbasat, I., Goldstein, D.K., Mead, M., 1987. The Case Research Strategy in Studies of Information Systems. MIS Quaterly 11, 369–386.

Beriha, G.S., Patnaik, B., Mahapatra, S.S., 2011. Safety performance evaluation of Indian organizations using data envelopment analysis. Benchmarking 18, 197–220. https://doi.org/10.1108/14635771111121676

Berkey, B.D., Dowd, W.M., Jones, P.I., 1993. Integration Of Quality, Process Safety, And Environmental Management Systems, in: International Process Safety Management Conference and Workshop. AIChE, pp. 93–100.

Bianchini, A., Donini, F., Pellegrini, M., Saccani, C., 2017. An innovative methodology for measuring the effective implementation of an Occupational Health and Safety Management System in the European Union. Saf. Sci. 92, 26–33. https://doi.org/10.1016/j.ssci.2016.09.012

Bigelow, P.L., Robson, L.S., 2005a. Occupational Health and Safety Management

Audit Instruments: A Literature Review, 4th ed, Institute for Work & Health. Institute for Work & Health, Toronto. https://doi.org/10.1002/9781118131350.ch25

Bigelow, P.L., Robson, L.S., 2005b. Occupational Health and Safety Management Audit Instruments: A Literature Review, 4th ed, Institute for Work & Health. Institute for Work & Health, Toronto. https://doi.org/10.1002/9781118131350.ch25

Bird, F., Germain, G., 1976. Practical Loss Control Leadership. International Loss Control Institute, Loganville, GA.

Blanco, M., Bares, L., Ferasso, M., 2022. Efficiency Analysis of Graduate Alumni Insertion into the Labor Market as a Sustainable Development Goal. Sustain. 14. https://doi.org/10.3390/su14020842

Bogetoft, P., Otto, L., Otto, M.L., 2022. R Package “Benchmarking”: Benchmarking with DEA and SFA.

Bonoma, T. V., 1985. Case Research in Marketing: Opportunities, Problems, and a Process. J. Mark. Res. 22, 199–208.

Bowen, H., 1953. Social Responsibility of the businessman, 1st ed. Harper, New York.

Brocal, F., González, C., Sebastián, M.A., Reniers, G.L.L., Paltrinieri, N., 2018. Standardized risk assessment techniques: A review in the framework of occupational safety, in: van Gulijk C. Haugen S., B.A.V.J.E.K.T. (Ed.), Safety and Reliability - Safe Societies in a Changing World - Proceedings of the 28th International European Safety and Reliability Conference, ESREL 2018. CRC Press/Balkema, pp. 2889–2896.

Brunton, G., Stansfield, C., Thomas, J., 2012. Finding relevant studies. An introduction to systematic reviews, 2nd ed. SAGE Publications Ltd, London.

BSI, 1999. OHSAS 18001. Occupational health and safety management systems: Specification.

BSI, 1996. BS 8800. Guide to occupational health and safety management systems.

Bunge, M., 1967. Scientific Research II: The Search for Truth. Springer Verlag, Berlin.

Camanho, A.S., 2018. Performance Assessment using Frontier Techniques (SLIDES - Part I).

Camanho, A.S., Dyson, R.G., 2008. A generalisation of the Farrell cost efficiency measure applicable to non-fully competitive settings. Omega 36, 147–162.

Carroll, A.B., 1999. Corporate social responsibility: Evolution of a definitional construct. Bus. Soc. 38, 268–295. https://doi.org/10.1177/000765039903800303

Cauchick-Miguel, P.A., 2007. Estudo de caso na engenharia de produção: estruturação e recomendações para sua condução. Production 17, 216–229.

Chambers, R.G., Chung, Y., Färe, R., 1996. Benefit and distance functions. J. Econ. Theory 70, 407–419. https://doi.org/10.1006/jeth.1996.0096

Chang, J.S., Han, S., Jo, S., 2020. Road safety performance across local governments: a data envelopment analysis approach. Int. J. Inj. Contr. Saf. Promot. https://doi.org/10.1080/17457300.2020.1810073

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision-making units. Eur. J. Oper. Res. 3, 339.

Cheng, C.-W., Lin, C.-C., Leu, S.-S., 2010. Use of association rules to explore cause-effect relationships in occupational accidents in the Taiwan construction industry. Saf. Sci. 48, 436–444. https://doi.org/10.1016/j.ssci.2009.12.005

Chung, Y.H., Fare, R., Grosskopf, S., 1997. Productivity and Undesirable Outputs : A Directional Distance Function Approach. J. Environ. Manage. 51, 229–240. https://doi.org/https://doi.org/10.1006/jema.1997.0146

Coelli, T.J., Rao, D., O'donell, C.J., Battese, G.E., 2005. An introduction to efficiency and productivity analysis, 2nd ed. Springer, New York.

Cohen, B.P., 1989. Developing sociological knowledge: Theory and method, 2nd ed. Nelson-Hall, Chicago, IL.

Cooper, M.D., 2019. The efficacy of industrial safety science constructs for addressing serious injuries & fatalities (SIFs). Saf. Sci. 120, 164–178. https://doi.org/10.1016/j.ssci.2019.06.038

Cooper, William W., Seiford, L.M., Zhu, J., 2011a. Handbook on Data Envelopment Analysis, 2nd ed. Springer.

Cooper, William W., Seiford, L.M., Zhu, J., 2011b. Data envelopment analysis: History, models, and interpretations, in: Cooper, W.W., Seiford, L.M., Zhu, J. (Eds.), Handbook on Data Envelopment Analysis. Springer.

Cooper, W W, Thore, S., Tarverdyan, R., 2011. A utility function approach for evaluating country performances with DEA - the twin goals of decent work and a fair globalization, Global Operations Management. Nova Science Publishers, Inc.

Cummins, J.D., Weiss, M.A., 2013. Analyzing firm performance in the insurance industry using frontier efficiency and productivity methods, in: Handbook of Insurance. Springer, New York, NY, United States, pp. 795–861.

de Castro, V.F., Frazzon, E.M., 2017. Benchmarking of best practices: an overview of the academic literature. Benchmarking 24, 750–774. https://doi.org/10.1108/BIJ-03-2016-0031

de Souza, I.G., Lacerda, D.P., Camargo, L.F.R., Dresch, A., Piran, F.A.S., 2018. Efficiency and internal benchmark on an armament company. Benchmarking 25, 2018–2039. https://doi.org/10.1108/BIJ-08-2016-0128

Deming, E.W., 1986. Out of the Crisis. Massachussetts Institute of Technology Press.

Denyer, D., Tranfield, D., Van Aken, J.E., 2008. Developing Design Propositions through Research Synthesis. Organ. Stud. 29, 249–269. https://doi.org/10.1177/0170840607088020

Diekemper, R.F., Spartz, D.A., 1970. A quantitative and qualitative measurement of industrial safety activities. ASSE J. 12–19.

Dou, E.-W., Zheng, X.-Y., 2011. Evaluation of contractors' safety performance based on DEA, in: International Conference on Management and Service Science, MASS 2011. Wuhan. https://doi.org/10.1109/ICMSS.2011.5998194

Dresch, A., Lacerda, D.P., Antunes Jr, J.A.V.A.V., 2015a. Design Science Research, 1st ed. Springer International Publishing. https://doi.org/10.1007/978-3-319-07374-3

Dresch, A., Lacerda, D.P., Miguel, P.A.C., 2015b. A Distinctive Analysis of Case Study, Action Research and Design Science Research. Rev. Bras. Gest. Negocios 17, 1116–1133. https://doi.org/10.7819/rbgn.v17i56.2069

Druker, P.F., 2006. The Effective Executive: The Definitive Guide to Getting the Right Things Done, Revised. ed. Harper Business.

Dubé, L., Paré, G., 2003. Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations. MIS Quaterly 27, 597–636. https://doi.org/10.4135/9781473915480.n65

Duroe, L., 2021. Why Health and Safety is Important [WWW Document]. https://www.ecoonline.com/. URL https://www.ecoonline.com/blog/why-health-and-safety-is-important (accessed 6.17.21).

Dyjack, D.T., 1996. Development and evaluation of an ISO 9000-harmonized occupational health and safety management system. Univ. Michigan. University of Michigan.

Dyjack, D.T., Levine, S.P., 1996a. Critical features of an ISO 9001/14001 harmonized health and safety assessment instrument. Am. Ind. Hyg. Assoc. J. 57, 929–935. https://doi.org/10.1080/15428119691014459

Dyjack, D.T., Levine, S.P., 1996b. Development pf an ISO 9000-Compatible Occupational Health Standard - II: Defining the Potential Benefits and Open Issues. Am. Ind. Hyg. Assoc. J. 57, 387–391.

Dyjack, D.T., Levine, S.P., 1995. Development of an ISO 9000 Compatible Occupational Health Standard: Defining the Issues. Am. Ind. Hyg. Assoc. J. 56, 599–609. https://doi.org/10.1080/15428119591016836

Dyjack, D.T., Levine, S.P., Holtshouser, J.L., Schork, M.A., 1998. Comparison of AIHA ISO 9001-based occupational health and safety management system guidance document with a manufacturer's occupational health and safety assessment instrument. Am. Ind. Hyg. Assoc. J. 59, 419–429. https://doi.org/10.1080/15428119891010686

Dyjack, D.T., Redinger, C.F., Ridge, R.S., 2003. Health and safety management system audit reliability pilot project. Am. Ind. Hyg. Assoc. J. 64, 785–791. https://doi.org/10.1080/15428110308984873

Dyson, R.G., Allen, R., Camanho, A.S., Podinovski, V. V, Sarrico, C.S., Shale, E.A., 2001. Pitfalls and protocols in DEA. Eur. J. Oper. Res. 132, 245–259.

Eden, Y., Ronen, B., 2007. Approximately right not precisely wrong: Cost accounting pricing and decision making. The North River Press, Great Barrington, MA.

Eisenhardt, K.M., Graebner, M.E., Eisenhardt, K.M., Graebner, M.E., 2007. Theory Building from Cases: Opportunities and Challenges. Acad. Manag. Rev. 50, 25–32.

El-Mashaleh, M.S., Rababeh, S.M., Hyari, K.H., 2010. Utilizing data envelopment analysis to benchmark safety performance of construction contractors. Int. J. Proj. Manag. 28, 61–67. https://doi.org/10.1016/j.ijproman.2009.04.002

Elsler, D., Takala, J., Remes, J., 2017. An international comparison of the cost of work-related accidents and illnesses. Eur. Agency Saf. Heal. Work 8.

Enrique, B., Marta, B., 2020. Efficacy, Effectiveness and Efficiency in the Health Care: The Need for an Agreement to Clarify its Meaning. Int. Arch. Public Heal. Community Med. 4, 3–5. https://doi.org/10.23937/2643-4512/1710035

Ermel, A.P.C., Lacerda, D.P., Morandi, M.I.W.M., Gauss, L., 2021. Literature Reviews: Modern Methods for Investigating Scientific and Technological Knowledge., in:

Literature Reviews: Modern Methods for Investigating Scientific and Technological Knowledge. Springer International Publishing. https://doi.org/10.1007/978-3-030-75722-9

Ertugrul Karsak, E., Goker, N., 2019. Common Weight DEA-Based Methodology for Ranking APEC Countries by Considering Sustainable Development Goals including Decent Work, Income Inequalities, and Gender Equality, in: Proceedings - 2019 3rd International Conference on Data Science and Business Analytics, ICDSBA 2019. Institute of Electrical and Electronics Engineers Inc., pp. 84–89. https://doi.org/10.1109/ICDSBA48748.2019.00028

EU-OSHA, 2021. The economics of occupational safety and health - the value of OHS to society [WWW Document]. EU-OSHA. URL https://visualisation.osha.europa.eu/osh-costs#!/ (accessed 6.20.21).

Färe, R., Grosskopf, S., Lovell, C.A.K., 1994. Production Frontiers. Cambridge University Press.

Färe, R., Grosskopf, S., Lovell, C.A.K., 1985. The measurement of efficiency of production. Kluwer.

Farrel, M.J., 1957. The Measurement of Productive Efficiency. J. R. Stat. Soc. Ser. A 120, 253–290. https://doi.org/https://doi.org/10.2307/2343100

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., 1996. From Data Mining to Knowledge Discovery: An Overview. AI Mag. 17, 1–54. https://doi.org/https://doi.org/10.1609/aimag.v17i3.1230

Fayyad, U.M., 1996. A concerted effort to develop knowledge. IEEE Expert Intellegent Syst. Their Appl. 11, 20–25.

Ferreira, C.M.C., Gomes, A.P., 2009. Introdução à Análise Envoltória de Dados: teoria, modelos e aplicações. UFV, Viçosa.

Fischer, J., 2004. Social responsibility and ethics: Clarifying the concepts. J. Bus. Ethics 52, 381–390. https://doi.org/10.1007/s10551-004-2545-y

Fisher, M., 2007. Strengthening the Empirical Base of Operations Management. Manuf. Serv. Oper. Manag. Publ. 9, 368–383. https://doi.org/10.1287/msom.1070.0168

Førsund, F.R., 2018. Economic interpretations of DEA. Socioecon. Plann. Sci. 61, 9–15. https://doi.org/10.1016/j.seps.2017.03.004

Ganji, S.S., Rassafi, A.A., Bandari, S.J., 2020. Application of evidential reasoning approach and OWA operator weights in road safety evaluation considering the best and worst practice frontiers. Socioecon. Plann. Sci. 69. https://doi.org/10.1016/j.seps.2019.04.003

Gauss, L., 2020. MARKET-DRIVEN MODULARITY: An Integrated Method to Conceptually Design Modular Product Families. UNISINOS.

Gauss, L., Lacerda, D.P., Cauchick Miguel, P.A., 2020. Module-based product family design: systematic literature review and meta-synthesis. J. Intell. Manuf. https://doi.org/10.1007/s10845-020-01572-3

Ghahramani, A., 2017. Assessment of Occupational Health and Safety Management Systems Status and Effectiveness in Manufacturing Industry, Behavioural Sciences Studies in Psychology.

Golany, B., Roll, Y., 1989. An Application Procedure for DEA. Omega 17, 237–250.

Goldratt, E.M., 2008. The Choice. The North River Press, Great Barrington, MA.

Goldratt, E.M., 1990. What is the thing called Theory of Constraints and it should be implemented? The North River Press, New York, NY, United States.

Gomes, R.F. de S., 2022. Safety Data Science: Basic Foundations and Critical Factors. Proc. Eng. Sci. 4, 217–220. https://doi.org/10.24874/PES04.02.011

Gomes, R.F.S., Gauss, L., Lacerda, D.P., 2021. Fast-response measures to mitigate the COVID-19 health and economic impacts within the organizations: the case of Thyssenkrupp Elevator Brazil. Production 31. https://doi.org/10.1590/0103-6513.20200062

Gomes, R.F.S., Gauss, L., Piran, F.S., Lacerda, D.P., 2022. Safety at Work: a Complex or an Exceedingly Simple Matter? Reliab. Theory Appl. 17, 267–287. https://doi.org/https://doi.org/10.24412/1932-2321-2022-167-267-287

Greenhalgh, T., Wong, G., Jagosh, J., Greenhalgh, J., Manzano, A., Westhorp, G., Pawson, R., 2015. Protocol-the RAMESES II study: Developing guidance and reporting standards for realist evaluation. BMJ Open 5. https://doi.org/10.1136/bmjopen-2015-008567

Hahsler, M., 2015. A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules.

Hahsler, M., Hornik, K., 2008. New probabilistic interest measures for association rules. Intell. Data Anal. 11, 437–455. https://doi.org/10.3233/ida-2007-11502

Halkos, G., Petrou, K.N., 2019. Treating undesirable outputs in DEA: A critical review.

Econ. Anal. Policy 62, 97–104. https://doi.org/10.1016/j.eap.2019.01.005
Han, S., Lee, S., Peña-Mora, F., 2010. Framework for a resilience system in safety management: A simulation and visualization approach, in: W., T. (Ed.), Proceedings of the International Conference on Computing in Civil and Building Engineering. Nottingham, University of Illinois at Urbana-Champaign, United States.
Hedlund, F.H., 2014. The relationship between the implementation of voluntary Five-Star occupational health and safety management system and the incidence of fatal and permanently disabling injury. Saf. Sci. 63, 94–103. https://doi.org/10.1016/j.ssci.2013.10.025
Hermans, E., Brijs, T., Wets, G., Vanhoof, K., 2009. Benchmarking road safety: Lessons to learn from a data envelopment analysis. Accid. Anal. Prev. 41, 174–182. https://doi.org/10.1016/j.aap.2008.10.010
Heuvel, S. van den, Zwaan, L. van der, Dam, L. van, Oude-Hengel, K., Eekhout, I., Emmerik, M. van, Oldenburg, C., Brück, C., Janowski, P., Wilhelm, C., 2017. Estimating the costs of work-related accidents and ill-health: An analysis of European data sources - European Risk Observatory, European Agency for Safety and Health at Work. European Agency for Safety and Health at Work, Luxembourg. https://doi.org/10.2802/566789
Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A., 2021. Cochrane Handbook for Systematic Reviews of Interventions, version 6. ed. Cochrane.
Holvad, T., Hougaard, J.L., Kronborg, D., Kvist, H.K., 2004. Measuring inefficiency in the Norwegian bus industry using multi-directional efficiency analysis. Transportation (Amst). 31, 349–369. https://doi.org/10.1023/B:PORT.0000025400.33939.6d
Ikeziri, L.M., Souza, F.B. de, Gupta, M.C., de Camargo Fiorini, P., 2018. Theory of constraints: review and bibliometric analysis. Int. J. Prod. Res. 57, 1–35. https://doi.org/10.1080/00207543.2018.1518602
ILO, 2022a. Decent work [WWW Document]. https://www.ilo.org/global/topics/decent-work/lang--en/index.htm. URL https://www.ilo.org/global/topics/decent-work/lang--en/index.htm
ILO, 2022b. Measuring decent work [WWW Document].

https://www.ilo.org/integration/themes/mdw/lang--en/index.htm. URL https://www.ilo.org/integration/themes/mdw/lang--en/index.htm (accessed 4.23.22).

ILO, 2020. International Labour Organization Web Page [WWW Document]. URL https://www.ilo.org/global/topics/safety-and-health- at-work/lang--en/index.htm (accessed 7.15.20).

ILO, 2013. Decent work indicators: guidelines for producers and users of statistical and legal framework indicators. ILO manual.

ILO, 2001. Guidelines on Occupational Safety and Health Management Systems.

ILO, 1999. Report of the Director-General: Decent Work, in: 87th Session, International Labour Conference. Geneva, Switzerland.

ILOSTAT, 2021. Statistics on safety and health at work [WWW Document]. URL https://ilostat.ilo.org/topics/safety-and-health-at-work/ (accessed 4.3.21).

İnan, U.H., Gül, S., Yılmaz, H., 2017. A multiple attribute decision model to compare the firms' occupational health and safety management perspectives. Saf. Sci. 91, 221–231. https://doi.org/10.1016/j.ssci.2016.08.018

Isa, N., Kamaruzzaman, N.A., Ramlan, M.A., Mohamed, N., Puteh, M., 2018. Market Basket Analysis of Customer Buying Patterns at Corm Café. Int. J. Eng. Technol. 7, 119–123. https://doi.org/10.14419/ijet.v7i4.42.25692

ISO, 2018. ISO 45001. Occupational health and safety management systems – Requirements with guidance for use.

ISO, 1996. ISO 14001. Environmental management systems — Specification with guidance for use.

ISO, 1994. ISO 9001. Quality systems — Model for quality assurance in design, development, production, installation and servicing.

Jain, R., Yadav, O.P., Rathore, A.P.S., 2008. The propagation of benchmarking concepts in Indian manufacturing industry. Benchmarking 15, 101–117. https://doi.org/10.1108/14635770810854362

Kang, L., Wu, C., Liao, X., Wang, B., 2020. Safety performance and technology heterogeneity in China's provincial construction industry. Saf. Sci. 121, 83–92. https://doi.org/10.1016/j.ssci.2019.09.005

Karanikas, N., 2016. Critical review of safety performance metrics. Int. J. Bus. Perform. Manag. 17, 266–285. https://doi.org/10.1504/IJBPM.2016.077244

Kerstens, K., Sadeghi, J., Van de Woestyne, I., 2019. Convex and nonconvex input-oriented technical and economic capacity measures: An empirical comparison. Eur. J. Oper. Res. 276, 699–709. https://doi.org/10.1016/j.ejor.2019.01.014

Ketokivi, M., Choi, T., 2014. Renaissance of case research as a scientific method. J. Oper. Manag. 32, 232–240. https://doi.org/10.1016/j.jom.2014.03.004

Kouris, I.N., Makris, C.H., Tsakalidis, A.K., 2005. Using Information Retrieval techniques for supporting data mining. Data Knowl. Eng. 52, 353–383. https://doi.org/10.1016/j.datak.2004.07.004

Kouzis-loukas, M., 2014. Analysing Customer Baskets. Erasmus School of Economics.

Krause, T.R., Bell, K.J., 2015. 7 Insights into Safety Leadership, 1st ed. The Safety Leadership Institute.

Krmac, E., Djordjević, B., 2018. Evaluation of the levels of safety at railway level crossings using data envelopment analysis (DEA) method: A case study on slovenian railways. Eur. Transp. - Trasp. Eur.

Kubota, F.I., Cauchick-Miguel, P.A., Tortorella, G., Amorim, M., 2021. Paper-based thesis and dissertations : analysis of fundamental characteristics for achieving a robust structure. Production 31. https://doi.org/10.1590/0103-6513.20200100

Lehtola, M.M., van der Molen, H.F., Lappalainen, J., Hoonakker, P.L.T., Hsiao, H., Haslam, R.A., Hale, A.R., Verbeek, J.H., 2008. The Effectiveness of Interventions for Preventing Injuries in the Construction Industry. A Systematic Review. Am. J. Prev. Med. https://doi.org/10.1016/j.amepre.2008.03.030

Li, H., Lu, M., Hsu, S.-C., Gray, M., Huang, T., 2015. Proactive behavior-based safety management for construction safety improvement. Saf. Sci. 75, 107–117. https://doi.org/10.1016/j.ssci.2015.01.013

Liao, C.W., Perng, Y.H., 2008. Data mining for occupational injuries in the Taiwan construction industry. Saf. Sci. 46, 1091–1102. https://doi.org/10.1016/j.ssci.2007.04.007

Lingard, H., Wakefield, R., Blismas, N., 2013. If you cannot measure it, you cannot improve it: Measuring health and safety performance in the construction industry. In: January 2013, in: 19th Triennial CIB World Building Congress. Brisbane, Aus.

Lynham, S.A., 2002. General Method of Theory Building in Applied Disciplines. Adv. Dev. Hum. Resour. 4, 221–241.

Mahmoudi, S., Ghasemi, F., Mohammadfam, I., Soleimani, E., 2014. Framework for continuous assessment and improvement of occupational health and safety issues in construction companies. Saf. Health Work 5, 125–130. https://doi.org/10.1016/j.shaw.2014.05.005

Mirabadi, A., Sharifian, S., 2010. Application of association rules in Iranian Railways (RAI) accident data analysis. Saf. Sci. 48, 1427–1435. https://doi.org/10.1016/j.ssci.2010.06.006

Mohammadfam, I., Kamalinia, M., Momeni, M., Golmohammadi, R., Hamidi, Y., Soltanian, A., 2017. Evaluation of the Quality of Occupational Health and Safety Management Systems Based on Key Performance Indicators in Certified Organizations. Saf. Health Work 8, 156–161. https://doi.org/10.1016/j.shaw.2016.09.001

Montella, A., 2011. Identifying crash contributory factors at urban roundabouts and using association rules to explore their relationships to different crash types. Accid. Anal. Prev. 43, 1451–1463. https://doi.org/10.1016/j.aap.2011.02.023

Morandi, M.I.W.M., Camargo, L.F.R., 2015. Systematic Literature Review, in: Design Science Research. Springer International Publishing, p. 161.

Nahangi, M., Chen, Y., McCabe, B., 2019. Safety-based efficiency evaluation of construction sites using data envelopment analysis (DEA). Saf. Sci. 113, 382–388. https://doi.org/10.1016/j.ssci.2018.12.005

Nahangi, M., Chen, Y., McCabe, B., 2018. Data envelopment analysis (dea) for safety-based efficiency evaluation of construction sites, in: J., R. (Ed.), CSCE General Conference 2018, Held as Part of the Canadian Society for Civil Engineering Annual Conference 2018. Canadian Society for Civil Engineering, pp. 193–201.

Naor, M., Bernardes, E.S., Coman, A., 2013. Theory of constraints: Is it a theory and a good one? Int. J. Prod. Res. 51, 542–554. https://doi.org/10.1080/00207543.2011.654137

Nations, U., 2022. Transforming our world: the 2030 Agenda for Sustainable Development [WWW Document]. https://sdgs.un.org/2030agenda. URL https://sdgs.un.org/2030agenda

Nations, U., 2015. Resolution adopted by the General Assembly on 25 September 2015. https://doi.org/10.1057/978-1-137-45443-0_24

Naveh, E., Marcus, A., 2005. Achieving competitive advantage through implementing a replicable management standard: Installing and using ISO 9000. J. Oper.

Manag. 24, 1–26. https://doi.org/10.1016/j.jom.2005.01.004

Nicolaidou, O., Dimopoulos, C., Varianou-Mikellidou, C., Boustras, G., Mikellides, N., 2021. The use of weak signals in occupational safety and health: An investigation. Saf. Sci. 139. https://doi.org/10.1016/j.ssci.2021.105253

Ning, Y., Liu, Z., Ning, Z., Zhang, H., 2018. Measuring eco-efficiency of state-owned forestry enterprises in Northeast China. Forests 9. https://doi.org/10.3390/f9080455

Nissi, E., Rapposelli, A., 2012. Data envelopment analysis (DEA) for safety-based efficiency evaluation of construction sites. AIEL Ser. Labour Econ. 6, 93–101. https://doi.org/10.1007/978-3-7908-2772-9_5

Oh, D., Suh, D., 2022. R Package "nonparaeff": Nonparametric Methods for Measuring Efficiency and Productivity.

Oliveira, R., 2008. The Assessment of Corporate Social Responsibility in the Mining Sector using Data Envelopment Analysis 50.

Oliveira, R., Camanho, A.S., Zanella, A., 2017. Expanded eco-efficiency assessment of large mining firms. J. Clean. Prod. 142, 2364–2373. https://doi.org/10.1016/j.jclepro.2016.11.039

Oliveira, R., Zanella, A., Camanho, A.S., 2020. A temporal progressive analysis of the social performance of mining firms based on a Malmquist index estimated with a Benefit-of-the-Doubt directional model. J. Clean. Prod. 267, 121807. https://doi.org/10.1016/j.jclepro.2020.121807

Oliveira, R., Zanella, A., Camanho, A.S., 2019. The assessment of corporate social responsibility: The construction of an industry ranking and identification of potential for improvement. Eur. J. Oper. Res. 278, 498–513. https://doi.org/10.1016/j.ejor.2018.11.042

Omrani, H., Amini, M., Alizadeh, A., 2020. An integrated group best-worst method – Data envelopment analysis approach for evaluating road safety: A case of Iran. Meas. J. Int. Meas. Confed. 152. https://doi.org/10.1016/j.measurement.2019.107330

OSHA, 2012. Recording and Reporting Occupational Injuries and Illness (Part Number 1904) [WWW Document]. URL https://www.osha.gov/laws-regs/regulations/standardnumber/1904 (accessed 8.26.21).

Pawson, R., Tilley, N., 1997. Realistic Evaluation. SAGE Publications Inc., London.

Piran, F.A.S., 2021a. AVALIAÇÃO DA EFICIÊNCIA ECONÔMICA DOS SISTEMAS PRODUTIVOS A PARTIR DA UTILIZAÇÃO INTEGRADA DA DATA ENVELOPMENT ANALYSIS E DA THROUGHPUT ACCOUNTING. UNISINOS.

Piran, F.A.S., 2021b. Case study (Notes).

Piran, F.A.S., 2015. MODULARIZAÇÃO DE PRODUTO E OS EFEITOS SOBRE A EFICIÊNCIA TÉCNICA: UMA AVALIAÇÃO EM UMA FABRICANTE DE ÔNIBUS. UNISINOS.

Piran, F.S., Camanho, A.S., Silva, M.C., Lacerda, D.P., 2022. Internal Benchmarking for Efficiency Evaluations using Data Envelopment Analysis: A Review of Applications and Directions for Future Research Fabio, in: Advanced Mathematical Methods for Economic Efficiency Analysis. Springer.

Piran, F.S., Lacerda, D.P., Camanho, A.S., Silva, M.C.A., 2021. Internal benchmarking to assess the cost efficiency of a broiler production system combining data envelopment analysis and throughput accounting. Int. J. Prod. Econ. 238, 108173. https://doi.org/10.1016/j.ijpe.2021.108173

Piran, F.S., Lacerda, D.P., Camargo, L.F.R., 2020. Analysis and Management of Productivity and Efficiency in Production Systems for Goods and Services. CRC Press, Boca Raton. https://doi.org/https://doi.org/10.1201/9780429351679

Podgórski, D., 2015. Measuring operational performance of OSH management system - A demonstration of AHP-based selection of leading key performance indicators. Saf. Sci. 73, 146–166. https://doi.org/10.1016/j.ssci.2014.11.018

Porter, L.W., Bigley, G.A., Steers, R.M., 2002. Motivation and Work Behavior, 7th ed. Mc-Graw Hill/Irwin.

Probst, T.M., Estrada, A.X., 2010. Accident under-reporting among employees: Testing the moderating influence of psychological safety climate and supervisor enforcement of safety practices. Accid. Anal. Prev. 42, 1438–1444. https://doi.org/10.1016/j.aap.2009.06.027

R core team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www. R-project. org.

Rae, A., Provan, D., Aboelssaad, H., Alexander, R., 2020. A manifesto for Reality-based Safety Science. Saf. Sci. https://doi.org/10.1016/j.ssci.2020.104654

Ramli, A.A., Watada, J., Pedrycz, W., 2011. Possibilistic regression analysis of influential factors for occupational health and safety management systems. Saf. Sci. 49, 1110–1117. https://doi.org/10.1016/j.ssci.2011.02.014

Ramli, N.A., Munisamy, S., 2013. Modeling undesirable factors in efficiency measurement using Data Envelopment Analysis: A Review. J. Sustentability Sci. Manag. 8, 126–135.

Rantanen, J., Muchiri, F., Lehtinen, S., 2020. Decent work, ILO's response to the globalization of working life: Basic concepts and global implementation with special reference to occupational health. Int. J. Environ. Res. Public Health 17. https://doi.org/10.3390/ijerph17103351

Redinger, C., Dotson, K., Leibowitz, A., 2011. Occupational health and safety management systems, in: Patty's Industrial Hygiene. pp. 2109–2162. https://doi.org/https://doi.org/10.1002/0471435139.hyg049.pub2

Redinger, C.F., 1998. Occupational Health and Safety Management System Conformity Assessment: Development and Evaluation of a Universal Assessment. Univ. Michigan.

Redinger, C.F., Levine, S.P., 1998. Development and evaluation of the Michigan Occupational Health and Safety Management System Assessment Instrument: A universal OHSMS performance measurement tool. Am. Ind. Hyg. Assoc. J. 59, 572–581. https://doi.org/10.1080/15428119891010758

Redinger, C.F., Levine, S.P., Blotzer, M.J., Majewski, M.P., 2002a. Evaluation of an occupational health and safety management system performance measurement tool—ii: Scoring methods and field study sites. Am. Ind. Hyg. Assoc. J. 63, 34–40. https://doi.org/10.1080/15428110208984689

Redinger, C.F., Levine, S.P., Blotzer, M.J., Majewski, M.P., 2002b. Evaluation of an occupational health and safety management system performance measurement tool—iii: Measurement of initiation elements. Am. Ind. Hyg. Assoc. J. 63, 41–46. https://doi.org/10.1080/15428110208984690

Robson, L.S., Clarke, J.A., Cullen, K., Bielecky, A., Severin, C., Bigelow, P.L., Irvin, E., Culyer, A., Mahood, Q., 2007. The effectiveness of occupational health and safety management system interventions: A systematic review. Saf. Sci. 45, 329–353. https://doi.org/10.1016/j.ssci.2006.07.003

RStudio, 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA.

Said, S.M., Sanwari, S.R., Said, F., 2013. Technical and scale efficiency in malaysian manufacturing industries in the presence of industrial accidents. World Appl. Sci. J. 24, 862–871. https://doi.org/10.5829/idosi.wasj.2013.24.07.13229

Saracino, A., Antonioni, G., Spadoni, G., Guglielmi, D., Dottori, E., Flamigni, L., Malagoli, M., Pacini, V., 2015. Quantitative assessment of occupational safety and health: Application of a general methodology to an Italian multi-utility company. Saf. Sci. 72, 75–82. https://doi.org/10.1016/j.ssci.2014.08.007

Sarkis, J., 2002. Preparing your data for DEA, in: J., Z., W.D., C. (Eds.), Productivity Analysis in the Service Sector with Data Envelopment Analysis. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-71607-7_17

Scheel, H., 2001. Undesirable outputs in efficiency valuations. Eur. J. Oper. Res. 132, 400–410. https://doi.org/10.1016/S0377-2217(00)00160-0

Schein, E.H., 1990. Organizational Culture. Am. Psychol. 45, 109–119. https://doi.org/10.1037/0003-066X.45.2.109

Schein, E.H., Bennis, W.G., 1965. Personal and organizational change through group methods: The laboratory approach. Wiley, New York.

Seyedalizadeh Ganji, S.R., Rassafi, A., Xu, D.-L., 2019. A double frontier DEA cross efficiency method aggregated by evidential reasoning approach for measuring road safety performance. Meas. J. Int. Meas. Confed. 136, 668–688. https://doi.org/10.1016/j.measurement.2018.12.098

Seyedalizadeh Ganji, S.R., Rassafi, A.A., 2019. Measuring the road safety performance of Iranian provinces: a double-frontier DEA model and evidential reasoning approach. Int. J. Inj. Contr. Saf. Promot. 26, 156–169. https://doi.org/10.1080/17457300.2018.1535510

Shah, S.A.R., Ahmad, N., Shen, Y., Pirdavani, A., Basheer, M.A., Brijs, T., 2018. Road safety risk assessment: An analysis of transport policy and management for low-, middle-, and high-income Asian countries. Sustain. 10. https://doi.org/10.3390/su10020389

Shen, Y., Hermans, E., Bao, Q., Brijs, T., Wets, G., 2020. Towards better road safety management: Lessons learned from international benchmarking. Accid. Anal. Prev. 138. https://doi.org/10.1016/j.aap.2020.105484

Shen, Y., Hermans, E., Ruan, D., Vanhoof, K., Brijs, T., Wets, G., 2010. A DEA-based malmquist productivity index approach in assessing road safety performance, in: Ruan D. Kerre E.E., R.D.C.G.L.T.X.Y. (Ed.), Computational Intelligence

Foundations and Applications - Proceedings of the 9th International FLINS Conference, FLINS 2010. World Scientific Publishing Co. Pte Ltd, pp. 923–928. https://doi.org/10.1142/9789814324700_0141

Shirali, G.A., Salehi, V., Savari, R., Ahmadiangali, K., 2018. Investigating the effectiveness of safety costs on productivity and quality enhancement by means of a quantitative approach. Saf. Sci. 103, 316–322. https://doi.org/10.1016/j.ssci.2017.12.017

Simon, H.A., 1996. The science of the artificial, 3rd ed. Massachussetts Institute of Technology Press.

Simon, H.A., 1957. Models of man. Wiley, Hoboken, NJ.

Skład, A., 2019. Assessing the impact of processes on the Occupational Safety and Health Management System's effectiveness using the fuzzy cognitive maps approach. Saf. Sci. 117, 71–80. https://doi.org/10.1016/j.ssci.2019.03.021

Strauss, A., Corbin, J., 1990. Basics of Qualitative Research: Grounded Theory Procedures and Techniques, 2nd ed. SAGE Publications Inc., Newbury Park.

Su, T.-S., Tsai, W.-Y., Yu, Y.-C., 2005. An integrated approach for improving occupational health and safety management: The Voluntary Protection Program in Taiwan. J. Occup. Health 47, 270–276. https://doi.org/10.1539/joh.47.270

Swanson, R.A., Chermack, T.J., 2013. Theory Building in Applied Disciplines. San Francisco, CA. https://doi.org/10.1108/jchrm-06-2013-0023

Tan, P.-N., Steinbach, M., Karpatne, A., Kumar, V., 2019. Introduction to Data Mining, 2nd ed. Pearson.

Tan, P.N., Steinbach, M., Karpatne, A., Kumar, V., 2018. Introduction to Data Mining, 2nd ed. Pearson, New York.

Terziovski, M., Power, D., Sohal, A.S., 2003. The longitudinal effects of the ISO 9000 certification process on business performance. Eur. J. Oper. Res. 146, 580–595. https://doi.org/10.1016/S0377-2217(02)00252-7

Terziovski, M., Samson, D., 1999. The link between total quality management practice and organisational performance. Int. J. Qual. Reliab. Manag. 16, 226–237. https://doi.org/10.1108/02656719910223728

Tešić, M., Hermans, E., Lipovac, K., Pešić, D., 2018. Identifying the most significant indicators of the total road safety performance index. Accid. Anal. Prev. 113, 263–278. https://doi.org/10.1016/j.aap.2018.02.003

Thomé, A.M.T., Scavarda, L.F., Scavarda, A.J., 2016. Conducting systematic literature review in operations management. Prod. Plan. Control 27, 408–420. https://doi.org/10.1080/09537287.2015.1129464

Thore, S., Tarverdyan, R., 2009. Using data envelopment analysis to quantify ILO objectives and identify policies conducive to decent work in a globalizing world. Socioecon. Plann. Sci. 43, 151–164. https://doi.org/10.1016/j.seps.2008.06.001

Uhrenholdt Madsen, C., Kirkegaard, M.L., Dyreborg, J., Hasle, P., Uhrenholdt, C., Louise, M., Dyreborg, J., Hasle, P., 2020. Making occupational health and safety management systems 'work': A realist review of the OHSAS 18001 standard. Saf. Sci. 129, 104843. https://doi.org/10.1016/j.ssci.2020.104843

Verma, A., Khan, S.D., Maiti, J., Krishna, O.B., 2014. Identifying patterns of safety related incidents in a steel plant using association rule mining of incident investigation reports. Saf. Sci. 70, 89–98. https://doi.org/10.1016/j.ssci.2014.05.007

Vilela, J., Castro, J., Martins, L.E.G., Gorschek, T., 2017. Integration between requirements engineering and safety analysis: A systematic literature review. J. Syst. Softw. 125, 68–92. https://doi.org/10.1016/j.jss.2016.11.031

Vinodh, S., Aravindraj, S., 2015. Benchmarking agility assessment approaches: A case study. Benchmarking 22, 2–17. https://doi.org/10.1108/BIJ-04-2013-0037

Wang, Y., Chen, H., Liu, B., Yang, M., Long, Q., 2020. A systematic review on the research progress and evolving trends of occupational health and safety management: A bibliometric analysis of mapping knowledge domains. Front. Public Heal. 8. https://doi.org/10.3389/fpubh.2020.00081

WBCSD, 2011. Eco-Efficiency: creating more value with less impact, 1st ed. WBCSD, Geneva.

Weick, K.E., 1995. What Theory is Not, Theorizing Is. Adm. Sci. Q. 40, 385. https://doi.org/10.2307/2393789

What is R? [WWW Document], 2021. URL https://www.r-project.org/about.html (accessed 6.2.21).

Whetten, D.A., 1989. What constitutes a Theoretical Development? Acad. Manag. Rev. 14, 490–495.

WHO, ILO, 2021. WHO/ILO Joint Estimates of the Work-related Burden of Disease and Injury, 2000–2016.

Wilcoxon, F., 1945. Individual Comparison By Ranking Methods. Biometrics Bull. 1, 80–83.

Wohlin, C., 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. ACM Int. Conf. Proceeding Ser. https://doi.org/10.1145/2601248.2601268

Yeh, L.-T., 2017. Incorporating workplace injury to measure the safety performance of industrial sectors in Taiwan. Sustain. 9. https://doi.org/10.3390/su9122241

Yin, R.K., 1994. Case study research: design and methods, 2nd ed. SAGE Publications Ltd, Beverly Hills, CA.

Zhang, S., Zhang, C., 2002. Association Rule Mining - Models and Algorithms. Lecture Notes in Artificial Intelligence, Springer. Springer, Berlin.

This page is intentionally left blank.

“If I have seen further, it is by standing on the shoulders of giants”

Sir Isaac Newton (1675)