Master's thesis

2022

Eivind Larsen

NTNU
Norwegian University of
Science and Technology
Faculty of Economics and Management
Department of Industrial Economics and Technology
Management

Master's thesis

Eivind Larsen

Safety performance indicators in shipping

A study of current application and potential contribution to improved safety management

February 2022

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology

Norwegian University of
Science and Technology

Safety performance indicators in shipping

A study of current application and potential contribution to improved safety management

Eivind Larsen

IØ6901 - Organisasjon og ledelse, masteroppgave
Submission date: February 2022
Supervisor: Trond Øystein Kongsvik
Co-supervisor: NA

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

ABSTRACT

The shipping industry is a key actor in the global society with regard to transportation of goods and personnel. The extent and nature of the industry creates considerable risk related to safety, assets and environment, as well as commercial aspects, affecting a broad spectre of stakeholders. Although, safety performance has developed in the positive direction through the past decades, the need to maintain a proactive safety management remains. Performance indicators are commonly used by shipping companies as part of their safety management systems. The resent years increased attention on how the industry and individual shipping companies manage risk, have actualized the concept of performance indicators.

The main objective of this thesis is to explore how application of performance indicators can improve safety management in shipping. In order to do this, an understanding of how shipping companies currently apply indicators have been sought. This have been done through a multiple case-study involving four companies within different segments. In addition, an evaluation of how relevant theory can be applied in the development and application of performance indicators, have been carried out. The included theories, metho and models address both the linear and non-linear aspects of a socio-technical system. Perspectives of Resilience Engineering have been emphasized and is considered as a valuable supplement to the more traditional view of safety.

The case studies revealed that there are many similarities with regard to type of indicators and how they are applied as part of the management system. In general, it was perceived that application of indicators are a valuable constituent in the management of safety. By looking at the applied set of indicators, combined and individually, the majority are of the reactive type, focusing on accidents, injuries and deviations. The context of the individual company, and of the shipping industry as whole, is considered to have influence on the application. The regulatory framework and high attention to compliance, as an attribute of the general shipping context, is one example of this.

The thesis shows how different methods can be applied to develop indicators, individually and combined. They can be divided in two main categories; those that focus on safety promoting factors and those that focus on important activities and processes, including aspects of dynamics and variability. They are proactive oriented and can support the development of more proactive indicators and improved safety management, amongst the case study companies and the shipping industry as a whole. It is also shown how methods and models can be applied to assess and obtain a good balance within the set of indicators, focusing on correlation between well distributed indicators and the ability to see both safety-related and economy-related indicators in the same picture, avoiding inappropriate efforts on one at the expense of the other.

TABLE OF CONTENT

TABLES

FIGURES

1 INTRODUCTION

1.1 Background

The shipping industry is a key actor in the global society, carrying around 90% of the total traded goods (OECD, 2022) and employing about 1,9 million seafarers (UNCTAD, 2021). The extent and nature of the industry creates considerable risk related to safety and environment, as well as commercial aspects and the global value chains, affecting a broad spectre of stakeholders. How the industry manages its risk, is thus both an internal and external concern.

When it comes to safety performance in the shipping industry, at least in the sense of loss prevention, it shown positive development through the past decade, where the total number of lost vessels have declined by 68 % (Allianz, 2020). Also the rate of vessels detained for serious safety related deficiencies have in general decreased globally, although with variation from year to year and within the different Port State Control regions (Paris MoU, 2020; US Coast Guard, 2020; Tokyo MoU, 2020; BS MoU, 2020). Still, incidents happen with negative impact on humans, assets and environment. In 2019, a total of 41 merchant vessels was lost globally with a corresponding fatality of 2 815 (Allianz, 2020). The same year, 251 personal injuries were reported on Norwegian flagged vessels (Sjøfartsdirektoratet, 2020). Looking at the environmental side of it, shipping accounts for about 2,5% of the global greenhouse gas emissions.

Shipping is a highly regulated industry (Bloor, et al., 2013) through a broad set of international and national legislation which aim to establish an acceptable risk level related to both safety and environment. Within this regulatory framework, the introduction of the International Safety Management Code (ISM Code) (IMO, 1993) represented a shift from a solely compliance-based approach to more self-regulated approach. Meaning that the shipping companies is required to develop and implement a safety management system to ensure achievement of defined safety objectives, including compliance with rules and regulations. The effect of ISM code has been discussed in several studies, which conclude both a positive (Tzannatos & Kokotos, 2008) and a negative effect (Bhattacharya, 2011). According to a recent survey performed among seafarers on Norwegian vessels in 2020, it is strongly agreed that the safety management system is important to ensure safety on board (Sjøfartsdirektoratet, 2020). Irrespective of the current effects of the ISM Code and the implemented safety management systems, they remain an important instrument to promote safety in shipping. In this context, the responsibility for implementing the Code, rest with the vessels Flag States - through enforcement and verification - and the shipping company - through development and implementation.

The objective of the ISM Code and the safety management systems is to enable effective implementation of the company safety and environmental protection policy, i.e., to support safety performance. Assessment of the effectiveness of safety management systems and safety performance are carried out through periodical audits, as the authority's primary means. In addition, shipping companies conduct internal audits and periodical reviews of the system to evaluate its effectiveness. The output from audits, both external and internal, and periodical reviews, should be improved safety management systems and safety performance. Studies indicate that the intended improvements are not fully achieved through the current audit approach, as the focus on documentation and compliance, hamper the shipping companies in developing useful and effective safety management systems (Størkersen K. V., 2018). In addition, the shipping companies are subject to audits and inspections from other stakeholders, such as cargo owners/associations, cost- and port authorities, which add additional focus on specific requirements and compliance.

The common area of interest, for both the shipping company and its stakeholders, is to have confidence that the safety performance acceptable. In this respect, it may be relevant to study the potential value in application of performance indicators to build confidence to and improve safety performance in shipping. Although performance indicators are not explicitly required by the ISM Code, they are commonly used in the shipping industry as means to provide insight to the performance within key areas such as safety.

The attention to performance measures have been actualized in parallel with global developments. For instance, the heightened focus on environment and sustainability have led to new reporting requirements, such as the ESG (Environmental, Social, Governance) framework (EU, 2021), the IMO DCS on reporting on fuel oil consumption (IMO, 2016) and the similar EU MRV (EU, 2015) reporting requirements. The development within digitalization have also increased the availability and utilization of data. The Norwegian Maritime Authority have since 2017 applied a risk-based approach to inspection and audit activities for Norwegian vessels (Sjøfartsdirektoratet, 2017). The risk model include data related to deficiencies, detentions, age and accidents for individual vessels, i.e., possible measures of safety performance.

1.2 Research questions and objectives

Provided the constant need to maintain an acceptable safety level, the increased attention to performance and the potential improvements of assessment approaches, the potential value in application of performance indicators to build confidence to, and improve safety performance in shipping, is considered relevant. The main question for this thesis is:

How can application of performance indicators improve safety management in shipping?

In order to answer the main question, the following research questions are formulated:

1. How do shipping companies apply performance indicators as part of their safety managements systems?
2. How can safety science theory be used to develop indicators?
3. What recognize “good” balance between different types of indicators?

The primary objective of this thesis is to provide insights that can be used as basis for evaluation of existing performance indicators, and for development of potential new indicators to support improved safety performance. Indirectly, this may set focus on other parts of the safety management system that has potential improvements. Secondary, as safety performance coincide with the objective of audit and verification approaches, the same insights should be relevant to apply in considerations of current practices, both internally in the shipping companies and other actors in the industry.

The motivation behind the subject of the thesis is largely based on my professional background, including about 15 years' experience with i.a. safety management system certification within the shipping industry.

2 RESEARCH DESIGN

Research design can be defied as the logical sequence that connect the empirical data to a study's initial research questions and, ultimately, to its conclusion. Or in other words, a logical plan for getting from *here* to *there*, where *here* may be the defined as the set of questions to be addressed, and *there* is some set of conclusions about these questions (Yin, 2018). Research design can also be seen as a framework for collection and analysis of data (Bryman, 2016). The research design for this thesis is based on a set of generic phases that may be valid for any research projects (Jacobsen, 2018). The different phases are illustrated in Figure 1 and further elaborated in the subchapters.

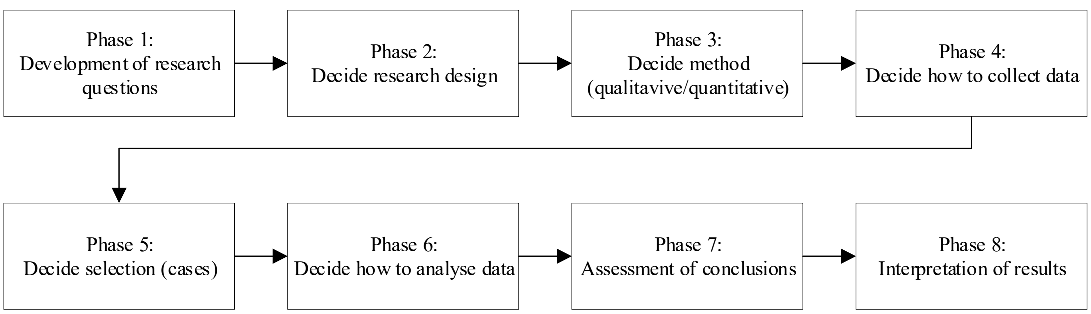


Figure 1: Phases of the research approach, reproduced from (Jacobsen, 2018)

Based on the overall problem formulation and the subsequent research questions, the selected research design incorporates a multiple case-study with the application of semi structured interviews, and a literature review. Hench, a qualitative method for collection of empirical data is applied.

In the process of defining the main question and research questions, a broader sampling approach was considered, including a survey design with use of questionaries. Although a broad sampling could provide valuable data with regard to generalization, a more in-depth approach was considered more appropriate for the thesis main question. Thus, the research questions were re-formulated accordingly.

2.1 Literature review

The objective of the literature review was to establish an understanding of the relevant concepts and terms associated with the main question of the thesis. The literature review was carried out in two stages, an initial review as part of the main question formulation and a subsequent through review. The literature search has been carried out by the use of the NTNU databases/library and Google Scholar. Citations in relevant publications have also been a valuable source for identification of other relevant publications and sources. The aim of the literature search was to identify sources that has both a general and domain specific relevance for the thesis research questions. It was discovered that the majority of research related to safety performance indicators was mainly done within high-risk industries, such as the nuclear power generation, chemical processing, aviation and offshore petroleum. Despite this, examples of relevant research within the shipping industry exists. Although the majority of research is related to other industries than the shipping industry, the theory and principles are to a large degree considered to be adoptable also for the shipping industry.

2.2 The case studies

A cases study is an empirical method that investigate a contemporary phenomenon (the "case") in depth and within its real-world context, especially when the boundaries between the phenomenon and the context may not be clearly evident (Yin, 2018). Case studies can be of a single-case or multiple-case design, where a multiple-case study occurs when the number of examined cases exceed one. The main argument in favour of a multiple-case study is that it improves theory building. By comparing two or more cases, the researcher is in a better position to establish the circumstances that a theory will or will not hold (Bryman, 2016).

The main question that is sought answered in this thesis is how application of performance indicators can improve safety management in shipping. In order to do say something meaningful about improvement potentials, an understanding about current application is needed. This understanding is sought through the case-studies where the main objective is to establish an understanding of how shipping companies apply performance indicators as part of their safety management system (the phenomenon that is subject for investigation). As the subject in this formulation is shipping companies (plural) it forms a logical argument for selecting more than one cases, i.e., a multiple-case design. Thus, there is an aim to provide external validity and a multiple-case study is therefore considered more appropriate than a single-case design. The external validity will in any case be limited, as the context of shipping companies (globally) is far too complex compared to the context represented by the selected cases.

The selected cases consist of four (4) Norwegian shipping companies. The companies referred to as Case 1 and Case 4 operates within the passenger transport segment. The company referred to as Case 2 and Case 3 operates within the oil/chemical tanker segments. The size of their fleets is in the range of 15 to 50 vessels. By this selection, it was possible to compare the cases within the same segments and the cases crosssegment. All cases have been subject to the same method as described in Chapter 3.3 and 3.4. The cases are numbered based on the order of the performed interviews.

2.3 Interviews

Qualitative interviewing was selected as method for collecting empirical data through the cases. As the main objective of the interviews was to establish an understanding of how the individual shipping company apply performance indicators as part of their safety management system, the form of the interviews was considered to factual (Kvale & Brinkmann, 2021). A factual interview does not only focus on the respondents' own perspectives and believes but emphasises on collecting valid and factual information on the subjects in questions. The interviews were further designed as semi-structured interview, which mean that it follows an interview guide with defined topics, a preferred sequence and some closed questions (Jacobsen, 2018). The same guide was used for all interviews.

The interview guide (Table 1) was prepared based on the performed literature review in order to identify key aspects related to application of performance indicators, both with regard to how indicators are developed and used within the company. The guide was structured as a set of main questions with related follow-up questions. Prior to the interviews, the current applied set of performance indicators was submitted for review and was further used as part of the guide during the interviews.

The respondents were selected based on the ability to provide an accurate and complete response to the sum of questions, subsequently how the company apply performance indicators. Relevant positions that could be involved was proposed and communicated to company prior to the interviews. In total, the interviews involved 10 respondents in management positions (Managing director, QHSE manager, HR manager, Technical manager and Operations manager).

The interviews were conducted at the company's location and by use of digital meeting. The interviews were documented by making notes on a pre-prepared log based on the interview guide. This in order to ascertain that all the main questions were provided with an appropriate response. The use of recording was considered, but it was concluded to more feasible to focus on obtaining a well formulated notes during the interviews as a continual quality check of the result.

Table 1: Interview guide

Main questions	Supporting questions
What are the main objectives/motivation for applying indicators?	
How is the process leading to this set of indicators?	Is it based on a defined and documented methodology (e.g. risk based)? Major accidents risk vs. personnel risk? Proactive vs. reactive How are the indicators linked to the overall objectives and policies? Who is involved? Are the applied indicators influenced by external stakeholders (customers, owners, authorities)? If so, how and which indicators?

	Are there relevant indicators that have been identified/considered but for some reasons are not monitored? How are time/resource needs, available tools etc. related to monitoring influencing the composition of indicators (total number, type, etc.)
Which criteria represent "good" indicators?	
How are the indicators monitored?	How often are sampling/reporting carried out? Which tools/methods are used in the monitoring? What is the time perspective (long/short term) for the individual indicators?
How are the indicators used?	Basis for initiate actions, changes? Any examples How are targets set? How is status of indicators communicated in the Company? Are any of the indicators used as incentive measures? If yes; which indicators and what incentives, including experienced effect? How often is the indicator set reviewed and changed?
To which degree do you feel that the set of indicators provide a representative picture of the effectiveness of the safety management system?	
Are there any negative effects (experienced/anticipated) related to application of performance indicators?	

2.4 Analysis of data

The main objective of the data analysis is to extract information from the data that is meaningful in relation to the main thesis question and research questions. The collected data from the cases studies consist of the documented set of indicators and notes from the interviews performed within the respective companies. The data analysis follows two different approaches. The applied indicators are analyzed by adapting a defined model (ref. Chapter 4.1) in order to categorize individual indicators with regard to type, including distinction between proactive and reactive. The analysis focus on both the individual indicators and indicators set for respective cases and the indicators combined. For the interview data, a thematic analysis approach has been applied. Thematic analysis is a commonly used approach to analyze qualitative data, such as interview records or notes. The objective is to identify relevant themes that emerge from the data. A relevant theme is defined as something that has a certain level of pattern or meaning in relation to the research questions (Bryman, 2016). The analysis follows six basic steps:

1. Perform a readthrough to get familiarized with the content
2. Begin coding of the records
3. Elaborate many of the codes into themes
4. Evaluate the higher-order codes ore themes/sub-themes
5. Examine possible links and connections between concepts and/or how the concepts vary in terms of features of the case.
6. Write up the insight from the previous stages to provide a compelling narrative about the data.

The initial step in the approach was to compile the notes from the interviews and perform a thorough readthrough to get an overall picture of the content. The following step was to code the content. In practice this was done as open coding, which means that the coding was developed simultaneously through the process. The main focus through this process was to capture parts of the content that have relevance to research questions, including looking for recurring topics. Through this process, it was acknowledged that the applied interview guide had considerable influence on the feedback and corresponding records. Thus, the developed codes showed great resemblance to the topics and questions in the guide. It was further experienced that the coding process overlapped with the subsequent step which was to identify relevant themes. Based on this process, it was identified two main aspects of the relevant parts of the content; (1) the respondents perceptions related to the relevant subjects and (2) the factual descriptions regarding how they apply indicators. The main themes that were identified reflects these aspects with a set of corresponding sub-themes, listed in Table 2.

The objective of the thesis is to answer the main question: How can application of performance indicators improve safety management in shipping? In order to be able to say something about improvement potentials, an understanding of current practices need to be established. This is the motivation for the research question: How do shipping companies apply performance indicators as part of their safety managements systems? The themes address the core aspects of this question. The theme "development and use" are considered more factual oriented than the two others. These themes are relevant in the sense that it shed light on current application, as it is assumed that organizations perceptions of the key concepts and perception of value have an influence on how things are done.

The themes also capture key aspects of the theory that is found to be relevant to the thesis main question and research questions. In sum, the themes are aimed to support a narrative about: overall perceptions and motivations, how things are done, the perceive value, improvement potentials and potential pitfalls. In writing up of the findings, citations have been included and are given reference to the applicable case by entering the number in brackets.

Table 2: Themes and sub-themes

Perceptions of key concepts	Development and use	Perception of value
Objectives	Influence by external stakeholders	Improvement potentials
Characteristic of good indicators	Monitoring	Potential pitfalls
Types of indicators	Targets	
	Actions	
	Incentives	

2.5 Data quality

2.5.1 Internal validity

Internal validity relates to mainly to the issue of causality and the question of whether a conclusion that incorporate a causal relationship between two or more variables holds water (Bryman, 2016). In addition, internal validity, relates to the concern about how well the data represent the reality, represented by three questions; to which degree have respondents provided a truthful description of the reality; to which degree is the researchers representation and interpretation of data correct, and to which degree is the findings and conclusions drawn by researcher representing reality (Jacobsen, 2018).

The overall problem formulation and the research questions for the thesis does not seek conclusion related to causalities (as a main objective). For the cases-studies and the interviews, the main objective was to establish a factual understanding of how the individual company apply performance indicators as part of their safety management system. In order to ascertain that the description obtained from the respondents reflects reality, (Jacobsen, 2018) argue that it is important to question whether the correct sources have been used/made available, whether the sources provide correct information and when in the research process are the information collected. The relevant respondents for the interviews were selected based on their involvement and knowledge of how performance indicators are applied in the company. It was also requested to have 2-3 respondents involved, as information from more than one respondent increase the validity of the sum of information that is collected (Jacobsen, 2018). As the topics for the interviews can be considered as less sensitive/controversial and not personal oriented, the respondent's motivation to provide untruthful information was considered as low.

When it comes to where in the research process the information is collected, it should be acknowledged that the researcher will gain new knowledge which can result in a change in focus compared to initial assumptions of the research process. This can both represent a strength and a weakness, where increased knowledge can result in more focus on important elements (based on initial assumptions), and at the same time there is a risk of getting blind for eventually new elements that could be valuable for validity of the research. It is though argued that the data gathered at a later stage in the research process are the best. The aim of the developed interview guide was to ensure that it was possible to maintain a relative uniform focus during the series of subsequent interviews and at the same time allow for some degree of unconstrained information sharing. This was also obtained during the performed interviews, without identified need for changes in the interview guide. By comparing the first and the last interview it was recognized that the results differ. This was mainly related to the way the interview guide was used with regard to ensure sufficient time to address the main questions. Still the results from all the interviews are considered valid. The interview results are based on a defined approach (thematic analysis) which aims to provide both a realistic and relevant representation. In addition, the records from each interview were submitted to respondents for verification of correctness.

2.5.2 External validity

External validity deals with problem of knowing whether a study's findings area generalizable beyond the immediate study/research context (Bryman, 2016; Jacobsen, 2018; Yin, 2018). Generalization can be done in two ways, analytical generalization and statistical generalization, where the latter is the most common. In statistical generalization inference is made about a population (universe) on the basis of empirical data collected form a sample of the population (Yin, 2018). In case-studies statistical generalization is not relevant as the case/cases cannot be considered “samples” and is far too few in numbers to serve as an adequately sized sample to represent a larger population. The aim for this thesis is thus to utilize the potential for analytical generalization, where the empirical data form the case studies will be used to shed light on the theoretical concepts and principles identified as relevant for the main research question.

2.5.3 Reliability

Reliability is concerned with the question of whether the results of a study are repeatable and is especially an issue in connection with quantitative research (Bryman, 2016). In case-study research, this means studying the same case over again, not just replicating the result of the original case study by studying another case (Yin, 2018). According to Yin (2018) opportunities for repeating case studies rarely occur. However, documenting the procedures followed for the case(s) should be done to allow for possible repetition. The core procedure for the case studies in this thesis is the documented interview guide. In addition, the description of selection of cases and respondents and how the information is presented and analysed.

3 THEORY

3.1 Introduction

The objective of this chapter is to present theory that is relevant to main question of the thesis and provide insight to key concepts and support answering the research questions. The main question actualizes three central concepts, namely, *safety*, *safety management* and *performance indicators*. In order to manage and monitor safety, we need to have clear understanding of what safety is and how it is achieved. This understanding is sought in Chapter 4.2 and includes a brief description of developments in safety research, including methods and models and how they can be categorized. This is considered valuable, as they have strengths and limitations when used as basis for “safety understanding” depending on the context and system in question. In addition, the chapter include a description of current influential theories, such as resilience engineering, seen as a valuable contribution to the traditional view of safety. Chapter 4.3 focus on safety management and the ISM Code which is a central part of the regulatory framework of the shipping industry. Chapter 4.4 describes the concept of performance indicators, including examples of models that can be applied in the development of performance indicators. These methods and models should be seen in the light of the perspectives included in Chapter 4.2.

3.2 Safety theory, methods and models

Safety theory, as term, refers to a broad spectre of perspectives on how safety can be understood and how safety is influenced under various contexts. The different perspectives, or line of thoughts, forms the basis for a variety of specific models and methods that can be found in the literature. When navigating the available literature, it is therefore important that different models and methods are understood in the light of the underlying perspectives, in order to decide their suitability and enable their application. A common objective of the theories, methods and models found in the literature seek to explain how accidents happen and are often referred to as accident theory, accident models, and so forth. This coincides with a common understanding of safety as “the absence of unwanted outcomes such as incidents or accidents”. Thus, when it comes to use of terms, safety theory and accident theory are concerned with the same issues.

3.2.1 Developments in safety research

Accident theories and models have developed over time and are driven by influential accidents and developments in technology and in organizations (socio-technical systems). Figure 2 illustrate this development together with influential accidents, back to 1931 represented by the domino model, developed by Heinrich (1931). The developments can be divided into five periods, where each are recognized by different perspectives (Kongsvik, et al., 2018). Throughout the development, past research has been supplement by new perspectives and many of the theories form earlier periods are still valid today.

Figure 2: Accidents and developments in safety research (Herrera, 2012)

3.2.2 Categorization of methods and models

The different accident theories and models that have been developed over the years can be divided into two main categories, linear and non-linear (Gui, et al., 2019). Linear models usually examine the causes of various stages of an accident (such as direct causes, indirect causes, root causes) and form a chain according to the logical sequence, so that a human can clearly see the various causes of an accident and the relationship between them. Such linear models are also frequently referred to as epidemiological models that emphasize the significance of latent failures, or system pathogens (Reason, 1997). Non-linear models generally focus on analysing one or a few factors of an accident, and do not logically distinguish the cause of the accident at different stages. Hollnagel (2012) have made a similar categorization by referring to three accident philosophies (Figure 3). Compared to categorization by Gui, et al (2019), Hollnagel have divided the linear category in two: simple linear and complex linear. In addition, Figure 3 include brief but illustrative description of what recognize the different philosophies and how they differ with regard to basic principles, purpose of investigation and focus of recommendations.

	Basic principles	Purpose of investigation	Focus of recommendations
Simple, linear model (for example, Domino)	Causality (single or multiple causes)	Find specific causes and cause–effect links	Eliminate causes and cut cause–effect links
Complex linear model (for example, Swiss cheese)	Latent conditions, hidden dependencies	Find combinations of unsafe acts and latent conditions	Strengthen barriers and defences
Non-linear (systemic) model	Dynamic couplings, functional resonance	Find tight couplings and complex interactions	Monitor and manage performance variability

Figure 3: Accident philosophies categories (Hollnagel E. , 2012)

3.2.3 Methods and models and system complexity

As described above, the developments in safety research are influenced by the developments in technology, organizations, and context. At the early age of the development, the socio-technical systems can be considered as relatively simple, assembly lines, mechanical and with low complex organizational structures. Today's organizations differ with regard complexity, depending on the industry domain they are part of. Perrow (1999) illustrate the diversity of complexity through a coupling-interaction diagram where the coupling can be described on scale from loose to tight, and the interaction on a scale from linear to complex (Figure 4). Perrow's key argument was that the current society and the technological development had become so complex that accidents had become inevitable and normal rather than rare occasions. The work of Perrow include an analysis of a relatively large number of accidents within different industries and based on this, he plotted the different industries in the coupling-interaction diagram. Marine transportation, which is the subject industry for this thesis, is plotted in the first quadrant.

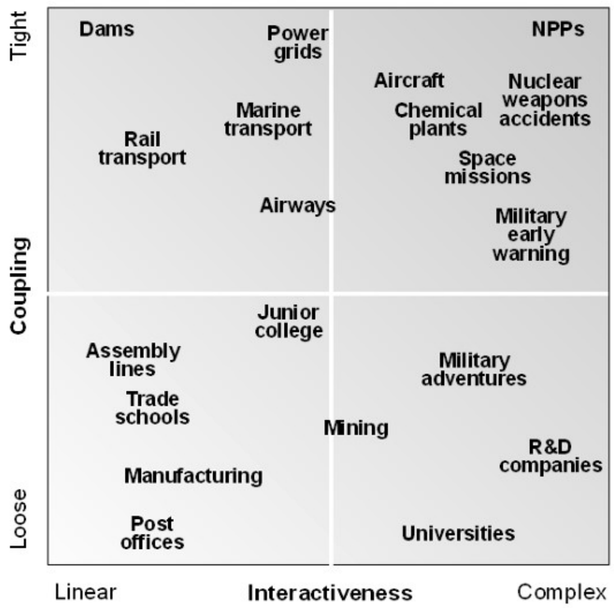


Figure 4: Perrow's coupling-interaction chart, updated by Hollnagel E. (2009)

The different accident models that are found in the literature today, e.g., Gui, et al. (2019) have strengths and weaknesses with regard to how they cope with complexity in the socio-technical systems. It is argued that the traditional chain-of-event models (linear models) are limited in their ability to handle accidents in complex systems (Hollnagel, Woods, & Leveson, 2006). Based on this, an emergent question is how to decide which model (and underlying understanding) is best suited given the industry in question (i.e., shipping) and the specific organization? A contribution to an answer can be found in a report by Hollnagel & Speziali (2008) which include a study of current accident investigation methods and their suitability with respect to the socio-technical system and situation in question. The different models are further linked to the coupling-interaction diagram developed by Perrow as shown in Figure 5. The report does not conclude that there is one ideal model that fits all conditions, and it is emphasized that today's models need to be updated in parallel to the continuous development and increased complexity of the socio-technical systems. It is further pointed out that individual socio-technical systems can comprise both liner and non-linear relationships (Hollnagel E. , 2014). This underlines the point that there is no ideal model that fits all conditions, hench the need to apply different models and perspectives. That would also count for a shipping company. When it comes to addressing complexity or non-linear aspects of the system, the perspectives of Resilience Engineering can be a valuable contribution.

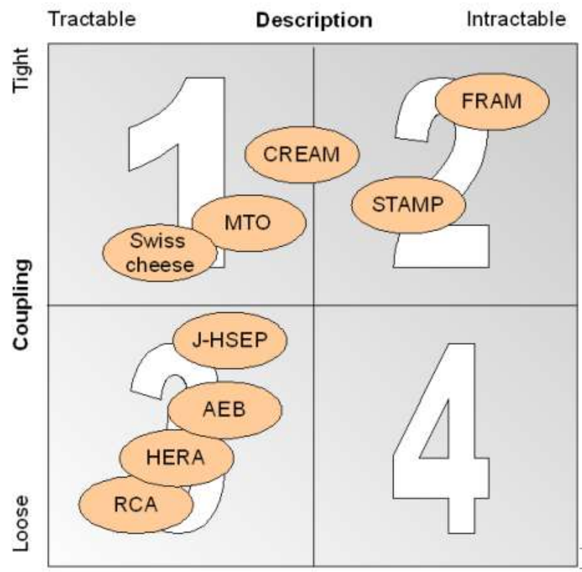


Figure 5: Characterization of accident investigation methods related to Perrow's coupling-interaction chart (Hollnagel & Speziali, 2008)

3.2.4 Resilience Engineering

Resilience Engineering (RE) has been an influential contribution to the safety research during the recent years (Patriarca, Bergström, Di Gravio, & Costantino, 2018). RE is a paradigm of safety management that focuses on systems coping with complexity and balancing productivity with safety (Hollnagel, Woods, & Leveson, 2006). RE aims at providing tools to proactively manage risk, acknowledging the inherent complexity of system functioning and the corresponding need for performance variability. This perspective becomes crucial if linked to the risk-related needs of current socio-technical systems. A socio-technical system can be understood based on the premiss that an organization is a combination of social and technical parts that must work together in order to accomplish the organizations tasks (Appelbaum, 1997). In these systems, safety is not a constant or permanent property; its presence or absence is a continuous function, i.e., it emerges from, the interactive properties and activities of its constituent components (Patriarca, Bergström, Di Gravio, & Costantino, 2018). One measure of resilience is ability to crate foresight, i.e., to anticipate the changing shape of risk that arise as organizations and its environment changes, and to act before failure and harm occurs (Hollnagel, Woods, & Leveson, 2006). Thus, RE and its objective of crating foresight crates a contrast to theories and practices that represent hindsight. The two perspectives referred to as "Safety I" and Safety II" (Hollnagel E. , 2014) can be seen as representation of this contrast, where the Safety I perspective represent hindsight and the Safety II perspective coincide with RE perspective. By comparing these two perspectives, the essence of both perspectives is effectively visualized.

3.2.5 Safety I and Safety II

According to Hollnagel E. (2014) the perspective of Safety I can be described as safety management through centralized control, while Safety II can be described as safety management through guided adaptability. Figure 6 show how the two perspectives differ with regard to some key aspects.

	Safety–I	Safety–II
Definition of safety	As few things as possible go wrong.	As many things as possible go right.
Safety management principle	Reactive, respond when something happens, or is categorised as an unacceptable risk.	Proactive, continuously trying to anticipate developments and events.
Explanations of accidents	Accidents are caused by failures and malfunctions. The purpose of an investigation is to identify causes and contributory factors.	Things basically happen in the same way, regardless of the outcome. The purpose of an investigation is to understand how things usually go right as a basis for explaining how things occasionally go wrong.
Attitude to the human factor	Humans are predominantly seen as a liability or a hazard.	Humans are seen as a resource necessary for system flexibility and resilience.
Role of performance variability	Harmful, should be prevented as far as possible.	Inevitable but also useful. Should be monitored and managed.

Figure 6: A comparison of Safety I and Safety II (Hollnagel E. , 2014)

Safety I define safety as a condition where the number of adverse outcomes (accidents/incidents/near-misses) is as low as possible. The objective of safety management is then to reduce the number of adverse events to an acceptable level. Safety I represent a traditional view of safety as absence of accidents that, when they occur, are caused by identifiable failures or malfunctions of specific components, such as technology, procedures the human workers and the organizations where they are embedded. A key concept in this respect involves work-as-imagined and work-as-done. Work-as-imagined represents the way personnel in the blunt end imagine that work should and carried out. While work-as-done represents the way personnel in the sharp end actually carry out the work. The main point is that the work carried out by personnel in the sharp end requires constant adjustment, which is impossible to fully identify and describe looking from the blunt end. Thus, there will be a difference between work-as-imagined and work-as-done. The functioning of the system is considered as bimodal; it succeeds (normal) or fail (abnormal). When it succeeds (number of adverse events are acceptable), it is because work is performed as imagined, in particular that people follow procedures. When the system fails, this is because something went wrong, like failure or malfunction of a technical or human "component", i.e., a deviation compared to work-as-imagined. The purpose of safety management is to maintain the system in the normal functioning by constraining performance through structural measures (regulations, procedures, standardization, etc.) and by eliminating errors when something has gone wrong. This view of safety coincides with the linear and epidemiological theories and models referred to in the earlier part of this chapter.

In safety II, safety is defined as a condition where as much as possible goes right, and alternatively as the ability to succeed under expected and unexpected conditions alike, so that the number of intended and acceptable outcomes (in other words, everyday activities) is as high as possible. A central aspect of safety II is that the functioning of a socio-technical system is fully dependent on performance variability in everyday work, meaning that people are able to adjust what they do to match the conditions of work. According to this understanding, both success and failures (accidents and incidents) happens in the same way, through the required performance variability as part of everyday work.

Based on this, Safety II focus on understanding everyday work and how successful outcomes are produced, especially as the successful outcomes represent the majority of the total outcomes. This understanding challenge the one-sided view of humans as a source of error as it implies that the variability that humans (and organizations) represent, and most of the time is the source of success, is wrong. When performance variability is understood as a prerequisite for functioning of the system, eliminating unwanted outcomes by constraining performance variability (safety I), would also affect the desired acceptable outcomes. Instead, efforts are needed to support the necessary improvisations and performance adjustments by understanding why and how they are made, by clearly representing the resources and constraints of a situation and by making it easier to anticipate the consequences of actions. It is important to point out that these two perspectives do not exclude each other but needs to be combined as illustrated in Figure 7. The relativeness in the figure, provides guidance on which perspectives that should be focused.

Figure 7: Relationship between Safety–I and Safety–II (Hollnagel E. , 2014)

3.2.6 The ETTO-principle

Another key principle in the understanding of everyday work and performance variability is the efficiency-thoroughness-trade-off (ETTO) principle (Hollnagel E. , 2009). This principal deal with the need of individuals and organizations to be both efficient and through. Efficient because resources always are limited, and in particular because time is limited. It is likewise necessary to be thorough both to make sure that we do things in the right way, so that we can achieve what we intend, and to avoid adverse consequences – incidents and accidents. An efficiency-trade-off can happen due to several reasons, such as limited time, tendency to not use more effort than needed, a need to maintain reserve (time) in case of unexpected events, social and organizational pressure, or individual priorities. The ETTO-principle is further about the dilemma between "time to think and time to do". In most cases, the time available and the time needed to both think and do does not correspond, with a resulting trad-off situation. Thus, people (and organisations) as part of their activities frequently – or always – have to make a trade-off between the resources (time and effort) they spend on preparing an activity and the resources (time and effort) they spend on doing it. The trade-off may favour thoroughness over efficiency if safety and quality are the dominant concerns, and efficiency over thoroughness if throughput and output are the dominant concerns. It follows from the ETTO principle that it is never possible to maximise efficiency and thoroughness at the same time. Nor can an activity expect to succeed, if there is not a minimum of either.

3.2.7 Dynamic safety model

The dynamic safety model (Rasmussen, 1997; Cook & Rasmussen, 2005) is another perspective that relates to perspectives of Resilience Engineering. The model (Figure 8) can be seen as a good illustration of how performance variability is influenced, and supplements the perspectives above, including the ETTO-principle and the understanding of safety as an emergent property (Chapter 4.2.8). The model describes the operating space, or envelope, of socio-technical systems, where its operating point is influenced by gradients that drive operations away from unacceptable workload and economic failure. These forces lead to a possible drift towards the marginal boundary, and possible past the acceptable performance boundary (i.e., accidents). The drift is dampened by a counter gradient formed by safety campaigns, e.g., safety culture **(Rasmussen, 1997)**. The authors point out that the actual location of the operating point is uncertain due the dynamics of the system and the influential forces. For real systems, the operations are frequently located close to the marginal boundary as they seek to optimize workload and cost. They further point out that the crossing of the marginal boundary is recognized by near-miss events, which leads to actions to bring the operations back on the right side of the boundary. In these setting, there is a possibility that the boundary can creep towards the acceptable boundary, if repeated near-misses are judged to be a result of the original boundary being too conservative.

In Figure 9, three examples of systems and their operating pointing are plotted. This illustrates the importance of precise insight to both the location of the margin boundary and the systems operating point, including its extent, represented by the variation in the systems operations. Such insight is attributes of systems included in B), in contrast to systems in C).

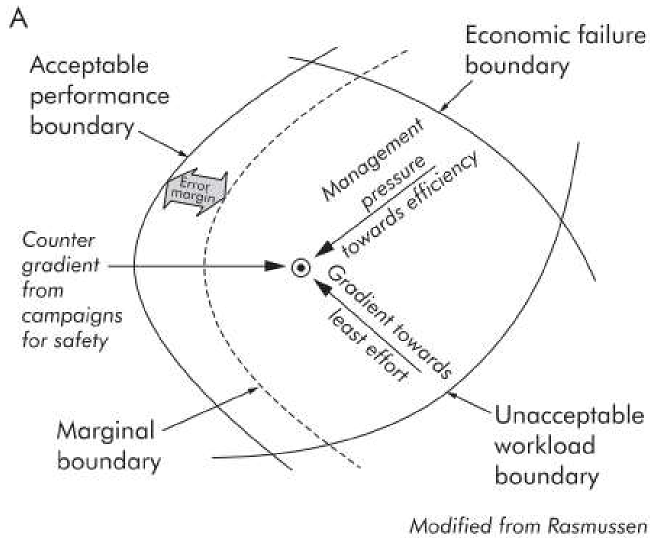


Figure 8: Dynamic safety model (Cook & Rasmussen, 2005)

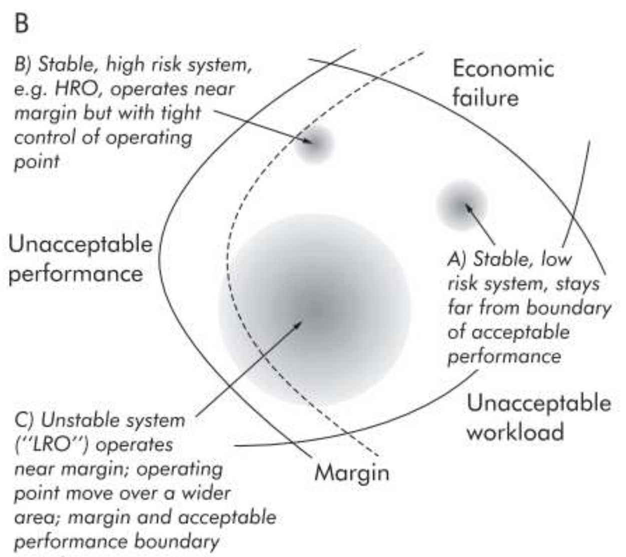


Figure 9: Dynamic safety model with examples of systems operating points (Cook & Rasmussen, 2005)

3.2.8 Dynamics of safety and manageability

The inevitable presence of performance variability and the need of adjustment can also be linked to the coupling-interaction chart by Perrow (1999), visualized in an updated version by Hollnagel (2009), as shown in Figure 10. Here it can be noted that the interaction dimension has been remained to "manageability", and the scale redefined from tractable to intractable. The changes do not alter the concept of the chart, but it highlights the effect of complexity when it comes to the ability to manage/control the systems activities and a corresponding need for performance adjustments. For intractable system, the actual functioning of the system is only partly known and subject to continual changes. The ability to describe and control such system is therefore low and the need for performance adjustments are subsequently high.

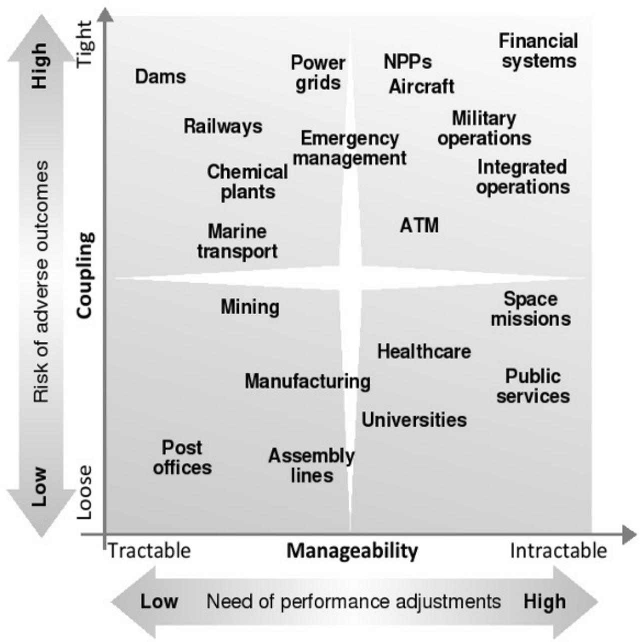


Figure 10: Coupling-manageability chart, based on Perrow's coupling-interaction chart (Hollnagel E. , 2009)

In line with the above, in the perspective of Safety II, safety can be viewed as dynamic non-event, where the non-event represents the absence of unwanted outcomes (incidents and accidents). That it is dynamic tell us that non-event cannot be guaranteed. This is an important aspect as it clearly underlines that safety is not a static property of the system but needs to be constantly managed and monitored. Hollnagel (2014) further point out that the non-event is consistent with *freedom* in another common definition of safety, as "freedom from unacceptable risk". Safety can then be defined as an "ability of a system to perform its intended purpose, whilst preventing harm to persons (..assets and environment)" and where safety, or the lack of safety, is an *emergent* property of the system through the combined result of the decision and actions of all persons with an ability to interact with the system (Provan, Woods, Dekker, & Rae, 2019). Emergent can be defined as something that is "starting to exist" or "to become known" (Cambridge Dictionary). The meaning however is not that emergences happens “magically”, but simply that it happens in such a way that it cannot be explained using principles of decomposition and causality (Hollnagel E. , 2012). These perspectives are important to keep in mind as they provide a valuable insight to how safety management can/should be practiced.

3.3 Safety management

As pointed out in the introduction of the theory chapter, safety management is central term in the main question the thesis. The former chapters provide insight in how safety can be understood and how safety is influenced through the functioning of the socio-technical system. Safety management can thus be considered as the adaptation of this insights and principles in the organization, and can be defined as *practices that can direct, monitor and intervene in core operations for the purpose of generating or maintaining safety* (Provan, Woods, Dekker, & Rae, 2019), or *as the intentional organisation or arrangement of parts (components, people, functions, subsystems) that makes it possible to achieve specified and required goals* (Hollnagel E. , 2009), or *as systematization of organizational conditions that should lead to safe operations* (Størkersen K. V., 2021). For instance, the two perspectives Safety I and Safety II is a good illustration of how the underlying understanding of safety, influence how safety management is practiced (Provan, Woods, Dekker, & Rae, 2019). This includes influence with regard to which activities that is incorporated, the way they are executed and which aspects that are in focus. Based on this, it is not possible to provide a generic description of what constitute safety management practices in way of specific activities and their contents. I.e., it depends on the specific socio-technical system and the prevailing understanding of safety.

There are however activities that are commonly associated with safety management, such as risk assessments, safety audits, inspections, reporting and an analysis of unwanted events, accident investigations and application of safety performance indicators (Kongsvik, et al., 2018). Common for these activities is that they seek to provide information about the current or the anticipated safety level of the organization, to support decisions with regards to necessary preventive measures. Such preventive measures can be grouped in physical/technical, organizational and personal measures. Kongsvik, et al., 2018, further point out that safety management is to a large extent about organizational learning which include changes in the organization’s formal routines, the unformal cultural aspects and the tacit knowledge applied in work- and decision processes.

Performance indicators are an integrated part of a safety management, and its concept are further elaborated below. When referred to safety management, they can be seen as organizational tools for the evaluation and improvement of the functioning of the socio-technical system (Reiman & Pietikäinen, 2011). And further that they represent an important constituent of a safety management system involving the establishment, implementation and follow-up of corporate policies, acceptance criteria and goals related to safety. In this respect, Wreathall (2009) points out that indicators in itself do not provide improvements in safety. It is the quality of the safety management system that is important. It requires a motivation and an understanding for the organization to be seeking opportunities for improvements.

3.3.1 Safety management system (ISM Code)

When talking about safety management in shipping, the ISM Code (IMO, 1993) must be noted as an influential instrument through the past 3 decades. The Code is made applicable through the SOLAS convention (IMO, 1974) and requires shipping companies to develop safety management systems, according to the requirements of the Code. In the ISM Code, a safety management system is defined as a "structured and documented system enabling Company personnel to implement effectively the Company safety and environmental protection policy". The safety and environmental protection policy of the company shall describe how the objectives of the company will be achieved.

These objectives are defined in the Code and is to; (1) provide for safe practices in ship operation and a safe working environment; (2) assess all identified risks to its ships, personnel and the environment and establish appropriate safeguards; and (3) continuously improve safety management skills of personnel ashore and aboard ships, including preparing for emergencies related both to safety and environmental protection. In addition, the safety management system should ensure; (1) compliance with mandatory rules and regulations; and (2) that applicable codes, guidelines and standards recommended by the Organization, Administrations, classification societies and maritime industry organizations are taken into account.

The Code provide further requirements to what the safety management system shall incorporate, including defined responsibilities in general and for key roles (master and designated person), procedures for resources and personnel, shipboard operations emergency preparedness, reporting and analysis of unwanted events, maintenance, documentation and audit and periodical reviews.

The SOLAS convention (IMO, 1974) further requires that the shipping companies and its vessels holds a Document of Compliance and Safety Management Certificate which is based on periodical audits of the implemented safety management systems. The ISM Code include implementation guidelines for both the maritime administrations and the shipping companies.

By adhering to the former chapters, it can quite easily be identified that safety management and safety management system, as required to Code and its sub-chapters, are not necessary the same. The definitions in itself give such indication, i.e., a safety management system serve as means or a tool that supports the organizations management of safety. Here lies a potential for misunderstandings, that is important to be aware of. When it is referred to safety management systems in the remaining parts of the thesis, it should not be thought of as merely the "documented system", but in broader terms, in line with description in the former chapter, including the underlying understanding of safety that guides it.

3.4 Safety performance indicators

Safety performance indicators has been subject for studies more than 50 years (Kjellén, 2009). Resent years, major accidents, such as the BP Texas City refinery and Deepwater Horizon accident, have identified the need for *early waring* indicators that can reduce the risk for major accidents. In both cases, investigation revealed that injury rates were used as measure of process safety performance. Provided that personal safety and process safety (or system safety) require different understanding with regard to causes and measures, this approach was concluded inadequate as means to prevent major accidents. One of the recommendations after the BP Texas City accident was to develop, implement, maintain, and periodically update an integrated set of leading and lagging performance indicators for effectively monitoring of process safety performance (Baker, et al., 2007). Similarly, in the investigation report of the Deepwater Horizon accident it was concluded that the imperative to prevent another offshore catastrophe supports efforts by industry to actively monitor safety performance indicators that capture barrier and safety management system health (CSB, 2016). These accidents and the subsequent investigations illustrate the purpose of safety performance indicators, as well as key aspects of the concept, such as leading versus lagging, and process safety versus personal safety.

When it comes to purpose, Hale (2009) propose a three folded purpose for applying safety performance indicators: (1) to monitor the level of safety in a system (whether that is a department, a site or an industry), (2) to decide where and how to take action if the answer to (1) is that action is needed, and (3) motivating those in a position to take action, to take it. The purpose can further be understood through numeral definitions found in the literature, e.g. Kjellén & Albrechtsen (2017) defines safety performance indicators as "An indicator used to express an organization's ability to control the risk of accidents in its activities", or as "An indicator is a measurable/operational variable that can be used to describe the condition of a broader phenomenon or aspect of reality" (Øien, Utne, & Herrera, 2009). According to OECD (2008), the term "indicators" is used to mean "observable measures that provide insights into a concept – safety – that is difficult to measure directly". Wreathall (2009), defines safety indicators as "proxy measures for items identified as important in the underlying model(s) of safety". A common aspect of the definitions is that safety is something that cannot be measured directly – "proxy measures" - given that safety is seen as a "phenomenon" or "concept". This can be seen to coincide with an understanding of safety as an emergent property of the organization.

In my opinion, the definitions by Kjellén & Albrechtsen (2017) and Wreathall (2009) provide a clear and constructive understanding of what we are aiming to measures – the ability to manage risk - and how this need to be linked to the underlying models of safety, i.e., the organizations understanding of safety and how it is brought about.

3.4.1 Types of safety performance indicators

The definitions referred to in the former chapter can be seen as generic, in the sense that they do not differentiate between the different types of indicators that can be found in the literature. Within the different types of indicators, the majority can be grouped according to a time perspective, commonly referred to as leading- and lagging indicators, proactive- and reactive indicators and activity- and outcome indicators. The leading or proactive type of indicators may also be referred to as "early warning" indicators and the lagging or reactive type as "loss-based" indicators. There are also examples of indicators that aim to capture the present performance, such as current indicators (Herrera, 2012) and monitor indicators (Reiman & Pietikäinen, 2011). In addition to the time perspective, performance indicators are often given a notation that indicate which safety perspective (theory, method or model) the indicator is based on, such as "risk based" (Øien K. , 2001), "assumption-based" (Leveson, 2014) and "resilience-based" (Øien, Massaiu, Tinmannsvik, & Størseth, 2010). Some of these types of indicators are described in more detail under chapter 4.4.3.

3.4.1.1 Proactive-, current- and reactive indicators

As mentioned in the before chapter, there are several notations used to describe indicators related to time perspective. Although the leading-lagging notation is observed to be the most common in the literature, I choose to use the proactive-reactive notation in this heading, as a suitable common term for both the leading-lagging and activity-outcome notation. Irrespective of the notation, the distinction between the two types have been subject to much debate. Several articles have been published as part of this debate (Safety Sience, 2009), staring with an article by A. Hopkins (Hopkins, 2009) where he discusses how the concept of leading- and lagging indicators are presented in the "Baker report" (Baker, et al., 2007) and the HSE guide (HSE , 2006). Hopkins conclude i.a., that the term leading and lagging are not used consistently. The published articles (Safety Sience, 2009) have a common objective to clarify the understanding of the terms by addressing relevant aspects, such as process safety versus personal safety, use of models for development of indicators and the causality/correlation.

The authors seem in general to be aligned in the understanding that lagging indicators for personnel safety, such as LTI, cannot be used as basis for concluding risk of process safety, as the precursors leading to personnel accidents and process accidents often are different. Further it is argued that the understanding of leading/lagging must include understanding of the relationship between the underlying (safety) model and the way indicators are used (Wreathall, 2009).

Hale (2009) point out that the objective with the use of safety performance indicators is to decide where and how to take action if the indicator(s) reveal a non-satisfactory safety level. This requires indicators deeper in the system, which is casually linked, or has high correlation, to the harm/safety outcome. The importance of causality and corelation is widely supported by the researchers. Dyreborg (2009) argue that the casual relationship between leading- and lagging indicators is important, i.e., for experience feedback and organizational learning. Decreasing leading indicator performance levels calls for improvement of existing controls, whereas decreasing lagging indicator performance levels, e.g., incidents or adverse effects without such a leading indicator decrease, calls for a revision of the risk control, i.e., reconsidering the causal relation between leading and lagging indicators. Thus, changes in the performance levels for leading and lagging indicators respectively, have different consequences for organisational safety learning (Dyreborg, 2009). Hale (2009) also point out that it necessary that the leading indicators can be shown to correlate with the lagging ones, as this provides the proof to managers (regulators and workforce) that they are valid.

As a contrast to this, Reiman and Pietikäinen (2011) argue that the outcome (lagging) indicators do not have a direct relation to the to the current safety level of the organization when safety is viewed as the organizational potential for dealing with both expected and unexpected circumstances. However, if the three types of indicators (lead drive, lead monitor and outcome, ref. chapter 3.4.3.3) show inconsistent results, organization's underlying model of safety might be flawed. They further point out that outcome (lagging) indicators cannot be used to measure safety, as safety is not an outcome (event), but a dynamic non-event that cannot be characterized or counted. This represent a paradox, meaning that safety is measured indirectly by its opposite, not by what happens when it is present, but by what happens when it is absent or missing (Hollnagel E. , 2014). This understanding coincides with the Resilience Engineering perspectives of safety as described in the former chapters.

When it comes to definitions, the HSE guide (2006) apply the following definition of the terms:

> "Leading indicators are a form of active monitoring focused on a few critical risk control systems to ensure their continued effectiveness. Leading indicators require a routine systematic check that key actions or activities are undertaken as intended. They can be considered as measures of process or inputs essential to deliver the desired safety outcome", and

> "Lagging indicators are a form of reactive monitoring requiring the reporting and investigation of specific incidents and events to discover weaknesses in that system. These incidents or events do not have to result in major damage or injury or even a loss of containment, providing that they represent a failure of a significant control system which guards against or limits the consequences of a major incident. Lagging indicators show when a desired safety outcome has failed or has not been achieved"

Similar to the leading and lagging indicators, OECD (2008) apply the following definition of the terms:

> "Activities indicators are designed to help identify whether organisations are taking actions believed necessary to lower risks. Activities indicators are pro-active measures and are similar to what are called "leading indicators" in other documents. They often measure safety performance against a tolerance level that shows deviations from safety expectations at a

> specific point in time. When used in this way, activities indicators highlight the need for action when a tolerance level is exceeded", and
>
> "Outcome indicators are designed to help assess whether safety-related actions (policies, programmes, Outcome indicators procedures and practices) are achieving their desired results and whether such actions are leading to less likelihood of an accident occurring and/or less adverse impact on human health, the environment and/or property from an accident. They are reactive, intended to measure the impact of actions that were taken to manage safety and are similar to what are called "lagging indicators" in other documents. Outcome indicators often measure change in safety performance over time, or failure of performance"

Another way to distinguish leading- and lagging indicators is to treat lagging indicators as naive documentation of what has been going on, whereas leading indicators require understanding, knowledge and interpretation in order to point towards the future (Herrera, 2012). Kjellén (2009) propose a similar definition, in the sense of pointing to the future, when he considers a leading safety performance indicator as an indicator that changes before the actual risk level of the organization has changed. Another similar definition can be found in EPRI (1999), that focus on human performance and defines leading indicators as an indicator that provide information about developing or changing conditions and factors that tend to influence future human performance. And further; effective leading indicators provide a basis for predicting or forecasting situations in which the potential exists for a change in human performance, either for the better or worse.

In addition to the distinction between lading and lagging, the notations current (Herrera, 2012; EPRI, 1999) and monitor (Leveson, 2014; Reiman & Pietikäinen, 2011) are used. According to Herrera (2012), the current safety performance indicators are used to indicate how the system actual operates under various constraint. Then, the leading and indicators refer to what may happen or to possible states in the future and the lagging indicators refers to the events that have happened or system states in the past. Compared to the definition of leading indicators in the HSE guide (2006), it can be found that this is quite similar with the current/monitor indicators. This indicate that the distinction may be different depending on the perspective, i.e., if we are looking at individual activities/elements or the system as a whole. In this respect, I agree with Hopkins (2009) when he recommends that the terms should be clarified carefully each time it is use, to establish a common understanding in a specific context rather than generalization across different disciplines. The relevance of the distinction is supported by the value of causality/corelation (as described above), and with regard to support a balanced composition of different types of indicators (Herrera, 2012).

3.4.2 Criteria for indicators

In the literature there are several proposed criteria/characteristics the performance indicators should comply with in order to serve as good measures. Examples of such characteristics are shown in table 3. According to Herrera (2012) there are no single indicator that meets all criteria's, thus it will be necessary to provide a reasonable compromise. Hench, the criteria should be used as guide to development of a suitable and effective set of indicators. In the application of the criteria, I would suggest that the "meaningful", or "Valid or representative of what is to be measured", are prioritized and consulted initially.

Table 3: Performance indicator criteria

IAEA (2000)	Kjellén (2009)	(Herrera, 2012)
There is a direct relationship between the indicator and safety,	Quantifiable and permitting statistical inferential procedures.	Meaningful
The necessary data are available or capable of being generated	Valid or representative of what is to be measured.	Sensitive
Indicators can be expressed in quantitative terms.	Provide minimum variability when measuring the same conditions.	Reliable
Indicators are unambiguous.	Sensitive to change in environmental or behavioural conditions.	Measurable
Their significance is understood.	Cost of obtaining and using measures is consistent with the benefits.	Verifiable
They are not susceptible to manipulation.	Comprehended by those in charge with the responsibility of using them.	Inter-subjective
They are a manageable set.		Operational
They are meaningful.		Affordable
They can be integrated into normal operational activities.		
They can be validated.		
They can be linked to the cause of a malfunction.		
The accuracy of the data at each level can be subjected to quality control and verification		
Local actions able to be taken on the basis of indicators		

3.4.3 Methods and models for designing performance indicators

The former parts of this chapter provide an overview of key characteristics of the performance indicator concept, including definitions, distinction between proactive and reactive, and indicator criteria. The objective of this chapter is to present methods and models that can be used to develop indicators and support a common understanding of how safety is brought about. A large portion of the research on the development and application of these models are related to high-risk industries, such as nuclear power industry, chemical process industry and offshore petroleum industry (Øien, Utne, & Herrera, 2009). In general, these methods and models can be grouped in two main tradition, those that apply the organizational factors (e.g., safety culture/climate) perspective and those that apply the risk analysis perspective (Kongsvik, Almklov, & Fenstad, 2010). As an example of the latter category, it can be referred to the work by (Øien K. , 2001) which describe a model (ORIM) for development of organizational risk indicators based on quantitative risk analysis of a specific technical system. The step-by-step guide (HSE , 2006), further elaborated below, can also be seen as representative of this tradition. The REWI-method and the concept of "lead drive" and "lead monitor" indicators can be seen as representatives of the former tradition. Also, as mentioned in the introduction to the theory chapter, the perspectives of Resilience Engineering are considered as a valuable line of thought in this thesis. Thus, some of the methods and models are selected to incorporate these perspectives.

3.4.3.1 Input-process-output model

Kjellén & Albrechtsen (2017) presents an overview of types of safety performance indicators which is based on a framework for accident analysis (Figure 11). The framework builds on other accidents models, such as the energy model, process models, and system models. This is a good example of how one model can be used to both understand how accidents have happened (reactive) and how accident can be prevented (proactive). When referred to the categorization of accident models in chapter 4.2, this model can be considered as complex-linear model. When referred to the two main traditions above, I would consider this model a representation of both.

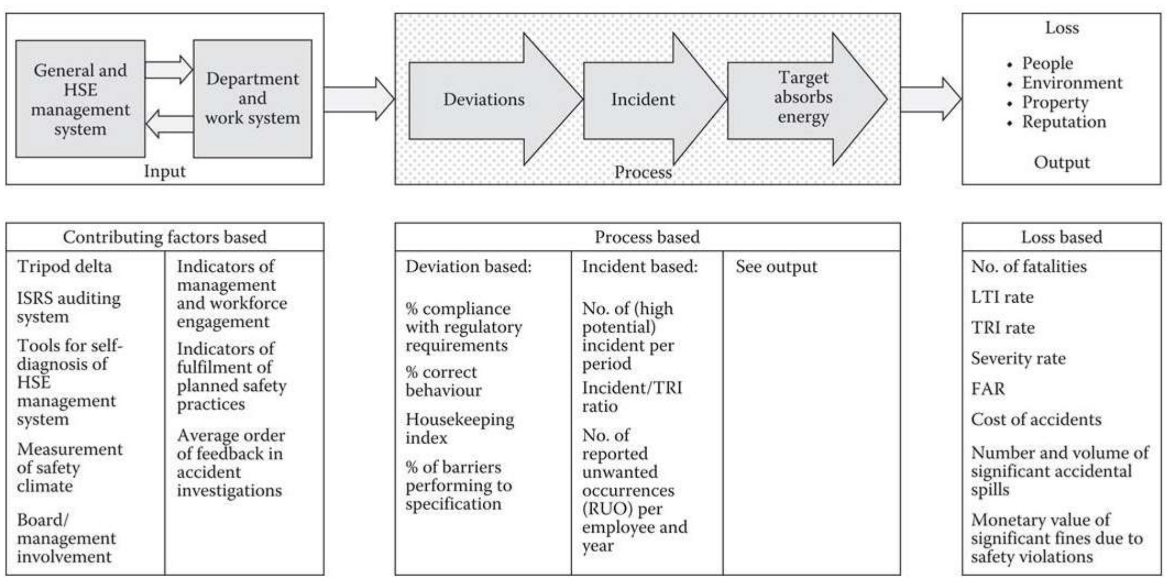


Figure 11: Input-process-output model (Kjellén & Albrechtsen, 2017)

According to this model, accidents can be understood based on an input-process-output model, where contributing factors and root causes in the human, technical and organizational system (input) generate an accident sequence through the systems activities and processes (process) that produce a loss (output). The types of indicators are related to this model and are labelled contributing factor-based, process-based and loss-based. Within the process-based type, two sub types are defined: deviation-based and incident-based.

Kjellén & Albrechtsen (2017) further presents examples of indicators within the different indicator category as shown in Appendix I. As I see it, the strength of this model is that it includes and illustrate important aspects related to both organizational factors and on process/activity level. Thus, the model, combined with example of indicators can be seen as valuable contribution to a performance indicator process.

3.4.3.2 REWI-method

This resilience based early warning indicator (REWI) method (Øien, Massaiu, Tinmannsvik, & Størseth, 2010) method is developed as part of a research project ("Building Safety"), to addressed safety opportunities and challenges in petroleum exploration and production in the northern regions (SINTEF, n.d.). The method is based on Resilience Engineering theory and aim to support the development of certain organizational and management factors, associated with attributes of a resilient organization. According to the method, these attributes are represented by a set of contributing success factors (CSF) at tow levels, as shown Figure 12. Below here and for each CSF a set of general issues are defined, and which form the basis for identification of indicator candidates. The method includes a step-by-step approach that starts with the review and assessments of the general issues, ending with the review and updating of the establish set of indicators. Examples of such indicators are shown in Appendix 2.

The authors acknowledge that indicator-based approaches may be seen as inconsistent with fundamental principles of resilience engineering. For instance, the resilience engineering line of thoughts would deny that we can establish valid relationships between measurements (indicators) and the concept they ultimately are supposed to measure; safety. Here it can be understood that that they refer to an understanding of safety as a dynamic non-event (Hollnagel E. , 2012), and that this represents a paradox i.e., how to measure something that "does not happen". This note is important as it clarifies the objective and scope of the method, which is to establish a common understanding of organizational factors that are important in order to continually produce this non-event (i.e., capacity or potential), and to support related contributing activities through a set of indicators. As the authors point out, the major benefit of application of the REWI-method is the process that such activity generates in the organization, in particular as it:

- provides a common language for talking about safety,
- requires the mobilization of resources for monitoring and evaluating safety-relevant activities over time
- requires a dialogue across departments/units/people on safety issues.

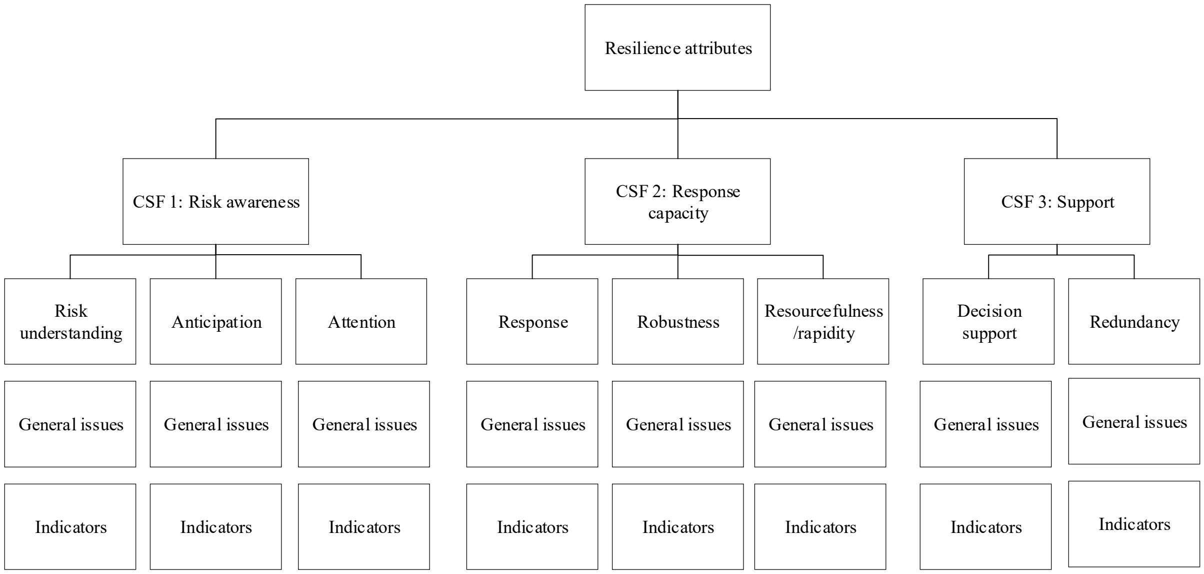


Figure 12: Principal parts of the REWI-method. Reproduced from (Øien, Massaiu, Tinmannsvik, & Størseth, 2010)

3.4.3.3 Lead-drive and lead-monitor indicators

The concept of lead drive and lead monitor indicators (Reiman & Pietikäinen, 2011) is a similar approach to the REWI- method in the sense that it builds on RE perspectives where the aim is to support the development of the organizational potential for safety. This potential for safety is represented by certain important factors, referred to as the "positive bunch" (Table 8) as a counterpart to a set of non-desirable factors, referred to as the "dirty dozen". These important factors ("positive bunch") can be seen as attributes of good safety culture which they describe as the ability and willingness of the organization to understand safety, hazards and means of preventing them, as well as ability and willingness to act safely, prevent hazards from actualising and promote safety. Safety culture refers to a dynamic and adaptive state (Reiman & Pietikäinen , 2010). This safety culture is further considered as basis for the safety model and safety boundaries in Figure 13, which again act as reference in the organization's conception of current safety level. The conception further influences which goals the top management sets for the organization, and thereby guides which criteria's to be used for the drive indicators. The drive indicators are defined as measures of the fulfilment of the selected safety management activities – the chosen priority areas of the organizational safety activity. The drive indicators are turned into control measures that are used to manage the system: to change, maintain, reinforce or reduce something. The main function of the drive indicators are to direct the sociotechnical activity by motivating certain safety-related activities. Examples of such indicators and measures are shown in Appendix 3.

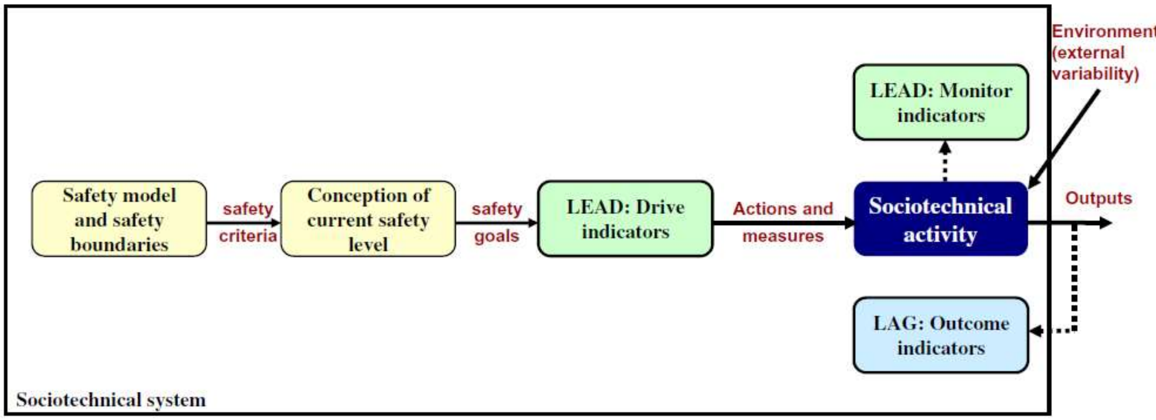


Figure 13: System model illustrating different types of indicators (Reiman & Pietikäinen, 2011)

Monitor indicators reflect the potential and capacity of the organization to perform safely. The indicators monitor the functioning of the system including but not limited to the efficacy of the control and development measures. These indicators seek to measure the internal dynamics of the sociotechnical system and provide information on the activities of the system, including safety culture aspects and changes in the organizations surroundings. Examples of such indicators and measures are shown in Appendix 3. The above illustrate the key distinction of two type of leading/proactive indicators; those that drive safety and those that monitor safety.

The outcome indicators measure the outcome of the socio-technical systems processes and activities, and examples of such indicators and measures are shown in Appendix 3. Here they point out that outcome indicators do not provide insight to the current safety potential of the organizations, but has its main function as a input to evaluate the adequacy of the monitor and drive indicators. In addition, they can serve as basis for improving safety barriers and by making corrective actions to safety system. The concept further illustrates (Figure 14) how the feedback from all the indicators are used to continuously update conception of the safety level, as well as safety model. In the latter case, such need may be warranted if inconsistencies are identified between the different indicators.

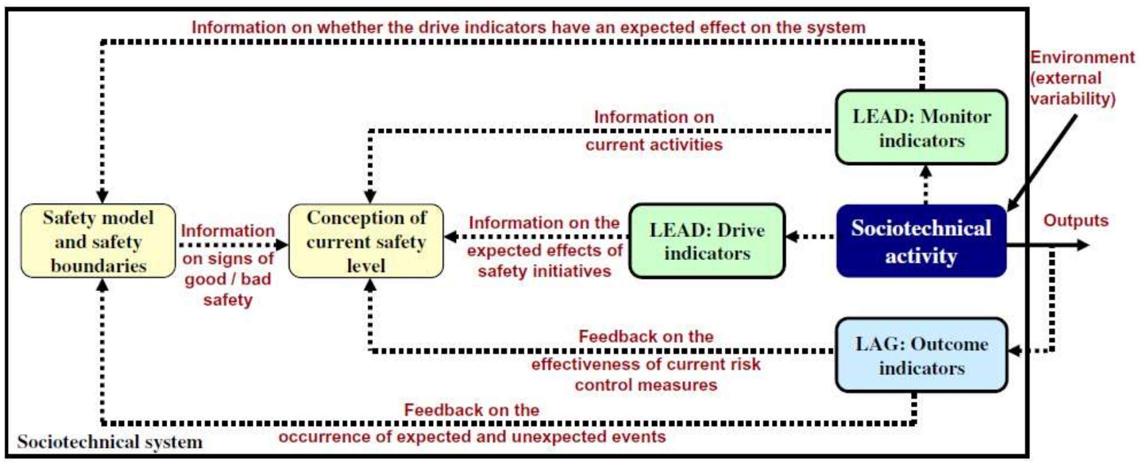


Figure 14: The extended system model showing the feedback from the indicators (Reiman & Pietikäinen, 2011)

Compared to the RWEI-method, the concept described above are similar in the way that thy both emphasize the importance of certain factors that can promote organizations safety potential or capacity. In addition to the REWI-method, the latter concept makes the distinction between lead drive and lead monitor indicators, whereas the monitor indicators are similar to those of the REWI-method, at least when comparing provided examples. Further, by looking at the examples of indicators from both contributions, it can be observed that examples of indicators according to REWI-method are more quantitative oriented than the other. By looking at the examples of both drive and monitor indicators, the majority would require a qualitative approach, such as audits or assessments.

3.4.3.4 Step-by-step guide

The step-by-step guide (HSE , 2006) is a guide for development of safety performance indicators based on key risk control system elements. The risk control system represents those system elements (processes/activities) that is found necessary to establish to control major hazard risk. The guide builds on the Swiss Cheese model (Reason, 1997), where the risk control systems correspond to the cheese slices in Reason's model (Figure 15). With reference to categories of safety theories and models presented in Chapter 4.2, the guide, together with Reason's model, are examples of linear models.

The leading indicators are connected to activities or inputs that is important for the risk control system to perform its purpose (outcome). The lagging indicators are connected to outcome of the risk control system, where unwanted outcomes correspond to holes in the mentioned cheese slices. Although the guide is developed for chemical and major hazards industries, it is considered to be generic and applicable for other industries as well. As the name indicate, the guide provides a step-by-step approach consisting of six consecutive steps, from arrangement of the development process to the periodical review of the set of indicators that is produced through the process.

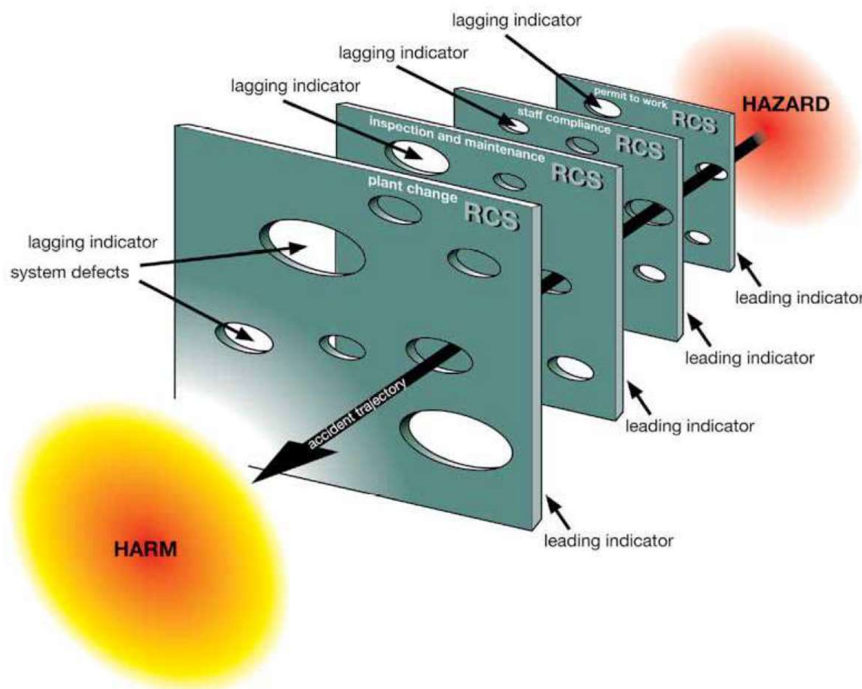


Figure 15: Leading and lagging indicators set to detect defects in important risk control systems (HSE , 2006)

When considering its suitability, a strength lies in the operational orientation that recognise a guide. Combined with the illustrative model based on Reason, the application should be relative manageable, whit regard to both understanding and results. Looking from a systemic perspective, the guide may be criticized for its shortcoming in addressing the performance variability that is normally present, including factors that influence such variability. Such shortcoming may be addressed by the application of supplementary methods, such as the Functional Resonance Analysis Method (FRAM) (Hollnagel E. , 2012).

3.4.3.5 FRAM

The Functional Resonance Analysis Method (FRAM) (Hollnagel E. , 2012) is a method that builds on Resilience Engineering perspectives and are designed on the knowledge that performance of socio-technical systems cannot be adequately understood by use of linear, or complex linear, thinking. Hollnagel argue that virtually all accident investigations and risk assessments are conducted in a state of relative ignorance of the full behaviour of the system. This condition contrasts with the fact that all established approaches to risk assessment require that it is possible to describe the system and the scenarios in detail; that is, that the system is tractable. He further points out that all socio-technical systems are more or less intractable, which means that the established methods are not suitable. Hence, the need for methods that can be used for intractable systems, i.e., the FRAM. Here it is important to add that there are differences between systems when it comes to being tractable (Perrow, 1999), and that individual socio-technical systems can comprise both liner (tractable) and non-linear (intractable) relationships (Hollnagel E. , 2014). The conclusion that established methods are not suitable could therefore be questioned. As I see it, the point here is that liner models have its limitation, but can be suitable if used for tractable system, acknowledging its limitations.

The FRAM can be seen as a concept, consisting of a terminology and a set of methods that make it possible to describe work-as-done rather than work-as-imagined. The primary purpose of the FRAM is to build a model of the functions of a system that describes how performance variability may occur in everyday operations and how the effects may spread through the system. This provides the necessary basis for identifying potential problem areas in the system's functioning, in addition to the more traditional analyses of failure modes, malfunctions and so on. Once such problem areas have been found, a number of well-known remedies can be applied (Elimination, Prevention, Protection and Facilitation)
The FRAM is further based on four main principles:

- The Equivalence of Failures and Successes
- The Approximate Adjustments
- Emergence
- Resonance

Application of the FRAM include four steps:

1. Identify the and describe the functions
2. Identification of variability
3. Aggregation of variability
4. Consequence of the analysis

As part of the description of the functions (first step in the application), six aspects can be used to characterize the individual functions. The function and associated aspects can be illustrated as hexagonal shape (Figure 16).

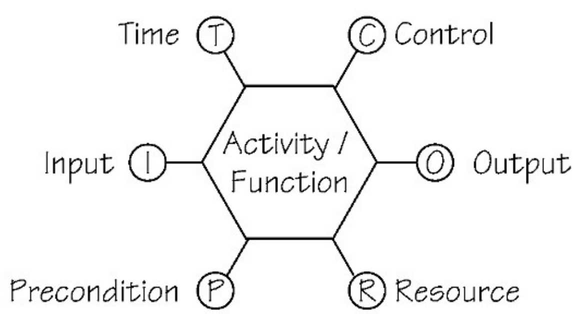


Figure 16: The functional representation used in the FRAM (Hollnagel E. , 2012)

In addition to analysis of the function of the system, the FRAM has also be used as basis for identifying relevant performance indicators, such as within the aviation industry (Herrera, 2012) and, more specific, on the basis of a road tanker offloading scenario (Sujan, Petriarca, Constantino, & Villani, 2021).
Within the maritime domain, FRAM has been applied in re-analysis of several accidents, such as the Herald of Free Enterprise (Praetorius, Lundh, & Lützhöft, 2011), the Sewol ferry capsizing (Kee, 2017), the Exxon Valdez grounding (Smith, Veitch, Khan, & Taylor, 2018) and the Prestige accident (Salihoglu & Bes¸ikçi b, 2011). These studies show that the FRAM can add new perspectives and understanding to the dynamics involved in accidents, compared to existing investigation results. In addition, FRAM have been used to model and analyse an ordinary shipboard operation (mooring at quay) (Patriarca & Bergström, 2017). In the latter study, the actual complexity of a perceived simple operation, is illustrated and provide an impressive graphic representation (se example). As the interpretation and analysis of such complex representation may be challenging, the FRAM has been combined with an abstraction/agency framework allowing for analysis at focused abstraction levels without losing the systemic perspective. Depending on the scope of the analyses, the FRAM can be a time-consuming method as it requires considerable data to be provided (Kee, 2017)

3.4.4 Safety indicators in the shipping industry

The application of performance indicators is not explicitly required through the maritime regulatory framework. Indirectly it can be seen as requirement by referring to the ISM Code (IMO, 1993), Clause 4, on the Designated Person (s) responsibility to monitor the safety and pollution prevention aspects of the operation of each ship. How this should be done specifically is not addressed in the Code.

Within the industry, the Shipping KPI system (BIMCO, 2020), which comprise both a standard (set of indicator candidates) and an online reporting/benchmark functionality, are the dominant instrument when it comes to performance indicators. The standard includes a concept of a 3-level hierarchy of indicators (KPI group, KPIs and PI), and a set of proposed indicators candidates. The KPI group consist of 8 defined areas, such as environment, health and safety, human resources, operational, etc. The majority of the indicators are related to deficiencies, violations, incidents and accidents.

Another influential instrument is the TMSA (Tanker Management Self Assessment) programme (OCIMF, 2017). The objective of the programme is to serve as means to improve safety management systems, encouraging shipping companies (tanker segment related) to assess their system against a set of KPIs, dived in 4 levels, whereas level 1 represents the minimum. In general, the KPIs focus on the presence of certain arrangements, activities and conditions that should be present, including a set of corresponding best practice guidance. For example, under element 2 - Recruitment and Management of Shore-Based Personnel, one KPI is: "Continual professional development of personnel is encouraged and supported". The results of the self-assessment are shared with clients (normally oil companies) and may be subject to audits as part of the vetting regime.

4 CASE STUDY RESULTS

The case studies consist of four Norwegian shipping companies which operate in the passenger segment (Case 1 and 4) and the oil/chemical tanker segment (Case 2 and 3). The size of their fleets are in the range of 15 to 50 vessels. For each case a semi-structured interview has been carryout out according to defined interview guide (Table 1). The cases are numbered based on the order of the performed interviews. In total the interviews involved 10 respondents in management positions, including Managing director, QHSE manager, HR manager, Technical manager and Operations manager. The results are presented in two parts, whereas the first focus on the indicator sets and second on the results of the interviews.

4.1 The indicator sets

In the following analysis I have used the types of indicators described by Kjellén & Albrechtsen (2017); loss-based, process-based (with sub types: incident-based and deviation-based) and contributing factor-based. With regard to proactive/reactive distinction, the loss-based, incident-based and deviation-based indicators are considered as reactive, where the loss-based indicators are considered as the most reactive. The loss-based indicators include incidents that cause harm to personnel (fatalities, injuries and illness), damage to property or environment. The incident-based indicators include reported near-accidents (near-misses) i.e., incidents that did not resulted in losses, but could have under slightly different circumstances. The deviation-based indicators include identified deviations towards requirements, correct behaviour, housekeeping norms and required barrier performance. The contributing factors-based (CF-based) indicators address management and work force engagement and fulfilment of planned safety practices. In addition, the CF-based indicators can be based on various organizational assessment tools, such as safety climate measurements.

When comparing the individual indicators applied by the four companies against the model, it is not always perfectly clear which label to be used. E.g., the percentage of overdue maintenance jobs are according to the model considered as a CF-based indicators, but it can also be considered as deviation (deviation-based indicator) if there are established internal requirements/targets for the amount of overdue maintenance jobs that is accepted. In the analysis I have labelled the indicators that are related to compliance with statutory requirements as deviation based, and indicators that are related to possible deviation to internal requirements/targets as CF-based. Another issue that may be subject to interpretations is the incident-based indicators which is based on reported near-accidents. In addition to measuring the numbers of near accidents it can also provide a measure of the reporting culture of the company. Thus, these types of indicators can be considered as both incident-based and CF-based. In the further analysis, I have labelled the indicators that refer to near accidents/misses and observations as incident-based and those that refer to suggestion for improvements and number of reports (in general) as CF-based, as they are perceived to serve as measure of reporting culture.

For the indicators that are related to commercial (including quality and cost) I have made an attempt to apply the same model and label them accordingly. As the model is safety oriented it is a need to adopt an additional perspective. A loss-based indicator in a commercial perspective is understood as failure to maintain operation and deliver products according to customer requirements and expectations. To define which label to be used for cost/budget related indicators is not that straight forward, especially with regard to the loss-based category if not directly related to other losses. For the cost related indicators applied by the companies I have found it reasonable to label them as deviation based, as they focus on, e.g., budget performance, that is influenced by both normal and unexpected circumstances, compared to a set target. In this case they are considered to be reactive. Alternatively, cost and budget performance could be considered as CF-based indicator looking from a safety perspective, as misbalance between actual needs and available finances can have a negative effect on safety performance.

4.1.1 Case 1

Of the 11 indicators applied by Case 1 (Table 4), five indicators are labelled as loss-based, one incident-based, one deviation-based and four CF-based. With respect to proactive vs. reactive, four of the indicators are considered proactive while seven are considered to be reactive.

Table 4: Performance indicators applied by Case 1

ID	Indicator	Area	Type of indicator	Proactive/reactive
1.1	LTI (per mil work hours)	Personal safety	Loss based	Reactive
1.2	Passenger injuries (per mil transported passengers)	Personal safety	Loss based	Reactive
1.3	Sick leave (sea and shore)	Personal safety	Loss based	Reactive
1.4	Number of spills of harmful substances to environment	Environment	Loss based	Reactive
1.5	Number of overdue improvement reports	Continual improvement	CF based	Proactive
1.6	Number of cancellations	Commercial	Loss based	Reactive
1.7	Number of overdue maintenance jobs (> 8 weeks)	Technical	CF based	Proactive
1.8	Planned vs. executed maintenance jobs	Technical	CF based	Proactive
1.9	Technical incidents leading to off hire	Technical	Incident based	Reactive
1.10	Overdue jobs for charging facility, mooring facility (>8 weeks)	Technical	CF based	Proactive
1.11	Non-conformity "elgrad" against defined targets in tender.	Technical	Deviation based	Reactive

4.1.2 Case 2

For the indicator set of Case 2 (Table 5), I have merged indicators that that are established for both the tanker and dry cargo vessels, and which measure the same aspect in order to avoid duplications. The indicator set is also made up by KPIs and PI, were some of the PIs are used as basis for the KPIs. Thus, I have excluded the PI's that are not considered to measure additional aspects to those covered by the KPI's. Of the 38 (merged) indicators, seven are labelled as loss-based, three as incident-based, 17 as deviation-based and 11 as CF-based. With respect to proactive vs. reactive, 11 of the indicators are considered proactive while 27 are considered to be reactive.

Table 5: Performance indicators applied by Case 2

ID	Indicator (KPI)	Area	Type of indicator	Proactive/reactive
2.1	On-hire (dry cargo)	Availability	Loss based	Reactive
2.2	Number of spills to sea	Environment	Loss based	Reactive
2.3	Average number of near-miss report per vessel per month	Health and safety	Incident based	Reactive
2.4	LTI index – All vessels	Health and safety	Loss based	Reactive
2.5	Number of Total Recordable Cases – All vessels	Health and safety	Loss based	Reactive
2.6	Overdue improvement reports	Health and safety	CF based	Proactive
2.7	Total Recordable Cases Index – All vessels	Health and safety	Loss based	Reactive
2.8	Retention rate – total – all vessels	Human resources	CF based	Proactive
2.9	Budget performance	OPEX	Deviation based	Reactive
2.10	CDI – Average number of findings per inspection	Quality	Deviation based	Reactive
2.11	Number of vessels with overdue critical work orders – all vessels	Quality	CF based	Proactive
2.12	OCIMF-SIRE – Average number of findings per inspection – All vessels	Quality	Deviation based	Reactive
2.13	Overdue technical work orders – Average all vessels	Quality	CF based	Proactive
2.14	PSC – Average number of deficiencies per inspection	Quality	Deviation based	Reactive
	Indicator (PI)			
2.15	Inspections leading to disapproval	Availability	Deviation based	Reactive
2.16	Number of printed paper sheets per day (office)	Environment	Deviation based	Reactive
2.17	Amount of waste (A+B+C+F+I) in dm3 per worked man hours	Environment	Deviation based	Reactive
2.18	Number of assessment hull washing per ship	Environment	CF based	Proactive
2.19	Number of contained cargo spills	Environment	Incident based	Reactive
2.20	Number of contained spills	Environment	Incident based	Reactive
2.21	Number of propeller polishing	Environment	CF based	Reactive
2.22	Fatality – all vessels	Health and safety	Loss based	Reactive
2.23	Crew disciplinary frequency	Human resources	Deviation based	Reactive
2.24	Crew evaluation status	Human resources	CF based	Proactive
2.25	HR deficiencies per SIRE/CDI inspection	Human resources	Deviation based	Reactive
2.26	Internal turnover	Human resources	CF based	Proactive
2.27	Promotion ready candidates	Human resources	CF based	Proactive
2.28	Sick leave (offices)	Human resources	Loss based	Reactive
2.29	Top 2 evaluation status	Human resources	CF based	Proactive
2.30	Violation of rest hours (average per vessel)	Human resources	Deviation based	Reactive
2.31	Cargo related deficiencies – tankers	Quality	Deviation based	Reactive
2.32	Cargo tank rejection (cases)	Quality	Deviation based	Reactive
2.33	Class – Number of CC and CA from planned surveys	Quality	Deviation based	Reactive
2.34	Class – Number of findings from planned surveys	Quality	Deviation based	Reactive
2.35	QHSE & Marine deficiencies per SIRE/CDI inspection	Quality	Deviation based	Reactive
2.36	Rescheduled work orders	Quality	CF based	Proactive
2.37	Technical deficiencies per SIRE/CDI inspection	Quality	Deviation based	Reactive
2.38	Terminal inspections – Average number of observations per inspection	Quality	Deviation based	Reactive

4.1.3 Case 3

For the indicators applied by Case 3 (Table 6), I have merged some of the indicators that measure the same issue but with different perspective, such as total numbers vs. frequency, and indicators that are measure both on vessel and fleet level. Indicators that are used as basis for presenting other indicators and that do not provide measure of a unique issue is also left out in the table below. Of the 33 (merged) indicators, 14 are labelled as loss-based, five as incident-based, seven as deviation-based and seven as CF-based. With respect to proactive vs. reactive, seven of the indicators are considered proactive while 26 are considered to be reactive.

Table 6: Performance indicators applied by Case 3

ID	Indicator (KPI)	Area	Type of indicator	Proactive/reactive
3.1	Fatality	Health and safety management and performance	Loss based	Reactive
3.2	Permanent Total Disabilities (PTD)	Health and safety management and performance	Loss based	Reactive
3.3	Permanent Partial Disability (PPD)	Health and safety management and performance	Loss based	Reactive
3.4	LTI's	Health and safety management and performance	Loss based	Reactive
3.5	Restricted Work Cases (RWC)	Health and safety management and performance	Loss based	Reactive
3.6	Medical Treatment cases (MTC)	Health and safety management and performance	Loss based	Reactive
3.7	First Aid Cases (FAC)	Health and safety management and performance	Loss based	Reactive
3.8	No of LTI's last 12 month	Health and safety management and performance	Loss based	Reactive
3.9	Total Recordable Cases (TRC)	Health and safety management and performance	Loss based	Reactive
3.10	Number of reports recorded pr month	HSE reports	CF based	Proactive
3.11	Expired (reports)	HSE reports	CF based	Proactive
3.12	Fleet No of Near Miss pr month	HSE reports	Incident based	Reactive
3.13	Fleet No of Non-conformities pr month	HSE reports	Deviation based	Reactive
3.14	Fleet No spill pr month	HSE reports	Loss based	Reactive
3.15	Fleet No safety observations pr month	HSE reports	Incident based	Reactive
3.16	Fleet No of suggestion for improvements pr month	HSE reports	CF based	Proactive
3.17	Company Security breaches pr months	HSE reports	Deviation based	Reactive
3.18	Number of Allisions	Operational related incidents	Deviation based	Reactive
3.19	Number of Collision	Operational related incidents	Loss based	Reactive
3.20	Number of Grounding	Operational related incidents	Loss based	Reactive
3.21	Number of Fire incidents	Operational related incidents	Loss based	Reactive
3.22	Number of Mooring incidents	Operational related incidents	Incident based	Reactive
3.23	Number of Explosion incidents	Operational related incidents	Loss based	Reactive

3.24	Incidents related to Cargo and ballast transfer	Operational related incidents	Incident based	Reactive
3.25	Nr of Alcohol violations	Operational related incidents	Deviation based	Reactive
3.26	Observations per inspection ISM	Flag state	Deviation based	Reactive
3.27	Observations per inspection MLC	Flag state	Deviation based	Reactive
3.28	Observations per inspection ISSC	Flag state	Deviation based	Reactive
3.29	Numbers of MOC		CF based	Proactive
3.30	Numbers of best practices identified		CF based	Proactive
3.31	Number of cyber security related incidents	Cyber security	Incident based	Reactive
3.32	Simulated Cyber Attacks (vessels)	Cyber security	CF based	Proactive
3.33	Simulated Cyber Attacks (office)	Cyber security	CF based	Proactive

4.1.4 Case 4

For the indicators applied by Case 4 (Table 7), I have two of the indicators related to overdue non-conformities as they are considered to measure the same issue, provided that the difference is represented by the origin of the non-conformity (all vs. audit non- conformities). Of the 27 (merged) indicators, 11 are labelled as loss-based, one as incident-based, nine as deviation-based and six as CF-based. With respect to proactive vs. reactive, six of the indicators are considered proactive while 21 are considered to be reactive.

Table 7: Performance indicators applied by Case 4

ID	Indicator	Area	Type of indicator	Proactive/reactive
4.1	LTI (per mil work hours)	Personnel	Loss based	Reactive
4.2	Disabling Injury Severity Rate (per mil work hours)	Personnel	Loss based	Reactive
4.3	Number of personal injuries without disability	Personnel	Loss based	Reactive
4.4	Number of passenger injuries	Personnel	Loss based	Reactive
4.5	Number of contractor injuries	Personnel	Loss based	Reactive
4.6	Number of vessels with high risk (NMA)	Compliance	Deviation based	Reactive
4.7	50% reduction of vessels with medium risk (NMA)	Compliance	Deviation based	Reactive
4.8	Sick leave	Personnel	Loss based	Reactive
4.9	Number of contact (berth) incidents	Incidents	Loss based	Reactive
4.10	Number of groundings	Incidents	Loss based	Reactive
4.11	Fuel consumption (per sailed km)	Environment	Deviation based	Reactive
4.12	Portion of suppliers which hold environmental certification	Environment	CF based	Proactive
4.13	Violation of rest hours	Personnel	Deviation based	Reactive
4.14	Violation of qualification requirements	Personnel	Deviation based	Reactive
4.15	Number of speed violations	Incidents	Deviation based	Reactive
4.16	Number of experience messages	Reporting	CF based	Proactive
4.17	Number of breakdowns	Incidents	Loss based	Reactive
4.18	Number of improvement suggestions	Reporting	CF based	Proactive
4.19	Number of customer complaints	Quality	Deviation based	Reactive
4.20	Number of near accidents	Incidents	Incident based	Reactive
4.21	Number of accidents	Incidents	Loss based	Reactive
4.22	Number of deficiencies from external inspections	Compliance	Deviation based	Reactive
4.23	Number of system non-conformities	Compliance	Deviation based	Reactive
4.24	Number of transport damages	Quality	Loss based	Reactive
4.25	Number of overdue non-conformities	Compliance	CF based	Proactive

4.26	Number of overdue critical maintenance work orders	Technical	CF based	Proactive
4.27	Number of overdue maintenance work orders (total)	Technical	CF based	Proactive

4.2 The interviews

In this chapter, the results of interviews are presented according to the thematic analysis described in chapter 3.4. The included quotes are referenced to the case respondents by adding the case number in brackets.

4.2.1 Perceptions of key elements of the indicator concepts

The objective and motivation of applying performance indicators is common for all the respondents, although they express this in different ways:

- To achieve improvements within important areas
- To trend performance, both positive and negative
- To increase the safety level within the organization
- To establish basis for management decisions and motivate desired behaviour to avoid injuries

One respondent refers to expression *"what gets measured, gets improved" (2)* and another use a comparison with navigation and says: "*you need to know where you are before you can proceed the voyage" (4)*

When asked about what recognize good indicators, the respondents suggested several characteristics, such as:

- Clearly defined and easy to understand (no room for interpretation)
- Measurable (not only intentions)
- Easy to measure and possible to compare
- Objective
- Relevant and considered as important (that it has influence on the important areas and not only used because the data are easily available)

In addition, it was said that it is important to maintain predictability and continuity when applying indicators and that "*we must be able to thrust the data, and that requires honest and real reporting" (2)*. Further they state that it is important to have a good link between the indicators and the system elements.

The distinction between proactive and reactive indicators became topic during all the interviews. One respondent state that the distinction between reactive and proactive indicators are not clearly defined within the organization, and further that it has been a discussion related to this, but it is not perceived to have high intention internally. At the same time, they mention that attempts have been made to identify indicators that can predict future unwanted incidents, but they have found it challenging to come up with good candidates Another respondent state that the distinction between proactive and reactive has not been a specific topic in the internal discussions related to the establishment of the performance indicators, and that this may be related to Company culture, where the awareness about this aspect is limited. Another respondent state that the distinction is difficult to define but suggest that a leading/proactive indicator can be understood as *"something that can give an indication before something wrong happens" (3).* Monitoring of near accidents was further suggested as an example of such indicator. Another respondent state that they have put more effort in identifying relevant proactive indicators, e.g., related to competence (familiarization) and risk assessments (task risk assessment). Risk assessments is also mentioned by another respondent in the sense that its application can contribute to more proactive indicators. During one interview it was noted that there was a slight change in the way the respondents considered the set of indicators with regard to proactive and reactive. An initial comment was that they had non proactive indicators. Through the discussion it was later commented that, e.g., retention rate and inspection results can be considered as leading indicators. It was also suggested that maintenance status can be considered as proactive indicator, and that setting of limits/targets in general can improve the proactiveness of the indicators. Maintenance status was also mentioned by another respondent as a more proactive indicator, compared to the set of applied indicators, and similar, the indicator related to speed violations. It was further suggested that e.g., sick leave can be considered as dual indicator that is both reactive, in the sense that it based on history, and proactive, if used to identify actions aimed at improvements.

The set of indicators applied by the companies are all quantitative. During the interviews, other implemented measures was brought up and suggested as possible qualitative indicators, such as safety culture/maturity assessments. One of the companies have carried out such assessment at on instance, while two of the companies state that they perform such assessments at periodical intervals, from 3-yearly (approximately) to 3 times per year as the most frequent. The results of the assessments are used, i.a. as basis for establishing action plans. The company which performs assessments 3 times per year, uses electronic questionaries which is responded to by the vessels crew. The questionaries include a set of topics which in sum are covered through the individual assessments during the year. The results of the assessments are reviewed and discussed by the crew itself, with subsequent establishment of action plans supported by the shore management. The respondents state that the crew's ownership to the assessment and its results, including the following discussions and actions, are perceived to have positive effect on the safety culture. Similar to the safety culture assessments mentioned here, one company have included a set of predefined questions as part of the periodical Masters Review, which is responded to and rated by vessels. The results can be considered as a qualitative indicator which can be monitored on a yearly basis, they suggest. Another example of qualitative measures brought up by one respondent relates to a implemented process for performance appraisals of the vessels crew. This process focusses on how the crew perform compared to wat is expected, related to both behaviour and technical aspects. The results of the appraisals can i.a. be used to identify training needs and as basis for evaluating promotions.

4.2.2 Development and use

The development process described by the respondents is in general much alike across the companies, involving representatives of the management team as part of the management review process. None of the companies have formalized this development process, e.g., by referring to a documented methodology or model. One respondent state that the identification of the current set of indicators used the different disciplines as a starting point, where the representatives of the management team shared their individual thoughts about important aspects/areas to monitor. Following this, the proposed indicator candidates were sorted and prioritized, whereas some indicators were omitted. Another respondent state that the starting point for the establishment and evaluation of the indicators are the company's deliverables and the established policy, related to quality, availability health & safety and environment. A common and key principle in the development process is to evaluate and adjust the current set of indicators based on the experience and performance during the last year/period. Thus, the indicators are to a large degree based on the aspects that the company consider important in order to ensure safety, including experience and areas where challenges are observed. *The process in general, and the result, is to a large degree influenced by the individuals involved and the culture of the Company (1)*

Although there are similarities between the indicator sets, the cases studies do not indicate that the use of predefined generic indicators are common for all. Still two of the company's state that they use/consider using a defined standard (Shipping KPI) as a reference source in the development and application of indicators. *"we use the shipping KPI standard as a reference source and have also some indicators that we have developed ourselves*". LTI is used as an indicator by all the companies and one respondent point out that its widespread use provides a good basis for benchmarking

The development is to some extent influenced by external stakeholders to the companies. There are however differences between the cases where a significant external influence is represented by oil companies expectations, and the TMSA (Tanker Management Self-Assessment) regime. The respondents sate that the clients (oil companies) have expectations with regard to specific indicators they expect to see implemented, e.g., serious accidents frequencies. They further point out that these expectations in practice is a requirement, as failure to implement it will count as a negative and may be commented on in audits and vetting's. Although the elements in the TMSA are all referred to as KPI's, the clients (oil companies) do not expect to find established indicators for all the elements. This use of the term can, according to the respondents, contribute to confusion. The other companies state that their clients do not require specific indicators, but for the two companies (Case 1 and 4), some of their clients requests reporting on specific areas, such as personal injuries and other serious injuries and contact damages. One respondent also states that the increased attention to sustainability, among clients and in general, have introduced new indicator candidates that have been considered, such as diversity (gender, culture, etc.).

All the companies have established targets related to the applied indicators, although not for the complete set. The targets are adjusted based on past performance, normally last year. One respondent state that "*It is important that the targets that are realistic and at the same time something to stretch for (2)"* Another state that not all the set targets are realistic, *"the target for personal injury is set to "0", at the same time the company have experienced relative high numbers of personal injuries historically (1)".* The targets are set without influence by external stakeholders, except for one case (3), which experience that their clients (oil companies) tend to comment on set targets, e.g., related to overdue maintenance orders for critical equipment.

Three of the four companies monitor and review their indicators on a quarterly basis, in addition to the annual review as part of the management review process. One of these have earlier tried to apply a monthly interval but have not found it feasible. It is through these reviews that actions are established, or adjusted, based on the past performance. "*If the indicators are not in line with the set targets, relevant actions are identified an incorporated in a established action plan"(2)* or as another respondent state; *"If the quarterly review show that the results are OK, we do not dig into the numbers" (3),* and further *"our safety performance have been good and we rarely see the need to identify actions based on the review of indicators" (3)*. Thus, they believe that the indicators show that the daily work is effective. When it comes to implemented actions effect on performance, one respondent state that *"It is not easy to identify the effect of individual actions/initiative, it is the sum of actions that contribute to improved performance" (2).* Another respondent state that they over time we have seen that actions and focus areas have given positive effect and at the same time also observed that the positive effects have declined when focus have been shifted to other areas. This relation between effect and focus/actions have been observed i.a. related to personal injuries; *"if we relax the measures, we see an increase in injuries" (4).* Another example is related to sick leave where they have worked persistent with good result; *"we see that focus change behaviour" (4).* They further point out that they have a lot of procedures and routines that shall contribute to good safety performance, but they see that this is not enough in itself. Another respondent describes an experience they had related to change of trade areas for the vessels which resulted in increased LTI, stress, and high workload for the crew. As a result of the analysis of this experience, the company implemented a management of change procedure for the changes in trade areas and which they claim had a positive effect. Yet another respondent refers to an example of identified need to initiate actions where they observed an increase in overdue maintenance tasks. This was analysed and it was found to be a reporting issue and corresponding actions was implemented with positive results.
Several respondents state that identification of actions is not limited to the results of the performance indicators. One respondent state that; *"the management team is "hands on" the daily operations and if we see negative results, we initiate required actions, not only based on the indicators" (3)*

Such negative results are often reported incidents, isolated ones or several similar, exemplified by one respondent which is looking into suitable actions based on three incidents involving fall from ladders on board (with injuries) last year. In this respect it was pointed out that *"It is luck that no lives have been lost" (4)*.
None of the respondents uses the indicators as part of incentive measures on a individual level. However, two of the companies practice a "ship of the year" award where i.a. some of the performance indicators are taken into account. The award includes additional funding to the awarded vessels which shall be used on welfare initiatives for the crew. The experience with this practise is positive and it is said to generate attention within the fleet.

4.2.3 Perception of value

In general, the companies perceive that the current set of indicators are suitable and provides a reasonable picture of the safety performance and the effectiveness of the system. One respondent state that; *"If they had been removed, they had been missed, and probably more accidents would happen" (2).* Another respondent point put that focus on the people is the result of the good performance; not the indicators as such. It was also commented that larger companies have more value of performance indicators than the smaller ones.

The effort invested in collection and analyses data is also perceived to be reasonable compared to the gained value, although one of the companies state that their data handling process is not fully developed and another respondent state that it can be a challenge to collect and summarize data from the different systems, such as the incident reporting system and the maintenance system. Here it is mentioned that the data collection process has become more effective with use of data warehouses.
It is perceived that the indicator sets and the way there are applied are continually improved as part of the periodical review, by adjustments and removal of indicators which provide limited value. Still the respondents state that there are improvement potentials. Common for all are the issue related to proactive vs reactive indicators where they all acknowledge that they have an overweight of reactive indicators, although the distinction between the two types were not clearly defined (ref. the theme above). One respondent state that the ambition is to look into how the structure according to Shipping KPI can be used with regard to levels of indicators. The aim is that department leaders can focus on the lower levels of indicators and the top management can focus on the upper level or overall picture. One respondent state that they see potential of increasing the effectiveness of processes, such as the periodical Management Review, if they manage to set and apply the indicators in good way. Several respondents state that it is important to set of adequate time in order to address the improvement potentials.

In the application of performance indicators there are potential pitfalls that in many ways can represent a failure to establish indicators with good characteristics as mentioned earlier. Here the respondent's states that it can be tempting to establish too many indicators with the risk of unbalance between effort and gained value. This temptation may be influenced by easily availability of data, which does not necessarily mean that they are the ones to be prioritized. One respondent point out that if unrealistic targets are set (not possible to achieve), this can lead to reduced credibility of the system within the organization. Another comment that the use of incentives can potentially lead to false reporting. In this context, one respondent mention that they have observed examples of "dishonest reporting" in the industry, which may be due to customer requirements and the possibility for losing jobs due to "poor" performance.

Further one respondent state that if wrong indicators are applied, they can give a false impression of the system and lead to wrong focus and potential sub-optimalization. Related to this, one respondent refers to an example where they changed system some time ago so that the vessels cannot changed due dates (postpone) of reported non-conformities. This resulted in an increased number of overdue NC's (as one of the applied indicators) from 12 to 80 to per vessels.

4.3 Summary of the results

The above analysis were the individual indicators applied by 4 companies have been labelled based on the model by Kjellén and Albrechtsen (2018), show that the majority (approximately 75%) of the applied indicators are considered to be of the reactive type. These reactive indicators include the three types; loss-based, incident-based and deviation based. Within these three types of indicators, the loss-based are considered as the most reactive and the deviation-based as the least reactive, and the incident-based in between. Within the loss-based category, which is the type of indictor that is most applied, 23 of the 37 indicators are related to either personal injuries, fatalities or illness. The indicator that is common for all is the LTI. The incident-based type of indicators represents the minority of the applied indicators. Within this category, number of near-miss/accidents are applied as an indicator by 3 of the 4 companies. The second largest category is the deviation-based indicators. Here, the majority of the indicators (20 of 34) are related to deficiencies or non-conformities from external inspections and audits. Among the CF-based indicators a large portion (12 of the 27) of the indicators are related to maintenance an improvement report backlog. Other indicators related to crew (retention/turnover/evaluations) and number of "positive" reporting (improvement/experience exchange) make up a portion of 10 of the 27 indictors. The remaining 5 indicators are; Number of assessment hull washing per ship (2.18), Number of propeller polishing (2.21), Numbers of Management of Change (MOC) (3.29), Simulated Cyber Attacks (3.32) and Portion of suppliers which hold environmental certification (4.12). Only one of these are considered to be directly safety oriented, i.e. Numbers of Management of Change (MOC). The overall picture is that the applied indictors are mainly reactive and focus on losses and lack of compliance identified by external parties, and backlogs in the maintenance system and improvement report handling process. Figure 17 and Figure 18 illustrate how the indicators are distributed within the different categories.

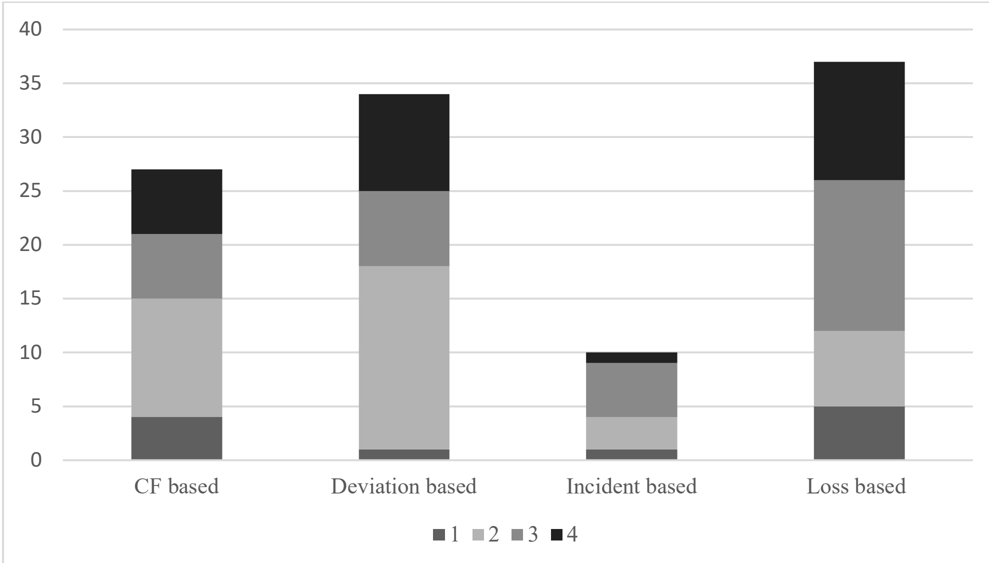
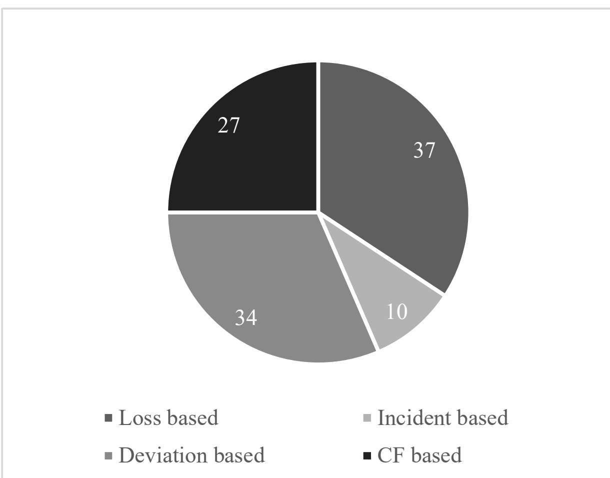


Figure 17: Distribution of applied indicators based on types, and reference to case number



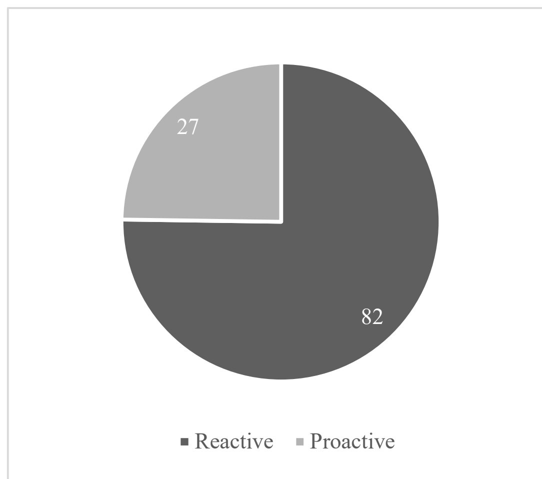


Figure 18: Distribution of applied indicators based on types and reactive vs proactive

According to the performed interviews, the involved companies apply performance indicators as a means to improvements, a key aspect of the performance indicator concept. The interviews provided insight to the respondents perceptions of other key elements as well, such as what recognise “good” indicators, distinction between proactive and reactive indicators, and possible qualitative indicators/measures, as supplement to the quantitative indicators currently applied. In general, the companies have clear perceptions of what constitutes “good” indicators. The distinction between proactive and reactive are in general perceived to be somewhat unclear, although definitions and examples of indicators are suggested. With regard to supplementary qualitative indicators/measures, the majority of the respondents presented examples of such, particularly related to safety culture.

Det process for development and use of the indicators are to large extent similar across the cases. They are monitored periodically (monthly/quarterly) and are updated/revised as part of the periodical management review process, based on the past years performance and perceived value/relevance. The development process mainly involves management representatives. The indicators, combined with set targets, are also used as basis for identification of improvement actions. At the same time, it is emphasized that it is not only the set of applied indicators that provide basis for initiation of actions. In general, the development process runs without substantial influence from external stakeholders, although some clients have specific expiations to which indicators that are established. With regard to perception of value, the dominant feedback is that the indicators are suitable and provide a good picture of safety performance. The value compared to efforts put into the application is also perceived to be reasonable. At the same time, all respondents see potential for improvements, such as addressing the overweight of reactive indicators.

5 DISCUSSION

In this chapter the case study's is discussed in relation to theory and the main question and the research questions of the thesis.

5.1 Application of indicators

Research question No. 1: How do shipping companies apply performance indicators as part of their safety managements systems?

As the research question imply, the main objective is to establish an understanding into both type of indicators and how the companies apply them as part their safety management systems. In this respect, the results of the case studies, as presented in the former chapter, provides the basis for such understanding. In the following, the results will be discussed in more detail with objective to provide enhanced insight to central aspects. This will include looking into the different type of indicators and how contextual circumstances, related to shipping in general and specific to induvial companies, may have an influence on application. Theory will be referred to when found relevant.

5.1.1 Context influence

The cases studies reveals that there are many similarities in the way the case study companies apply performance indicators, but there are also differences. The overall objective is, not surprisingly, in principle common across the cases and is aligned with the common understanding found in the theory. When comparing the set of indicators applied by the four companies, the number of indicators varies from 11 to 38. One explanation to these differences may be connected to the prevailing context for the individual companies. Two of the companies (Case 1 and 4) do not manage vessels in international trade and are not subject to specific industry related inspection regimes as the two other companies (Case 2 and 3). Indicators related to for instance port state control, ISPS, MLC, OCIMF, and CDI inspections are thus only relevant for

Case 2 and 3. It can also be noted that while Case 1 “only” have one indicator related to personnel injury (LTI), Case 3 have 8, including LTI. This mean that they both apply indicators related to personal injuries, but the difference is that one has measures on a broader aspect. Such comparison show that the set of indicators is not that different with regard to what they measure, as the numbers might indicate.

5.1.2 Loss-based indicators

According to the case studies, the loss-based indicators (together with the incident-based) represent the largest category among the indicators applied by the four companies. Provided that safety is commonly understood as “the absence of unwanted outcomes such as incidents or accidents”, it is no surprise that all the case study companies have established indicators within this category. A similar focus can be found among clients of the companies, as the indicators that they require/expect to see implemented, are of the loss-based type. Incidents and accidents are further hard to ignore, and their occurrence will subsequently be recorded in the internal reporting systems and communicated externally to authorities (Forskrift om melde- og rapporteringsplikt til sjøs, 2008). Thus, the data will in most cases be easily available.

Within the loss-based category, a large portion of the indicators are related to occupational health and safety (OH&S) such as LTI, RWC, MTC and sick leave. In this respect, it is important to distinguish between indicators that address OH&S risk and those that address major accident risks (process safety), such as collision, grounding, explosions or fires. The distinction is important because the risks and hazards represented by the two categories are different (Hopkins, 2009; CSB, 2016; Baker, et al., 2007).

In the context of OH&S, the focus is on hazards that are directly related to the individual workers and related incidents such as slips, trips and falls. In contrast, major accident risk arises from the overall functioning of the socio-technical system where the picture of causes and contributing factors are considerably more complex. For example, many of OH&S risks can be managed by separating energy sources from the personnel and by the use of personnel protective equipment (PPE), e.g., in line with Haddon’s (1980) “energy model”. While PPE can prevent personal injuries, it has very limited effect on prevention of major accidents, e.g., capsizing. The key message here is that OH&S related performance indicators cannot be used as a exclusive measure of the organization (socio-technical systems) ability to manage major accident risk. This was also one of the key findings from the Deepwater Horizon accident in 2010 (CSB, 2016).

The above does not conclude that OH&S based indicators are of non-value. All shipping companies are required to include OH&S objectives in their management system policy (IMO, 1993) and treated as an integrated element of the management system. Several shipping companies (whereas one of the case study companies) have also management systems that are developed and certified according to ISO 45001 (previous OHSAS 18001). I would also argue that the distinction between OH&S risk and major accident risk does not mean that there are no resemblance or interrelation in the cause and contributing factors, as it may be interpret based on the above. Increased sick leave can for example be caused by high workload and fatigue. The same issues may have negative impact on major accident risk, and thus actions aimed at reducing fatigue would have positive impact on both OH&S and major accident risks. High level of sick leave will also represent increased strain on the organization's resources with potential risk of inadequate production-safety trade-offs (Hollnagel E. , 2009).

The LTI indicator is a commonly used indicator (Kjellén & Albrechtsen, 2017) and also applied by all the case study companies. In addition to limited ability to tell how well major accident risk are managed, the indicator has also been criticised for its shortcoming as measurement of OH&S. This is due to the fact that such incidents are rare and only a few occurs during a year in any particular workplace. Thus, the variation from year to year will be statistically insignificant and will consequently give no guidance on changes in the OH&S levels (Hopkins, 1994). As a means to address the statistically issue with the LTI, alternative indicators, such as the Total Recordable Cases (TRC), are applied in the industry and also by the case study companies. Another critique on the LTI, is the possibility for manipulation (e.g., by assigning an injured person to a lighter job to avoid lost days) (Kjellén & Albrechtsen, 2017). In this context, one respondent mention that they have observed examples of "dishonest reporting" in the industry which may be induced by customer requirements and the possibility for losing jobs due to "poor" performance. Although it was not further elaborated, it illustrates the potential for manipulation, especially in settings where performance is linked to commercial incentives.

5.1.3 Compliance and deviation

The deviation-based indicators represent the second largest category of the applied indicators and are to a large degree based on result from external inspections and audits. Based on own experience, such indicators are commonly used by shipping companies in general. I would claim that this is connected to the fact that the shipping industry is a highly regulated industry (Bloor, et al., 2013), which naturally generate high focus on compliance. A merchant vessel must comply with numerous statutory requirements, related to technical, operational and personnel aspects in order to be licenced to operate. The requirements are enforced by maritime administrations and class societies throughout the vessel's operational lifetime by certification and periodical inspections and audits. Lack of compliance and failure to maintain valid certification leads to operational limitations and subsequent negative consequences for the shipping company in question. Hench, there is both commercial and safety aspects that motivate application of such indicators.

The attention on compliance performance is also influenced by the port state control regimes, such as Paris MOU. The mission of the port state control regime is to eliminate operation of sub-standard ships through a harmonized system of port state control (Paris MoU, 2021). Vessels that are found to be in non-compliance may be detained (prohibited for operation) if the non-compliance is found to be substantial. The results of the port state controls are further used to establish vessel risk profiles, including rating of the vessels flag states and recognized organizations. The risk profile is used as part of the selection scheme with regard to frequency and thoroughness of inspections. The port state control results can thus be viewed as common industry performance indicator that show how the different actors perform (as public available information) with regard to compliance. The two case study companies (2 and 3), that has vessels which is subject for port state control, apply performance indicators based on port state control results.

The deviation-based indicators provide an insight into which degree the vessels are in compliance with requirements. Provided that the requirements, if complied with, represent a defined safety standard, they may be considered as a valuable indicator. Another motivation for applying these types of indicators may be related to fact that these data are made available to the company's "free of charge", i.e., produce by actors outside the company itself. The data are also made relative easily available through the company' reporting systems. As Kjellén & Albrechtsen (2017) point out, individual indicators tend to have both advantages and weaknesses. When it comes to the indicators in question, the data they are based on is a result of assessments perform by different external actors, i.e., a process which the company itself has limited ability to control and verify. Considering that the process involves human involvement (capacity, competence, trade-offs, etc), and that inspections and audits often follow a sampling approach, the data have a degree of uncertainty attached to it. For instance, if an inspection is completed without any findings or deficiencies, it cannot be concluded that the vessel is fully compliant with the applicable requirements. In this respect it is relevant to adhere to the criteria of indicators mentioned in chapter 4.4, such as "accuracy and ability perform control and verification". If further referred to the criteria "valid or representative of what is to be measured" it can be understood that "what is to be measured" is the organizations' ability to operate in compliance with rules and regulations. This ability is the result of the socio-technical system and its constituent parts as a whole. If referred to the criteria that "local actions able to be taken on the basis of indicators" and taken into account that the indicators are generally formulated (e.g., PSC – Average number of deficiencies per inspection), relevant, local actions as response to negative performance would require some investigation. This is normally handled through the incident reporting system where individual deviations are handled and analysed individually. A key point in this discussion is that the deviation-based, as well as the loss-based, indicators can be seen as result or an outcome of activities and processes of the socio-technical system. These types of indicators are categorized as reactive (ref. Ch. 5.1) in contrast to the CF-based indicators (proactive). This actualize the aspect of causality or correlation, which is discussed further in following part.

5.1.4 CF-based and correlation

Within the CF-based category, indicators related to maintenance system are commonly applied amongst the shipping companies. Maintenance is a central part in the management of vessels and is essential with regard to both production and safety. Hench, the motivation to apply maintenance related indicators is easy to justify. According to Hale (2009), it is important that proactive indicators can be shown to correlate with the reactive ones. Provided that an overall objective of the implemented maintenance system is to ensure that vessels are maintained in compliance with rules and regulations (IMO, 1993), a correlation between maintenance-based and deviation-based indicators should ideally be present, i.e., if the maintenance is carried out as planned (no back log), the results of inspections should not reveal lack of compliance. Such correlation is however not that straight forward as it implies several assumptions and variables, e.g., the scope and interval of the maintenance routines are adequately defined, maintenance routines are satisfactorily executed, and equipment and systems remain operable (in compliance) between maintenance periods. The maintenance indicators applied by the cases study companies are in addition generic, i.e., they do not differentiate maintenance routines based on specific equipment and systems. Higher correlation could be obtained, for instance if an indicator measured the maintenance status of specific equipment/system, for instance lifesaving appliances, combined with an indicator that measured number of deficiencies from inspections related to the same equipment.

Retention rate is another example of CF-based indicators that are applied by the case study companies, and commonly in the industry alike. The indicator can be seen as a measure of the organizations ability to retain personnel and the competence that they possess. This is considered to be critical factor for maintaining a competitive advantage and ensuring safe operations. The indicator is further actualized through shortage of skilled seafarers that has been experienced worldwide recent years (Yuen, Loh, Zhou, & Wong, 2018)

Several of the case study companies apply indicators that measure number of reports, such as near-miss, improvement suggestions and total number of reports registered, including number of experience messages. The basic objective with the reporting system is for learning purposes and to identify improvement needs and relevant actions (Kongsvik, et al., 2018). Such indicators can also be seen as measure of the organizations ability to develop a reporting culture, or just culture, as encouraged through the ISM code (IMO, 1993). I agree with (Hasanspahić, Frančić, Vujičić, & Maglić) regarding the importance of well-functioning reporting system, but not their subsequent categorical argument that the causes are the same for near-misses and accidents. This is founded on the systemic perspectives which claim that accidents cannot solely be explained by casual chain-of-events models, but rather as an emergent property in the functioning of the sociotechnical system. The point here is that the nature of the event and the prevailing context need to be taken into account. That may necessitate the application of different perspectives, where many of the simpler events can be sufficiently understood based on a straightforward chain-of-event approach. Indicators that focus on reporting culture (e.g., Reason (1997)) can also be justified in this respect, as many of the other indicators rely on timely and accurate reporting, and thus promote the reliability of the indicator set as a whole.

As presented through the case study result, the applied performance indicators (formally referred as such) are all quantitative. At the same time, other measures, such as safety maturity/culture assessments, was brought up as possible qualitative indicators. When referring to the applied model in chapter 5.1, these types of measures fall under the category of CF-based indicators. This indicate that the overweight of reactive indicators may be more nuanced if relevant qualitative measures are taken into account. The importance of safety culture as means to avoid accidents is widely supported in the literature (e.g., Reason,1997; Pidgeon, 1991). At the same time, the use of various methods in assessment of safety culture/climate in organizations have been found to have both strengths and weaknesses (Filho & Waterson, 2018). For instance, Antonsen (2009) found significant differences in the results of a quantitative safety culture assessment (survey based) and a qualitative investigation after a major incident on an oil and gas platform. Here it is i.a. pointed out that a survey-based approach by use of questioners has a weakness in the way that respondents may answer based on what they think is the "correct" response, rather than what they actually feel. In contrast, a qualitative approach would allow for richer data collection by follow-up questions and extension of scope, as found feasible. Similarly, Kongsvik, Almklov, & Fenstad (2010) presents studies that have found relationships between safety culture and safety performance, and studies that have not. The struggle to establish relationship towards performance is experienced especially when studies are linked to rare events (major accidents). I agree with their conclusion that, based on the existence of identified relationships, safety culture/climate measures could be useful for preventing accidents, as the organisational attributes that a safety culture represent, must be considered as positive.

5.1.5 Perception of value

According to the case study results, the general perception among the respondents is that the application of performance indicators provide value. At the same time, it was pointed out by one respondent that the application of performance indicators might have greater value for larger companies. By this it is understood that company in question (Case 3) and its management team are "hands on" the daily operation, and thus have less need to consult performance metrics in order to decide the current safety level. This raises an interesting question about the potential value of applying indicators related to size of companies. It can quite easily be assumed that the larger the company (number of hierarchical layers, personnel, locations, etc), the more difficult it is for the management team to be "hands on" the daily operations. Subsequently, the value can be assumed to have greater value for larger companies. In this context, it is relevant to refer to the Shipping KPI standard (BIMCO, 2020), which incorporate a hierarchical concept for selecting indicators, i.e., taking into account that the need for performance measurements varies through the hierarchical layers and organizational departments. This concept was also something that case study company no 3 considered to implement.

5.2 Inputs to development of indicators

Research question No. 2: How can safety science theory be used to develop indicators?

In the context of this research question, it is important to keep in mind that the main objective of safety theory is to establish an understanding of what safety is and how it brought about. As the objective of applying safety indicators is to measure and improve safety performance, the value and necessity of safety theory and the understanding it can provide, must be acknowledged. The following discussion aim to evaluate how the relevant theory and its perspectives can be used to develop safety indicators, taking into account the understanding about current application as represented by the case studies.

5.2.1 Key perspectives

I would claim that the process for seeking understanding about safety and how it is brought about should begin with addressing the applicable context, which in this thesis is the shipping industry, and more specific shipping companies. A shipping company may be referred to as a socio-technical system, i.e., an organization which is a combination of social and technical parts that must work together in order to accomplish the organizations tasks (Appelbaum, 1997). As for all commercial organizations, their purpose is to deliver products and services that meets the markets needs and expectations, where cost and effectiveness are important drivers.

At the same time, the activities and processes, including technology, that are established to generate the organizations deliverables, represent potential harm to personnel, environment and assets, i.e., potential accidents. The need to be both effective (cost and time) and avoid accidents represent a potential challenging balance, as illustrated by the ETTO-principle (Hollnagel E. , 2009) and the Dynamic Safety Model (Cook & Rasmussen, 2005). In my point of view, these are fundamental principles that should be incorporated in the understanding of safety and how it is brought about, irrespective of which industry or socio-technical system in question. Further related to context, I consider it important to include the perspective of considering socio-technical systems as open systems (Kongsvik, et al., 2018), meaning that all socio-technical systems are influenced by its surroundings, represented by external factors, stakeholders and interested parties. This is important as the surroundings (or context) represents both threats and opportunities that the organization need to address in order to manage its risks. This include also influence on the organization's ability obtain a sound balance between efficiency and accident prevention. The organizations surroundings will vary from organization to organization with respect to its elements and overall complexity, which is well illustrated by the diversity that represent the shipping industry in general, and through the case studies in this thesis. Although they are all shipping companies, they all have their unique context (segments, clients, area of operation, market situation, suppliers, etc.). Still, there are elements in the general shipping context that have general influence, such as the regulatory framework.

As discussed under research question 1, the widespread use of compliance related indicators may be explained by the regulatory framework, which generate high attention to compliance. Consequently, I would argue that the shipping context have an influence on how safety is understood in the industry, i.e., through compliance with rules and regulations. This is also one of the main objectives in the ISM Code (IMO, 1993), which in my opinion have overshadowed the other objectives, such as managing risk and provide safe practices. This is partly, again my opinion, due to the Code itself, as it is highly general in its formulations and provide limited guidance, in particular with regard to how management of risk should be carried out. Although the intention with the code is to support proactive safety management, the implemented systems have not met that intention, in the way that reliance on extensive procedures, incident reporting and root cause analysis have been dominant. This does not mean that these system elements are important. Procedures for instance, are necessary in many of the high-risk activities in a shipping company, and valuable if adequately designed and effectively implemented. But this is not always the case, as illustrated by Bhattacharya (2011), Vandeskok (2015) and Batalden & Sydnes (2014). The reliance on procedures and compliance as means to prevent accidents corresponds to the attributes of Safety I (Hollnagel E. , 2014) and actualize the concept of work-as-imagined versus work-as-done. In this context, it can also be relevant to adhere to the concept of autonomy versus rules and different perspectives of rule developments, such as the “model 1” versus “model 2” described by Antonsen, Heldal, & Kvalheim (2017).

5.2.2 Context developments and complexity

As presented in theory chapter, safety theory has been/are in constant development. Taking into account that the ISM Code was developed some 30 years ago with only minor amendments, there have been considerable developments both in the context of shipping and in the understanding of safety in the elapsed time. This is unfortunate, and I support Hollnagel’s (2014) concern about the possibility that “…we may inadvertently create the challenges of the future by trying to solve the problems of the present with the models, theories and methods of the past”. When looking at the developments in other management systems standards, such as the ISO-standards, including the ISO 31000 on risk management, they have been subject for more frequent and substantial developments, where the latest versions have adopted i.a. the importance of context and the need to address the risks and opportunities that it represents.

Another relevant perspective in the discussion about context and complexity can be found in the work by Perrow (1999) and further developed by others such as Hollnagel (2009). According to Perrow’s initial work, Marin transport is placed in the upper left quadrant (Figure 4), close to the middle of the interactions dimension (linear to complex). This tell us that a shipping company can be characterised as a socio-technical system with fairly tight couplings and with a need for control that is correct, precise and with limited room for variability (in order to avoid accidents). This need corresponds well with the regulatory framework and the high attention to compliance as discussed above. On the interaction dimension, a shipping company can be characterised as fairly linear, i.e., that actual functioning of the system is fairly predictable and to a large extent can be understood. Referred to the updated version developed by Hollnagel (2009) (Figure 10), a shipping company can be characterised as system which is fairly manageable with limited need for performance adjustments and where the risk of accidents is fairly high. It is also to be noted that marine transport is placed further down and further to left in the updated version.

Without insight in the basis for this change, I would expect that an updated version would indicate change in the opposite direction, or at least a status quo, especially if the recent developments in the shipping companies' surroundings are taken into account. When said, that is not the most important aspect of this picture. What is most important, in my opinion, is the illustration of how the nature of a socio-technical system and its context (both internal and external) creates a valuable basis for deciding adequate safety management principles. Further to this, it is important to keep in mind that no shipping companies are alike, and I would predict that each individual company would be given a unique location in the chart, probably distributed within all four quadrants. For instance, a company within the ferry segment (fairly stable and known operational situations) would be categorized differently than a company within the search and rescue segment (highly variable and unknown operational situations), at least within the manageability/interaction dimension. Thus, also varying needs for performance adjustments. I would also suggest that this understanding can be extended to det different parts of the socio-technical systems, as they may be categorized differently, where the technical parts (vessels with equipment and systems) as more tractable/linear than other organizational parts. When going back to the findings from Bhattacharya (2011), Vandeskok (2015) and Batalden & Sydnes (2014), they might be explained by an understanding of the system as more tractable/linear than it actually is/was. The understanding can further be applied to avoid unrealistic confidence in documented procedures as means to prevent accidents. In addition, the above discussion emphasizes the importance of understanding; the constant need to balance between efficiency and thoroughness and how this crates dynamic forces that can shift performance towards unacceptable limits (accidents), including how this is affected by the context of the organization. It also indicates that the general context of the shipping industry has considerable influence on how safety is understood amongst the industry actors.

5.2.3 The input-process-output model

As presented in the theory chapter there are several proposed methods and models that can be applied to develop indicators, including the model by Kjellén & Albrechtsen (2017) which have been applied to categorize the indicators from the case studies. The model is based on different accident models and can be considered as a complex-liner model. I my point of view, the model has its strengths in its visualization, as for associated models such as the "Swiss Cheese" model (Reason, 1997). Another strength is that it categorizes the different types of indicators according to the elements of the model (input-process-output), and thus makes it relative clear which type of aspect that the indicator aims to measure. Whether the indicators categories are considered proactive or reactive is not explicitly defined in the model, although it may be intuitive that the "output-indicators" are reactive, and the "input-indicators" are proactive. If the importance of clearly distinction between proactive and reactive are considered to be subordinate, such model and approach can help overcome potential confusion and disagreements which may hamper the development process. Particularly if the model in question corresponds with the organizations own understanding of how safety is brought about (Wreathall, 2009), which I would expect is the case for a large portion of shipping companies. This is based on the above discussion and by referring to models commonly used in accident analysis/root cause analysis, such as the M-SCAT (DNV), which has great resemblance to the model by Kjellén & Albrechtsen (2017).

5.2.4 Safety culture approaches

In addition to the model itself, the authors (Kjellén & Albrechtsen, 2017) have presented examples of indicators within the different categories. Among these, are indicators based on safety culture/climate assessments. As revealed through the cases studies, such assessments are applied amongst the case study companies, although not formally included in the set of performance indicators. As mentioned under discussion related to research question No.1, I support the value of measures to promote safety culture, provided that its limitations (Antonsen, 2009) are taken into account when applied. It is noted that advocates of the Resilience Engineering (RE) perspectives have questioned the importance of safety culture, below here just culture and learning culture, e.g., due to its reactive orientation when it relies on reporting and learning from near-misses/accidents (e.g. Hollnagel (2012)). Related to this, RE perspectives have been used as a supplement to the traditional understanding of safety culture (e.g., Reason (1997)), referred to as resilience safety culture (Akselson, Ek, Floor, Stewart, & Marie, 2009; Shirali , Shekari , & Angali , 2018). The resilience safety culture includes the same aspects of the "traditional" safety culture and adds a resilience-based understanding to it. For instance, learning is focused on the everyday work and why things normally succeed, in addition to learning from near-misses and accidents. A further development of this concept, including suitable assessment methods, should be considered as valuable contribution, both with regard to understanding of safety and as a basis for safety performance.

5.2.5 Resilience based methods

The REWI (Resilience Based Early Waring Indicator) method (Øien, Massaiu, Tinmannsvik, & Størseth, 2010) is developed within the Norwegian offshore petroleum industry. Still the authors state that the method can be applied within other high-risk areas as well. In addition to the method itself, their work also includes a set of candidate indicators (ref. Appendix 2). According to the method, the process leading to a set of indicators, start at the three predefined Critical Success Factors (CSF) and their sub-levels which is derived from the attributes of resilience. In my opinion, the method appears to be good alternative, especially with regard to involvement and possibility to build a common understanding in the organization through workshops and evaluation of the general issues. The importance of common understanding is also supported by others as well, such as Kjellén & Albrechtsen (2017). With reference to case study results, the development process within the case study companies mainly involves the management team. Related to this process it was stated that i.a. it is difficult to identify good proactive indicators, that it is limited attention towards seeking proactive indicators, and that the development to large degree is influenced by what the individuals perceive as important. A method such as the REWI could therefore be valuable supplement as it supports common understanding through the predefined CSF that can guide the process towards a set of suitable indicators. The REWI method can also be seen in connection with the resilience safety culture, as a supplementary model, as it also builds on the aspects of traditional safety culture by adding RE perspectives.

Another relevant contribution can be found in the work by Reiman & Pietikäinen (2011). As for the REWI-method, the consept of leading-drive and leading-monitor indicators are inspired by Resilence Engineering and builds on safety cultur aspects. When compareing the concept of drive and monitor indicators with the CF-based indicators according model by Kjellén & Albrechtsen (2017), the drive indicators are quite similar the CF- Based (work place factors) in the sense that they both aim to motivate certain safety-related activies. Similarly, the monitor indicators and the CF-based (general and HSE MS) both focus on the internal dynamics of the soscio-technical system in order to reflect the potential and capacity to produce safety, e.g. through safety culture/climate assessments. Thus it can be understood that the latter category aim to measure the effect of the activities which is motivated through the former category of indicators. One could off course ague that there are incosistencies in the comparison and inference of similarities of these to models, when looking at the specific examples of indicators and the category notations, e.g. CF-based (work place facors) could be seen as more natural to be compared with monitor indicators, rather than the drive indicator. This aside, I would argue that the comparison is valid, at least on a principal level. It is also important, as it illustrate how new/different perspectives can be adopted to existing theories and potentially increase their value and aplicabillity.

In my point of view, one of the aspects of the work by Reiman & Pietikäinen (2011), is their illustation of how esential the underlying safety model and the conception of safety is for the development process. Another important point, is that the proctive safety management (measurements and activities) is about managing the orgaincations safety potential or capacity. Examples of factors that represent such potential/capacity is shown in table 8. As a supplement to this, I find it valuable to refere to the work by Provan, Woods, Dekker, & Rae (2019) which apply an operational view on how and which activities safety profesionals should spend energy on in order to build such capacity. Here they also illustrate how this “new” approach (corresponding to Safety II) differ from the current dominat approaches (corresponding to Safety I).

Table 8: Examples of safety promoting factors

Safety promoting factors (Reiman & Pietikäinen, 2011)	Critical success fcators (Øien, Massaiu, Tinmannsvik, & Størseth, 2010)	Organizational capacities for a safety mode of guided adaptability (Provan, Woods, Dekker, & Rae, 2019)
Vigilance and energy	Risk understanding – how do we achieve knowledge and experience about risk/hazards?	Anticipation - Create foresight about future operating conditions, revise models of risk
Motivation and mental resources	Anticipation – what can we expect?	Readiness to respond - Maintain deployable reserve resources available to keep pace with demand
Assertive attitude to safety issues	Attention – what should we look for?	Synchronization Coordinate information flows and actions across the networked system
Adequate task and safety knowledge	Response (including improvisation) – what must we do?	Proactive learning Search for brittleness, gaps in understanding, trade-offs, re-prioritisations
Situation awareness	Robustness (of response) – how can we ensure completion of the response (without suffering damage)?	
Social permission to carry out the work thoroughly	Resourcefulness/rapidity – how can we ensure timely and sufficient response?	
Norms that support safety	Decision support – how do we support the trade-off between safety and production?	
Clear and accurate communication	Redundancy (for support) – how do we compensate for degradation to uphold/maintain critical functions?	
Flexible organization and slack resources		
Good task and work design		
Functioning teamwork and cooperation		

5.2.6 Criteria and potential pitfalls

The theories and models discussed so far suggest indicators related to organisational factors, such as those representing safety culture. They further focus on promoting activities that are aimed at supporting such factors. Taking into account the proposed criteria of indicators (Chapter 3.4.2), such as “quantifiable and permitting statistical inferential procedures” and “provide minimum variability when measuring the same conditions”, it could be concluded that many of the proposed indicators, as they are formulated, struggle to comply. One reason for this, is that many of the proposed indicators are of the qualitative type and naturally fail the “quantification criteria”. This should however not automatically exclude the indicator as a relevant candidate. To find an indicator that fully comply with all the criteria is also unrealistic (Herrera, 2012). Here it is important keep in mind that the criteria serve as guide to assess the quality and suitability of the indicators. The assessment should avoid indicators that generate inappropriate attention to less important aspects related to safety at the expense of more important ones. In addition, they should prevent the organization spending inappropriate amount of efforts compared to gained value and available resources in the organization. The "meaningful" criteria should therefore be of high priority, which i.a. corresponds to the safety model and conception of safety as the reference source in the development process (Reiman & Pietikäinen, 2011)

The criteria can be seen as guide that can be used to avoid potential pitfalls when applying indicators. In this respect it can be valuable to be aware of the potential biases, such as the What-You-Look-For-Is-What-You-Find (Hollnagel E. , 2012), meaning that our perceptions about what is important guides the way we make interpretations and conclusions. As our perception about what is important is founded in our underlying understanding of how things work/should work (e.g., safety model), the result may be that we make false conclusions, i.e., find something that is not really there. For example, that we interpret and conclude that the socio-technical system, or parts of it, are more linear (in contrast to non-linear) than it actually is. This again, underline the importance of having an adequate understanding of how safety is brought about, as this guides the selection of indicators and the way we interpret them. In addition, it is important to keep in in mind that it is unrealistic to apply performance indicators on all relevant aspects of safety. Subsequently, focusing on some particular issues may lead to overseeing others and can be seen as a "tunnel effect" (Herrera, 2012).

5.2.7 Causality and correlation

In relation to criteria, it has also been argued that it is important that proactive and reactive indicators can be shown to have a causal relationship or high correlation (Hale, 2009; Dyreborg, 2009). In this I agree, provided the that parts of the system they are linked to are sufficiently linear in order to validly conclude that causality exists. If it is understood that the functioning of a socio-technical system comprises both linear cause-and-effect relationships and non-linear emergent relationships (Hollnagel E. , 2014) it must be accepted that indicators that for instance are aimed at promoting factors/capacities as exemplified in Table 8, hardly can be shown to have a causal relationship with indicators that are based on accidents. Correlation can however be found if we refer to the drive and monitor indictors (Reiman & Pietikäinen, 2011), but the degree of correlation that is possible to obtain in practice is questionable. Especially if quantitative safety culture/climate assessments are uses as basis for the monitor indicators. In my point of view, causality/corelation should be sought when relevant and found feasible, based on the unique context of the system in question, but should not constrain the use of non/less casual/correlating indicators, if they are found valuable in line with the prevailing safety model and perception of safety. This topic, concerning balance, is further discussed under chapter 6.3.

5.2.8 Functional resonance

The theories and models inspired by Resilience Engineering that is discusses above is to a large degree about strengthening the organizations safety potential/capacity (Table 8). In order to build such capacity and to utilize it, require an understanding of the dynamics in the socio-technical system, i.e., how things actually work. In this respect, the Functional Resonance Analysis Method (FRAM) (Hollnagel E. , 2012) can be considered as a relevant tool. As presented in the theory chapter, the application of the method have been tested and evaluated in several settings, such as method for accident investigation (Kee, 2017; Smith, Veitch, Khan, & Taylor, 2018), to develop understanding of shipboard operations (Patriarca & Bergström, 2017) and identification of performance indicators (Herrera, 2012; Sujan, Petriarca, Constantino, & Villani, 2021). As illustrated in the referred to articles, the application of the method can be time consuming, and the visual representation and interpretation can be quite complex. This in itself provides a valuable insight, as operations normally considered as simple, such as mooring at quay, can be shown to be far more complex when modelled according to the FRAM. Subsequently, this should be seen as a motivation to seek understanding of how things actually work/should work. In my point of view, the FRAM has a potential application as a tool for developing performance indicators. The obvious question is how can it be adopted in practice, and can the effort be justified compared to the value? As mentioned above, I think that the approach should be to consider the method as a supplementary contribution to the current implemented methods and practices (the organization and its management system in general) and evaluate how it can be adopted to increase value. In the evaluation it should be taken into account that the method does not, as it is described in the literature, represent a procedure that needs be strictly "complied with" in all its steps and details. As illustrated by Patriarca, Bergström, Di Gravio, & Costantino (2018), the method can be modified/supplemented in order to tackle complexity issues without losing the systemic perspective. When looking at the method which is centred around an activity with a set of aspects, a parallel can be drawn to the process approach that is incorporated in the ISO 9001:20015 standard, illustrated in Figure 19. This approach imply that the organization shall decide and describe its required processes, including their, inputs and outputs, mutual relationships, the need for control measures and recourses, and measures of performance. The FRAM representation (Figure 16) can thus, as I see it, quite easily be added as supplement to how organizations adopt the process approach, including how they identify performance indicators. Here it must be noted that the ISO 9001 is a voluntary instrument and not formally applied in the majority of the shipping industry (none of the case study companies have formally obtained certification towards this standard). A relevant question is how much the method can be "simplified" before the intention/value (systemic perspective) is lost. In my point of view, the method can have a value just by the insight it can provide to the involved personnel, and the organization as a whole. For instance, I would expect that the perception of an operational procedure, with regard to how it is developed and its effect as a safety barrier (incorporeal) (Hollnagel E. , 2004), would be quite different when looked at in the light of the FRAM, compared to a “traditional” view, e.g. represented by Safety I.

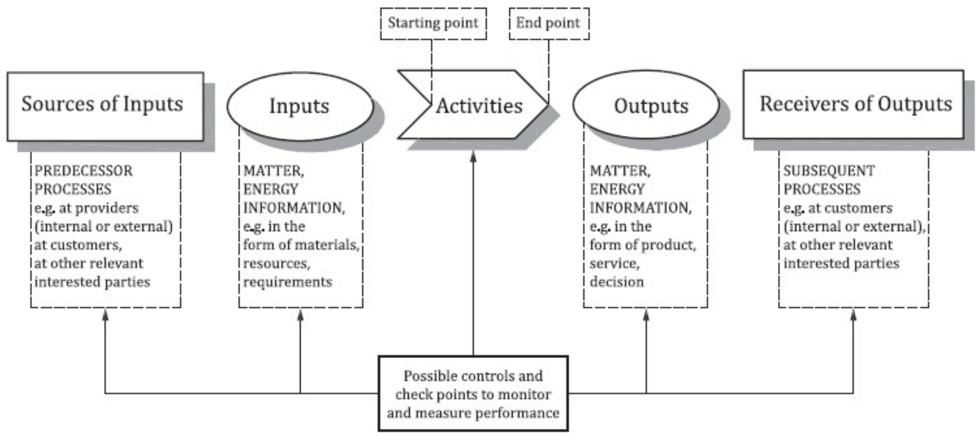


Figure 19: Illustration of the process approach (ISO 9001:20015)

5.2.9 Evaluation of suitability

As touched upon in former part of this discussion, the understanding of the unique context of the organization, including how it can be seen with regard to complexity is considered to be of great importance. This understanding also imply that an organization (socio-technical system) consists of both linear and non-linear relationships. With reference to work by Hollnagel & Speziali (2008), which include an evaluation of different methods and models with regard to suitability compared to the coupling/interaction relationships (ref. Ch. 4.2), the methods mentioned in the above discussion can be recognized. The model by Kjellén & Albrechtsen (2017), which is based on other models, such as Swiss Cheese model, can be considered suitable for system in the upper-left quadrant (Figure 5) corresponding to same quadrant where marine transportation is plotted (Figure 4). Further, the FRAM can be found in the upper-right quadrant. The REWI method and the model by Reiman & Pietikäinen (2011) are not directly linked to methods and models evaluated by Hollnagel & Speziali (2008) (as they are not considered as accident investigation models).

Based on the above, one could conclude that, in the case of a shipping company, the understanding of safety and how to measure performance, can be based on liner/complex-linear theories and methods. This is however partly correct, provided that it is acknowledged that a socio-technical system, as a shipping company is, consists of both linear and non-linear relationships. Subsequently, I would conclude that all the perspectives that these contributions represent in general are suitable. As Hollnagel & Speziali (2008) also state, there are no ideal model/method that fits all needs, and a combination would be necessary. With this I agree, and I would consider that the model by Kjellén & Albrechtsen (2017) can be used as basis and be supplemented by the other methods and models, as illustrated in Figure 20.

The REWI-method and the concept of lead drive and lead monitor indicators can be a valuable supplement to the input part of the model with regard to promoting the organizations "safety potential" or "capacities for a safety mode of guided adaptability", including the monitoring of such "safety potential", or resilience safety culture as representation of such potential/capacity. In this respect I find it important to point out the potential weaknesses in assessments of culture/climate as illustrated by Antonsen (2009) and that this is taken into account when such assessments are applied.

The step-by-step approach presented in the HSE guide (2006) can be used as supplement to the "process" part of the model, focusing on activities and process that make up the organizations risk control system. The process approach included in the ISO:9001.2015 standard could also easily be adopted here. I would also claim that the FRAM, at least as a supplementary perspective, provide value in this respect, in particular with regard to understanding of performance variability and how this influence interrelated activities and processes.

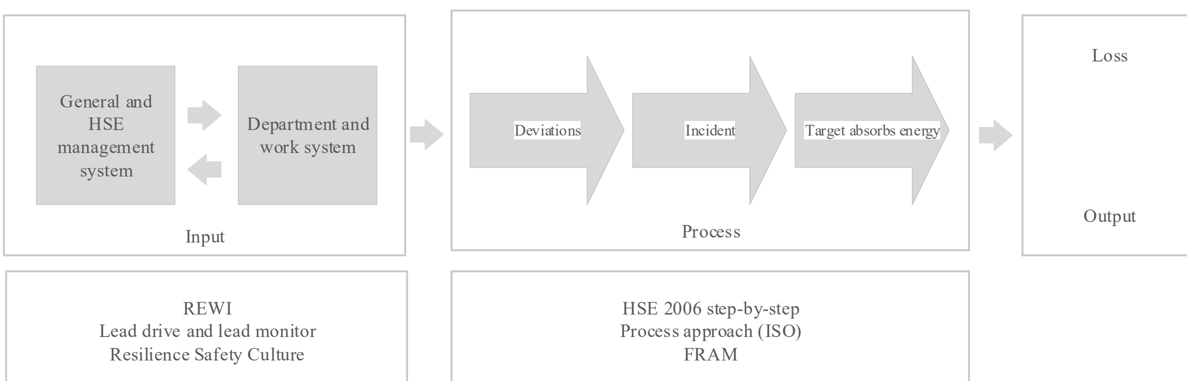


Figure 20: Illustration of how different models/methods can be combined. Reproduced and modified after (Kjellén & Albrechtsen, 2017)

5.3 Balancing the indicator set

Research question No. 3: What recognize "good" balance between different types of indicators?

The research question is based on an assumption that the different types of indicators that are applied, including the balance between those, represents the focus of the organization. Subsequently, which areas are prioritized with regard efforts and recourses, including their orientation, proactive or reactive. According to the case study results, most of the applied indicators are considered to be of the reactive type. Here it is important to point out that the indicators themselves do not necessarily represent the actual orientation of the organization. It was for example stated by one respondent (Case 3), that good performance is the result of focus on the personnel, and not the indicators as such. Nevertheless, it is considered valuable to address the question, as the composition of the indicator set can influence the orientation of the organization. Here it is also acknowledged that the indicators that are applied require certain efforts that potentially could be allocated to more valuable areas. This should be a valid aspect, as resources in most organizations are considered as scarce.

5.3.1 Proactive versus reactive

The question of good balance is not isolated to aspect of where the organization have its focus and prioritize it resources. It can also be related to the robustness or validity of the indicator system itself. As pointed out by several researchers (Dyreborg, 2009; Hale, 2009; Herrera, 2012), it is important that there is causality or high correlation between the different indicators, i.e., between the proactive/leading and the reactive/lagging ones. Based on this, the presence of causality/correlation can be seen as an attribute of "good" balance. Such balance is well illustrated in the HSE guide (2006) where the indicators are identified based on primary outcomes (lagging) of certain risk control system (RCS) elements and important (and correlating) actions or activities (leading) that are necessary in order for the risk control element to deliver its desired outcome. By establishing such balance or correlation, the monitoring and review process can identify if the correct indicators are applied. For instance, if the leading indicator indicate poor performance and the corresponding lagging indicator indicate good performance, this can imply that the leading indicator is connected to an activity that is too distant compared to the risk control element or connected to an activity that has less influence on the outcome of the risk control system. Another scenario can be that the leading indicator show good performance and that the lagging indicator show poor performance. This can indicate that the activity is not performed the correct way, or in other words that the RCS elements are ineffective. Thus, the aim here is to have indicators that correlate, i.e., when the lagging indicator indicate good performance, the leading indicators should indicate the same, and vice versa. If not, the validity of the indicator set can be questioned.

If we refer to model by Kjellén & Albrechtsen (2017), the balance that is represented by measurements of RCS elements outcomes and inputs/activities can seen to correspond to the process element. Although, the types of indicators that are related to this element (deviation- and incident-based) are considered as reactive in this thesis (ref. Ch. 5), the approach in the HSE guide (2006) show that it makes sense to use the proactive/reactive notations, if it is within a defined context or scope. In most cases, a certain activity/process within the socio-technical system, is the subject for measurements. Provided that all activities/processes in a socio-technical system have mutual relationship, where output/outcome from one activity/process are input to another, the inputs and outputs (indicators) can be considered as both proactive and reactive depending on the activity/process in focus.

As an example, the HSE guide (2006) consider "inspection and maintenance" as an element in the RCS, and suggests a lagging indicator based on "number of loss due to component failure" and a corresponding leading indicator based on, i.a. “percentage of maintenance actions completed to specific time scale”. Here the correlation can be considered as valid, provided the assumption that maintenance activities ensure appropriate reliability of equipment and prevent losses, hench considered leading relative to the outcome (losses). At the same time, the “percentage of maintenance actions completed to specific time scale” can also be considered as performance measure of other activities such as “maintenance planning”. In this case, the “percentage of maintenance actions completed to specific time scale” is considered as lagging indicator. A corresponding leading indicator could then be related to availability of maintenance resources (manpower, competence, materials, etc.). This could also be continued by looking at the activities and interrelations related to ensuring adequate resources. Inevitable, activities related to overall management, including influences in the external surroundings would become a subject for consideration. This illustrate that balance and correlation between proactive and reactive measures can be found within isolated elements (local) and at the overall perspective of the socio-technical system (global). Figure 21, which is inspired by the model of Kjellén & Albrechtsen (2017) attempts to illustrate this relationship. If we consider a scenario where the indicator set consist of only loss-based and CF-based indicators, that may be considered as balance, at least with regard to proactive versus reactive. However, when it comes to correlation it can be a problem as losses (accidents) are rare events (Øien K. , 2001).

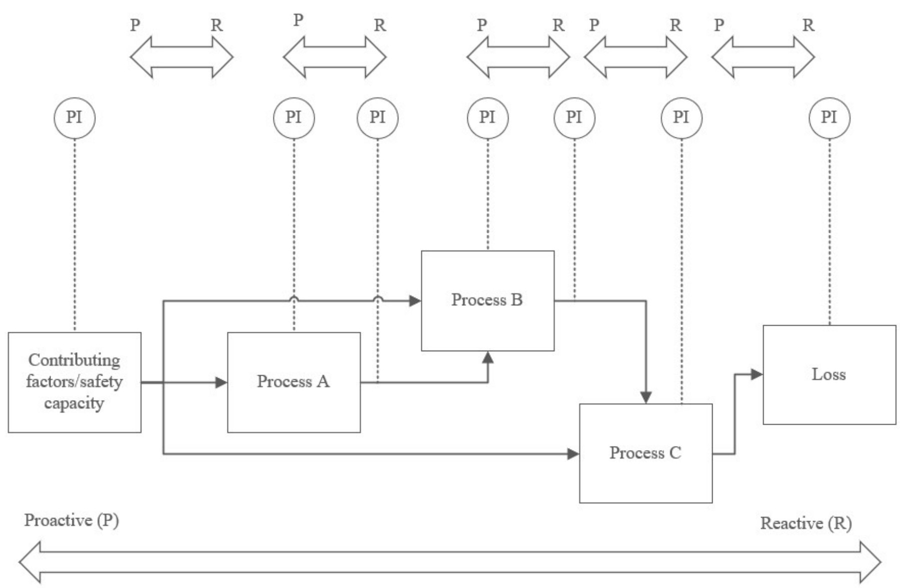


Figure 21: Balance and correlation between different types of indicators. Inspired by Kjellén & Albrechtsen (2017)

A good balance, in my point of view, is then recognized by a well-considered distribution of indicators across the entire socio-technical system. The distinction between proactive and reactive will vary depending on the orientation, i.e., global or local. By well-considered, I refer to discussion under research question No. 2, with respect to ensure an adequate understanding of how safety is brought about, and the ability to apply suitable methods/and models, taking into account that socio-technical system comprise both linear and non-linear relationships. If it is further acknowledged that the main objective with application of performance indicators is to support proactive safety management, i.e., to be able to anticipate what can happen, to take appropriate action and build capacity to handle variability, the types of corresponding indicators should be prioritized in the consideration of "good" balance. Although the prioritization should be maintained (on a principal level), I would also argue that a "good" balance does not represent a static picture, as both the organization and its surrounding is subject to continual changes, e.g., as illustrated by Herrera (2012) (Figure 22)

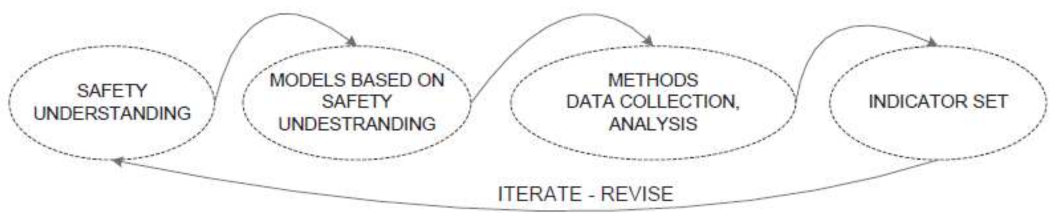


Figure 22: Safety understanding and performance indicators (Herrera, 2012)

5.3.2 Safety versus economy

As indicated in the beginning of this discussion, balance is not limited to the consideration of whether the indicators are considered proactive or reactive on superficial level, it is also about how the organization focus its attention and efforts. In this respect, it is found relevant to include the "Q4 Framework" developed by Woods, Branlat, Herrera , & Woltjer (2015) as a means to evaluate balance of the applied set of indicators. The framework builds on the acknowledged potential that organizations can experience safety-economy conflicts resulting in trade-offs and discounting of safety information and reduced ability to act proactively. This is done by plotting the applied indicators in a chart (Figure 23) consisting of two axis, proactive/reactive and safety/economy, which make four quadrants: reactive-economy, proactive-economy, reactive-safety and proactive-safety. The framework can also be used to analyse how the organization spend their "safety energy" where safety energy is seen as the finite amount of resources (time and manpower). In this analysis, it should be taken into account that the individual indicators do not necessarily consume an identical amount of resources, i.e., certain indicators may be considered more important than others with more efforts invested in the monitoring and associated actions.

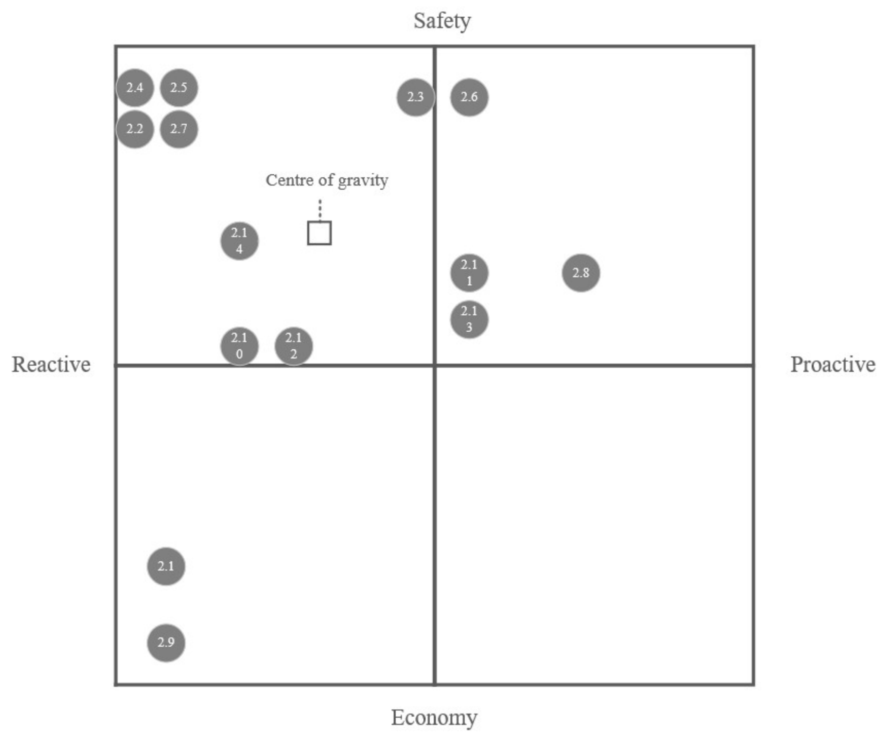


Figure 23: Illustration of how the Q4 - Balance framework can be applied

In the design and execution of the case studies, the primary focus was on safety-related indicators. Thus, the balancing issue of safety and economy has not been an explicit subject. The plotting of the indicator set below are therefore carried out solely based on the indicator set itself. It is also worth noting that the applied indicator sets that have been made available as part of the case studies, include few economy-related indicators. Even so, it would be reasonable to assume that all the companies use some efforts on financial monitoring, provided they are commercial actors.

Based on this, the plotted indicators (Figure 23) should be considered as an illustration of how the framework can be used, and not as actual representation. The indictor set of case company No.3 have been used as it includes some indicators that can be considered as economy related. These are indicators related budget performance (2.9) and On-hire (2.1). The latter are not a directly economy related indicator per se, but indirectly as off-hire will have and impact on financial performance. In this respect, I would, in the light of the framework's objective and theoretical foundation, consider other potential operational-related indicators as relevant under the economy dimension, i.e., a token of potential operational/commercial/economy safety trade-off.

I have further made a selection within the complete set of indicators limited to those that are categorized as KPI by the company (14 in total). The indicators are further plotted in the chart based on the analysis in chapter 5, with regard to proactive vs. reactive. Her can also the illustration in Figure 21 above be referred to when it comes to "degree of proactiveness/reactiveness". The plotted chart show that, based on the selected indicators, the focus, or "centre of gravity" is located in the reactive-safety quadrant. If the effect of safety energy were known and taken into account, the centre of gravity might have been shifted. For instance, if the indicators in the reactive-safety quadrant was found to consume more attention and effort compared with the other indicators, the centre of gravity would be shifted further towards the corner of the reactive-safety quadrant. If this was the case, it could be concluded that the indicator set represent minor potential for economy safety trade-off, and that the majority of the safety energy is allocated to reactive safety management. Provided that proactive safety management is the desired state, the current set should revise in order to shift its centre of gravity towards proactive quadrant. Again, it must be underlined that the set of indicators does not tell whole story about how the organization focus its attention and effort. This would of course need to be taken into account by the those that might apply the framework.

6 CONCLUSION

The main question: How can application of performance indicators improve safety management in shipping?

Application of performance indicators can improve safety management in shipping by enabling organizations to be proactive, i.e., to identify conditions and factors that influence risk of accidents and initiate actions before they occur. By adhering to relevant theories, methods and models, as included in this thesis, the organization can obtain the necessary understanding about how safety is brought about and use this to develop indicators that address the most important aspects of the organization and its context, while at same time obtaining a sound balance between the potential conflicting aspects of production and safety. The case studies reveal that there is a potential to shift the balance towards a more proactive safety management, supported by perspectives of this thesis.

6.1 The understanding of safety and the indicator concept

The main question of the thesis includes the key concepts, safety, safety management and performance indicators. One of the objectives of this thesis have been to establish an adequate understanding these elements. As the concept of safety performance indicate, the aim is to measure safety. Hench, we need to understand what safety is before we can know what to measure. This perspective, as a starting point for establishing understanding of what safety is and how it is brought about, is in my opinion a prerequisite for the concept to serve its purpose: improve safety management. Another closely related and important perspective is the view of an organization as socio-technical system, which comprise the interrelated functioning of both social and technical parts. This is important because it acknowledge that the system (organization) consist of both linear and non-linear relations, an insight that should be used to identify which model or methods to apply as basis for understanding safety. It is also important to recognise that a socio-technical system is an open system, meaning that its function is influenced by its surroundings, crating both risks and opportunities. Finally, the fundamental understanding of safety as dynamic emergent property, i.e., something that the system produces through the interrelated functioning of the system elements and its surroundings, balancing the need to be both efficient and thorough. This crate a contrast to a view of safety as static property that can established and left alone. There are of course other perspectives that are relevant, but the above mentioned is in my view the basic principles that must be included in order to utilize the potential of the indicator concept.

6.2 How indicators are applied by the shipping companies

When it comes to the question of how shipping companies apply safety performance indicators, the case studies reveal that there are both similarities and differences. The development and monitoring process can to a large extent concluded to be similar. Based on the analysis of the indicator sets, both individual, and combined, it can be concluded that the there is an overweight (combined app. 75 %) of reactive indicators. It can further be concluded that the majority of these indicators are related losses and deviations. The application of loss-based indicators can be explained based on the commonly understanding of safety as "the absence of unwanted outcomes such as incidents or accidents" and that such outcomes inevitable will have attention. They are after all the outcomes one wish to avoid. The widespread use of deviation-based indicators can, in my opinion, be explained by the compliance focus that is an attribute of the shipping industry in general, influenced by extensive regulatory framework and how it is enforced by authorities. A similar loss/deviation-based focus can also be found in the Shipping KPI framework which is the dominant performance indicator standard in the shipping industry. It can also be concluded that the indicator sets are influenced by the context of the shipping companies, e.g., the relevance of some indicators are dependent of vessel type and trade area, and some companies have clients that have specific requirements or expectations to performance indicators. The cases studies also revealed that respondents acknowledge the overweight of reactive indicators. At the same time, it was stated that they have made attempts to identify more proactive indicators have been made, but this was found to be challenging. Some also stated that the subject of potential proactive indicators has limited attention at the top management. It is important to point out that the indicators that are included in the indicator sets does not give a full representation of how the individual company practice safety management. Other measures are applied, such as safety culture/climate assessments, although not formally included as part of the indicators sets.

6.3 How theory, methods and models can be applied

The theory, methods and models included in this thesis have a two folded objective; to establish a general understanding of key aspects, as addressed in the former part of this chapter, and to support the development of indicators. When it comes to the question of how safety science theory can be used to develop indicators, both the theory that provide a general understanding of the key aspects and the specific methods and models must be taken into account. The methods and models for development of performance indicators that are included in this thesis, can be considered to represent the two categories, the complex-linear (Kjellén and Albrechtsen' s model and the HSE guide) and the non-linear (REWI, Lead drive and lead monitor, and FRAM)). Provided that a shipping company can be seen as socio-technical system which comprise of both linear and non-linear relations in its functioning I would conclude that all the perspectives that these contributions represent in general are suitable, and individually serve as valuable contribution. As all the methods and models have their limitations, I would recommend that a combination is sought, as illustrated in Chapter 6.2. Her I consider the model by Kjellén and Albrechtsen' (2017) as good "based-model" that can be supplemented by other perspectives. In addition to the described methods and models, several examples of indicators are included. In my point of view, these can serve as valuable inspiration and relevant candidates, but should not be adopted without an adequate process beforehand. By referring to the indicators applied by case study companies, which is mainly reactive oriented, the methods and models included in this thesis, which is mainly proactive oriented, can be applied to support a more proactive safety management.

6.4 How to obtain good balance

The aspect of balance relates to where the organization focus its attention and efforts represented by the indicator set as a whole, and correlation within the set of indicators, often represented by the distinction between proactive and reactive. By adhering to the above principles, the widely debated distinction between proactive and reactive indicators becomes subordinate. By this I mean that if the indicators are established based these principles and understanding, they will (most likely) be effective as a tool to support improvement of proactive safety management. As illustrated under the discussion of research question No. 3, the distinction can change depending on orientation (local or global) and by viewing interrelated activities and processes relative to each other. Thus, I would argue that it makes more sense to consider the degree of proactiveness or reactiveness, rather than a clear cut proactive or reactive. As I see it, the main value of adhering to the distinction is represented by the ability to establish correlation/causality between indicators (where relevant) as means to promote validity to the indicator set, and as means to assess the balance of the indicator set as a whole. A good balance is thus achieved when there is corelation between indicators and when the efforts related to monitoring and actions, in sum, are proactive, rather than reactive oriented. A good balance is also represented by the ability to see both safety-related and economy-related indicators in the same picture, avoiding appropriate efforts on one at the expense of the other. The "Q4 Framework" can be valuable tool in this respect.

6.5 Further research

Further research could address the following:

- The cases studies include a limited sample of the population shipping companies. A broader study with regard to number of samples, different segments, nationality, etc, would provide a richer understanding of how performance indicators are applied in the industry.
- The regulatory framework of the shipping industry is considered to have an influence on how performance indicators are applied and possible how safety is understood. A possible study could be to explore how instruments within this framework, such as the ISM Code, could be utilized or developed to better support proactive safety management, e.g., inspired by Resilience Engineering.
- The FRAM can provide valuable insight to the actual functioning of the socio-technical system. The method is however perceived to be time consuming. A possible study could be to explore how it can be adopted as cost-effective method throughout the system, e.g., combined with an ISO-process approach, while maintaining the systemic perspective.

7 REFERENCES

Akselson, R., Ek, Å., Floor, K., Stewart, S., & Marie, W. (2009). Resilience Safety Culture. *17th World Congress on Ergonomics 2009 IEA*.

Allianz. (2020). *Safety and shipping review 2020.* Allianz. Retrieved from https://www.imo.org/en/KnowledgeCentre/Pages/MaritimeFactsFigures-Default.aspx

Antonsen, S. (2009). Safety Culture Assessment: A Mission Impossible? *Journal of Contingencies and Crisis Management, 17*.

Antonsen, S., Heldal, F., & Kvalheim, S. A. (2017). *Sikkerhet og Ledelse.* Gyldendal Norsk Forlag AS.

Appelbaum, S. B. (1997). Socio-technical systems theory: an intervention strategy for organizational development. *Managment Decision*(35/6 452-463).

Baker, J., Bowman, F., Erwin, G., Gorton, S., Hendershot, D., Leveson, N., . . . Duane Wilson, L. (2007). *The Report of The U.S. Refineries Independent Safety Review Panel.*

Batalden, B.-M., & Sydnes, A. K. (2014). Maritime safety and the ISM code: a study of investigated casualties and incidents. *WMU Journal of Maritime Affairs, 13(1), 3-25*. doi: 10.1007/s13437-013-0051-8

Bhattacharya, S. (2011). The effectiveness of the ISM Code: A qualitative enquiry. (Marine Policy 36 (2012) 528-535). doi:10.1016/j.marpol.2011.09.004

BIMCO. (2020). The Shipping KPI Standard V4.0. Retrieved from https://www.shipping-kpi.org/public/downloads/documentation/Shipping_KPI_Standard_V4.0.pdf

Bloor, M., Sampson, H., Baker, S., Walters, D., Dahlgren, K., Wadsworth, E., & James, P. (2013). Room for Manoeuvre? Regulatory Compliance in the Global Shipping Industry. *Social & Legal Studies, 22(2) 171-189*. doi: 10.1177/0964663912467814

Bryman, A. (2016). *Social Research Methods.* Oxford University Press.

BS MoU. (2020). *Port state control in the black sea region- Annual report 2019.* Retrieved from http://www.bsmou.org/2020/07/annual-report-for-2019/

Cook, R., & Rasmussen, J. (2005). "Going solid": a model of system dynamics and consequences for patient safety. *BMJ Quality & Safety, 2005;14:130-134*. doi:https://doi.org/10.1007/s10111-021-00668-x

CSB. (2016). *Drilling rig explosion and fire at the Macondo Well.* U.S. CHEMICAL SAFETY AND HAZARD INVESTIGATION BOARD.

Dyreborg, J. (2009). The causal relation between lead and lag indicators. *Safety Science*(47 (2009) 474-475). doi:10.1016/j.ssci.2008.07.015

EPRI. (1999). *Guidelines for Leading Indicators of Human Performance: Preliminary Guidance for Use of Workplace and Analytical Indicators of Human Performance.* Palo Alto, CA.

EU. (2015). *REGULATION (EU) 2015/757 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the monitoring, reporting and verification of carbon dioxide emissions from maritime transport, and amending Directive 2009/16/EC.* Retrieved from https://eur-lex.europa.eu/eli/reg/2015/757/oj

EU. (2021). *Amending Directive 2013/34/EU, Directive 2004/109/EC, Directive 2006/43/EC and Regulation (EU) No 537/2014, as regards corporate sustainability reporting.* Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0189

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

Forskrift om melde- og rapporteringsplikt til sjøs. (2008). Retrieved from https://lovdata.no/forskrift/2008-06-27-744

Gui, F., Xuecai, X., Qingsong, J., Zonghan, L., Ping, C., & Ying, G. (2019). The development history of accident causation models in the past 100years: 24Model, a more modern

accident causation model. *Process Safety and Environmental Protection*(134 (2020) 47-82). doi:https://doi.org/10.1016/j.psep.2019.11.027
Hale, A. (2009). Why safety performance indicators? *Safety Science*(7 (2009) 479–480). doi:10.1016/j.ssci.2008.07.018
Hasanspahić, N., Frančić, V., Vujičić, S., & Maglić, L. (n.d.). Reporting as a Key Element of an Effective Near-Miss Management System in Shipping. *MDPI - Safety, 6 (4), 53*. doi: https://doi.org/10.3390/safety6040053
Herrera, I. A. (2012). *Proactive safety performance indicators (Doctoral thesis).* NTNU.
Hollnagel, E. (2004). *Barriers and accident prevention.* Taylor & Francis.
Hollnagel, E. (2009). *The ETTO Principle: Efficiency-Thoroughness Trade-Off.* Ashgate Publishing Limited.
Hollnagel, E. (2012). *FRAM: The Functional Resonance Analysis Method.* Ashgate Publishing limited .
Hollnagel, E. (2014). *Safety-I and Safety-II - The Past and Future of Safety Management.* Taylor & Francis Group.
Hollnagel, E., & Speziali, J. (2008). *Study on Developments in Accident Investigation Methods: A Survey of the "State-of-the-Art".* Swedish Nuclear Power Inspectorate (SKI).
Hollnagel, E., Woods, D. D., & Leveson, N. (2006). *Resilience Engineering - Concepts and Precepts.* CRC Press, Taylor & Francis Group.
Hopkins, A. (1994). *The limits of lost time injury frequency rates.* Worksafe Australia.
Hopkins, A. (2009). Thinking about process safety indicators. *Safety Science*(47 (2009) 460-465). doi:10.1016/j.ssci.2007.12.006
HSE . (2006). *Developing process safety indicators - A step-by-step guide for chemical and major hazard industries.* HSE (Health and Safety Executive) Books.
IAEA. (2000). *Operational safety performance indicators for nucelar power plants.*
IMO. (1974). International Convention for the Safety of Life at Sea.
IMO. (1993). *ISM Code (International Safety Management Code).*
IMO. (1993). Resolution A.741(18). *INTERNATIONAL MANAGEMENT CODE FOR THE SAFE OPERATION OF SHIPS AND FOR POLLUTION PREVENTION*.
IMO. (2016). *RESOLUTION MEPC.278(70) - Amendments to MARPOL Annex VI (Data collection system for fuel oil consumption of ships).* Hentet fra https://wwwcdn.imo.org/localresources/en/OurWork/Environment/Documents/278(70).pdf
Jacobsen, D. (2018). *Hvordan gjennomføre undersøkelser? Innføring i samfunnsvitenskapelig metode.* Cappelen Damm AS.
Kee, D. (2017). Comparison of Systemic Accident Investigation Techniques Based on the Sewol Ferry Capsizing. *Journal of the Ergonomics Society of Korea*(Volume 36, Issue 5, p. 485-498). doi:dx.doi.org/10.5143/JESK.2017.36.5.485
Kjellén, U. (2009). The safety measurement problem revisited. *Safety Science*(47 (2009) 486-489). doi:10.1016/j.ssci.2008.07.023
Kjellén, U., & Albrechtsen, E. (2017). *Prevention of Accidents and Unwanted Occurrences.* Taylor & Francis Group.
Kongsvik, T., Albrechtsen, E., Antonsen, S., Herrera, I. A., Hovden, J., & Schiefloe, P. (2018). *Sikkerhet i arbeidslivet.* Fagbokforlaget.
Kongsvik, T., Almklov, P., & Fenstad, J. (2010). Organisational safety indicators: Some conceptual considerations and a supplementary qualitative approach. *Safety Science, 48 (2010) 1402-1411*. doi:doi:10.1016/j.ssci.2010.05.016
Kvale, S., & Brinkmann, S. (2021). *Det kvalitative forskningsintervju.* Gyldendal Norsk Forlag AS.

Leveson, N. (2014). A systems approach to risk management through leading safety indicators. *Reliability Engineering and System Safety*(136 (2015) 17-34). doi:http://dx.doi.org/10.1016/j.ress.2014.10.008

OCIMF. (2017). *Tanker Management and Self Assessment - A Best Practice Guide.*

OECD. (2008). *Guidance on developing performace indicators.* Retrieved from https://www.oecd.org

OECD. (2022). *OECD - The ocean*. Hentet fra https://www.oecd.org/ocean/topics/ocean-shipping/

Paris MoU. (2020). *Annual report 2019.* Retrieved from https://www.parismou.org/publications-category/annual-reports

Paris MoU. (2021). Retrieved from Paris MoU on Port State Control: https://www.parismou.org/

Patriarca, R., & Bergström, J. (2017). Modelling complexity in everyday operations: functional resonance in maritime mooring at quay. *Cognition Technology & Work*(november 2017). doi:10.1007/s10111-017-0426-2

Patriarca, R., Bergström, J., Di Gravio, G., & Costantino, F. (2018). Resilience engineering: Current status of the research and future challenges. *Safety Science, Safety Science 102 (2018) 79-100*. doi:http://dx.doi.org/10.1016/j.ssci.2017.10.005

Perrow, C. (1999). *Normal Accidents.* Princton University Press.

Pidgeon, N. F. (1991). Safety Culture and Risk Management in Organizations. *Journal of cross-cultural psychology, 22 (1) 129-140*. doi:10.1177/0022022191221009

Praetorius, G., Lundh, M., & Lützhöft, M. (2011). Learning from the past for pro-activity – A re-analysis of the accident of the MV Herald of Free Enterprise. doi:https://doi.org/10.4000/BOOKS.PRESSESMINES.1089

Provan, D. J., Woods, D. D., Dekker, S. W., & Rae, A. J. (2019). Safety II professionals: How resilience engineering can transform safety. *Reliability Engineering and System Safety*(195 (2020) 106740). doi:https://doi.org/10.1016/j.ress.2019.106740

Rasmussen, J. (1997). Risk Management i a Dynamic Society: A Modelling Problem. *Safety Science, 27 (1997) 183-213*. doi:https://doi.org/10.1016/S0925-7535(97)00052-0

Reason, J. (1997). *Managing Risks of organizational Accidents.* Ashgate Publishing.

Reiman, T., & Pietikäinen , E. (2010). *Indicators of safety culture – selection and utilization of leading safety performance indicators.* Swedish Radiation Safety Authority.

Reiman, T., & Pietikäinen, E. (2011). Leading indicators of system safety – Monitoring and driving the organizationalsafety potential. *Safety Science*(50 (2012) 1993-2000). doi:http://dx.doi.org/10.1016/j.ssci.2011.07.015

Safety Sience. (2009). Process Safety Indicators. (47 (2009) 459-568).

Salihoglu, E., & Bes¸ikçi b, E. B. (2011). The use of Functional Resonance Analysis Method (FRAM) in a maritime accident: A case study of Prestige. *Ocean Engineering, 219(2011) 108223*. doi:https://doi.org/10.1016/j.oceaneng.2020.108223

Shirali , G., Shekari , M., & Angali , K. (2018). Assessing Reliability and Validity of an Instrument for Measuring Resilience Safety Culture in Sociotechnical Systems. *Safety and Health at Work, 9 (2018) 296-307*. doi:http://dx.doi.org/10.1016/j.shaw.2017.07.010

SINTEF. (n.d.). *Building Safety*. Retrieved 2022, from SINTEF: https://www.sintef.no/projectweb/building-safety/

Sjøfartsdirektoratet. (2017). *Risikobasert tilsyn*. Retrieved 2022, from Sjøfartsdirektoratet: https://www.sdir.no/sjofart/fartoy/tilsyn/risikobasert-tilsyn/

Sjøfartsdirektoratet. (2020). *Personskader 1981 - 2019 [Data set].* Retrieved from https://www.sdir.no/sjofart/ulykker-og-sikkerhet/ulykkesstatistikk/

Sjøfartsdirektoratet. (2020). *Spørreundersøkelse maritim sikkerhet [Datasett].*

Smith, D., Veitch, B., Khan, F., & Taylor, R. (2018). Using the FRAM to Understand Arctic Ship Navigation: Assessing Work Processes During the Exxon Valdez. *TransNav*(Voulme 12, No. 3). doi:10.12716/1001.12.03.03

Størkersen, K. V. (2018). *Bureaucracy overload calling for audit implosion (Doctoral thesis, NTNU).* Retrieved from https://samforsk.no

Størkersen, K. V. (2021). Safety management in remotely controlled vessel operations. *Marine Policy*(130 (2021) 104349). doi:https://doi.org/10.1016/j.marpol.2020.104349

Sujan, M., Petriarca, R., Constantino, F., & Villani, M. L. (2021). Developing Leading Safety Indicators using the Functional Resonance Analysis Method. *Safety-Critical Systems Symposium 2021.* York. Retrieved from https://www.researchgate.net/publication/347556295_Developing_Leading_Safety_Indicators_using_the_Functional_Resonance_Analysis_Method

Tokyo MoU. (2020). *Annual report on port state control in the asaia-pacific region.* Retrieved from http://www.tokyo-mou.org/

Tzannatos, E., & Kokotos, D. (2008). Analysis of accidents in Greek shipping during the pre- and post-ISM period. *ELSEVIER*(Marine Policy 33 (2009) 679-684). doi:10.1016/j.marpol.2009.01.006

UNCTAD. (2021). *Review of maritime transport.* Hentet fra https://unctad.org/system/files/official-document/rmt2021_en_0.pdf

US Coast Guard. (2020). *Port state control ain the united states - 2019 annual report.* US Coast Guard. Retrieved from https://www.dco.uscg.mil/Our-Organization/Assistant-Commandant-for-Prevention-Policy-CG-5P/Inspections-Compliance-CG-5PC-/Commercial-Vessel-Compliance/Foreign-Offshore-Compliance-Division/Port-State-Control/Annual-Reports/

Vandeskok, B. (2015). The Legitimacy of Safety Management Systems in the Minds of Norwegian Seafarers. *TransNav, 9(1)*. doi:DOI: 10.12716/1001.09.01.12

Woods, D. D., Branlat, M., Herrera , I., & Woltjer, R. (2015). Where Is the Organization Looking in Order to Be Proactive about Safety? - A Framework for Revealing whether It Is Mostly Looking Back, Also Looking Forward or Simply Looking Away. *Journal of Contingencies and Crisis Management, 23 (2) 2015*. doi:10.1111/1468-5973.12079

Wreathall, J. (2009). Leading? Lagging? Whatever! *Safety Science*(47 (2009) 493-494). doi:10.1016/j.ssci.2008.07.031

Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods .* SAGE Publications, Inc.

Yuen, K. F., Loh, H. S., Zhou, Q., & Wong, Y. D. (2018). Determinants of job satisfaction and performance of seafarers. *Transportation Research, 110 (2018) 1-12*. doi:https://doi.org/10.1016/j.tra.2018.02.006

Øien, K. (2001). A framework for the establishment of organizational risk indicators. *Reliability Engineering & System Safety*(74 (2001) 147-167). doi:10.1016/S0951-8320(01)00068-0.

Øien, K., Massaiu, S., Tinmannsvik, R. K., & Størseth, F. (2010). Development of Early Warning Indicators based on Resilience Engineering. *Paper presnted at PSAM 10, June 7-11 2010, Seattle, USA.*

Øien, K., Utne, I., & Herrera, I. (2009). Building Safety indicators: Part 1 – Theoretical foundation. *Safety Science*(49 (2001) 148-161). doi:10.1016/j.ssci.2010.05.012

APPENDIX 1: Examples of indicators – input-process-output

Examples retrieved from Kjellén & Albrechtsen (2017)

Type of contributing factor at the workplace	Example of performance indicators
M – Human/behavioural	Management and workforce engagement: • % of inspections attended by management • Staff attitude survey outcome • % of workforce suggestions implemented % of staff assigned to hazardous tasks that meet task qualification requirements % of required safety induction of contractors conducted on schedule
T – Technical/physical	% of inspection and preventive maintenance tasks performed on schedule ('backlog')
O – Organisational/economic	% of job procedures for hazardous work based on risk assessments % of regulatory requirements implemented in job procedures % of job descriptions defining safety responsibilities % of sampled permit-to-work (PTW) that identified the required controls % of plant design changes subject to risk assessment % of completed statutory training requirements % of planned safety training completed % of high-potential (HIPO) incidents not identified in risk assessment % of HIPO incidents subject to Level 3 investigation Degree of learning from incident investigations % of close-out of inspection findings (within one week) Number of stop-work events per employee and year % of emergency response elements that are fully functional when tested % of on-scene emergency staff that has participated in a drill per quarter/year

Figure 24: CF based (work place factors) indicators

Corporate general and HSE management systems	Examples	Reference
General management		
	Tripod Delta, standardised questionnaire on 'general failure types' in the organisation	Van der Want 1997; HSE 2003
HSE management system		
Performance measurement systems	% of compliance with international standard (e.g. OHSAS 18001) ISRS Self-diagnostic tools	Alteren 1999; SCS 2003; Roy et al. 2008; DNV GL 2014
Measurement of specific elements (structure, activity)	HSE management elements established (policy) Resources devoted to HSE in % of budget % of HSE management activities completed as planned	
Safety culture		
Performance measurement systems	Employee safety culture perception survey Safety Climate Tool Safety Culture Maturity assessment	Energy Institute 2008c; Zohar 2010; HSE 2013
Measurement of specific elements	Staff perception of management commitment to safety Number of senior management site safety tours completed per individual and year Board and top-management visibility and involvement	SCS 2003; ICMM 2012

Figure 25: CF based (General and HSE MS) indicators

Type	Definition
Employee self-reporting	
Incident (near-miss) rate	Number of reported incidents (near misses) per period.
RUO rate	Average number of employee reported unwanted occurrences (RUO) including incidents, unsafe acts, and unsafe conditions per employee and year. This includes spontaneous reporting by the employees, excluding reports by health, safety, and environment (HSE) staff and from regular safety activities such as inspections.
Incident/TRI ratio	Ratio between the number of reported incidents in a period, and the number of total recordable injuries (TRIs) in the same period.
High-potential (HIPO) incidents	
HIPO incident rate	Number of high-potential incidents per period.
Process safety events (PSE)	
PSE rate	Total number of PSE, that is, events involving loss of primary containment of any material from primary containment above a defined threshold per million hours worked in the activity (OGP 2011).

Figure 26: Incident based indicators

Type of deviation and examples	Example of performance indicators	Reference
General	% of compliance with statutory requirements	ICMM 2012
	RUO rate (i.e. the average number of reported unwanted occurrences including incidents, unsafe acts, and unsafe conditions per employee and year)	
Work situation		
1. *Human error* (e.g. wrong action, wrong sequence, omission)	Behavioural sampling (% of correct behaviour)	Krause et al. 1999
2. *Technical failure* (e.g. machine failure, missing equipment/tools)	% of safety critical equipment that performs to specification when inspected/tested	Health and Safety Executive 2006
3. *Disturbance in material flow* (e.g. poor quality, delays)	Mean time between failures (MTBF)	Rausand and Høyland 2004
4. *Personnel deviation* (e.g. absence, inadequate qualifications, indisposed)	% of staff having required level of training/qualifications when checked	
5. *Inadequate information* (e.g. inadequate or missing instructions/procedures, permit-to-work [PTW])	% of PTW filled in correctly	Health and Safety Executive 2006
Environment		
1. *Intersecting or parallel activities* (e.g. disturbance from other work team)	% of simultaneous operations complying with simultaneous operations procedure	
2. *Bad housekeeping*	Housekeeping index (rating of housekeeping standard from 1 to 5 per plant area, see Section 17.2.3)	
3. *Disturbances from the environment* (e.g. excessive noise, wind speed, precipitation, high or low temperature)		
Safety systems		
1. *Failure of active or passive safety systems*	% of safety systems/barriers performing to specifications	Health and Safety Executive 2006
2. *Inadequate guarding*	Covered by failure of passive safety systems	
3. *Inadequate personal protective equipment or clothing*	Behavioural sampling	Krause et al. 1999
4. *Inadequate emergency response* (e.g. delayed notification and mobilisation, failure to evacuate)	Number of emergency response elements not fully functional when activated in exercise/real emergency	OGP 2011

Figure 27: Process based (deviation based) indicators

APPENDIX 2: Examples of indicators – RWEI-method

Examples retrieved from Øien, Massaiu, Tinmannsvik, & Størseth (2010)

No.	CSF level 2	General issue
1.1	Risk understanding	
1.1.1		System knowledge
1.1.2		Information about risk through e.g. courses & documents
1.1.3		Reporting of incidents, near-misses and accidents
1.1.4		Information about the quality of barriers (technical safety)
1.1.5		Information about the quality of barrier support functions (oper. safety)
1.1.6		Discussion of HSE issues/status in regular meetings
1.1.7		Safety performance matters requested by senior management
1.1.8		Communicating risk/resilience at all levels of the organization
1.2	Anticipation	
1.2.1		Risk/hazard identification
1.2.2		Learn from own experiences & accidents
1.2.3		Learn from other's experiences & accidents
1.3	Attention	
1.3.1		Process disturbances; control and safety system actuations
1.3.2		Bypass of control and safety functions
1.3.3		Activity level/simultaneous operations
1.3.4		Trends in reported events and quality of barriers
1.3.5		Early warnings/weak signals (e.g. from whistle blowers)
1.3.6		Changes; technical, organizational, external (weather, …)
1.3.7		Focus on safety (safety versus other issues)

Figure 28: General issues ford CSF1 - Risk Awareness

No.	General issue	Candidate indicator
1.1.1	System knowledge	
1.1.1.1		Average no. of years experience with such systems
1.1.1.2		Average no. of years experience with this particular system
1.1.1.3		Portion of operating personnel involved during design & construction
1.1.1.4		Average no. of hours system training last 3 months
1.1.1.5		Portion of operating personnel receiving system training last 3 months
1.1.1.6		No. of violations to authorized entrance of systems
1.1.1.7		Portion of operating personnel familiar with design assumptions
1.1.1.8		Turnover of operating personnel last 6 months
1.1.2	Info. about risk	
1.1.2.1		Portion of operating personnel taking risk courses last 12 months
1.1.2.2		Portion of staffing taking risk courses last 12 months
1.1.2.3		Portion of operating personnel informed about risk analyses last 3 months
1.1.2.4		Average no. of SJA[1] operating personnel have attended last month
1.1.2.5		No. of different persons having facilitated/led SJA during last month
1.1.2.6		No. of tool-box meetings last month
1.1.2.7		No. of violations to assumptions/limitations in the risk analysis (QRA[2])

Figure 29: General issues for CSF1 - Risk Awareness

APPENDIX 3: Examples of indicators – lead drive and lead monitor

Examples retrieved from Reiman & Pietikäinen (2011)

Indicator area	Measures
Safety management and leadership	E.g., (1) management is actively committed to, and visibly involved in, safety activities, (2) number of management walk arounds per month, (3) number of times safety is a topic in the management meetings and (4) conservative decision-making principles are applied in making decisions about the operational safety of the plant
Strategic management	E.g., (1) safety is visibly and systematically considered in the organization's official plans and strategy documents, (2) systematic ageing management program exists for systems, components and structures, (3) a program of preventive maintenance is in place and it is revised according to maintenance history and (4) there is a system for documenting history data on equipment and their maintenance actions
Supervisor activity	E.g., (1) superior provides positive feedback on safety-conscious behavior of the personnel
Proactive safety development	E.g., (1) a system for reporting and analyzing incidents is implemented, (2) independent safety reviews and audits are carried out regularly and proactively, (3) there is a system for gathering development initiatives from the personnel and (4) there is a system for analyzing the common safety-related findings (trends, root causes, changes, variety of corrective actions, generalizability to other components/equipment) from the maintenance history as well as events and near misses in the organization
Competence management	E.g., (1) an adequate system exists for the identification of current competence profiles, (2) there are clear objectives established for training programs, (3) a mechanism is in place to ensure that the scope, content and quality of the training programs are adequate and (4) feedback is gathered from the trainees and is utilized in developing the training program
Change management	E.g., (1) there is a clear definition of what constitutes a technical change or an organizational change in the safety policy of the organization, (2) risk assessment is done for organizational changes, (3) there is a procedure for planning, implementing and follow-up of technical and organizational changes and (4) the effects of the implementation period to organizational practices is monitored during the change
Work conditions management	E.g., (1) the availability of sufficient workforce is controlled, (2) procedures are updated regularly and (3) number of outdated procedures
Work process management	E.g., (1) The bottlenecks of information flow are identified and controlled and (2) tasks and situations where routines may develop and where they might have consequences for safety are identified
Contractor management	E.g., (1) there is a process for purchasing outside work, (2) a record of contractor safety performance is utilized in decision making concerning contracts and (3) contractors are trained on safety culture issues and work practices of the client organization
Hazard control	E.g., (1) a systematic corrective action program is in place to deal with deviations, (2) hazard identification and risk assessments are used to develop policies, procedures and practices, (3) adequate barriers are set against the identified hazards, (4) the organization has analyzed potential accident scenarios and set barriers to prevent them, (5) there are adequate human performance tools (HPT) to facilitate safe behavior and (6) surgical checklist is used in all surgeries (%)
Contingency planning and emergency preparedness	E.g., (1) the organization has an adequate on-site emergency preparedness plan and (2) there is regular training on emergencies on-site

Figure 30: Examples of drive indicators

Table 3
Examples of monitor indicators.

Indicator area	Measures
Work and safety motivation	E.g., (1) the extent to which the personnel report that their work is meaningful and important, (2) the extent to which human performance tools are utilized in daily practice and (3) the extent to which personnel consider safety as a value that guides their everyday work
Controllability of work	E.g., (1) employees' reported sense of control over their work, (2) staffing, (3) the extent to which work is carried out in accordance to the processes described in the management system and (4) the amount of slack resources to cope with unexpected or demanding situations
Understanding of hazards	E.g., (1) the extent to which the personnel understands the hazards that are connected to their work, (2) the extent to which the personnel has been trained in accordance with the planned training program, (3) the extent to which the personnel are aware of the limitations of human performance capacity, (4) the extent of personnel's awareness of the technical/physical condition of systems, structures and components and (5) the findings from external audits concerning hazards that have not been perceived by personnel/management previously
Understanding of safety	E.g., (1) The extent to which the personnel have basic knowledge of human performance issues, (2) the extent to which the defense-in-depth principle is understood among the personnel, (3) the extent to which Human Factors are considered neutral phenomena and not something to be avoided (i.e., a negative phenomenon) and (4) the extent to which changes and improvements are considered at system level as opposed to unit or group level
Felt responsibility for the entire organization	E.g., (1) the extent to which the personnel are willing to spend personal effort on safety issues and take responsibility for their actions and (2) the extent to which the personnel make initiatives in improving organizational practices or report problems to the management
Mindfulness and vigilance	E.g., (1) the extent to which the personnel continuously seek to identify new risks and enhance their view on the hazards of their work, (2) the extent to which the personnel at all levels exhibit a questioning attitude and (3) the extent to which external audits provide results that are in accordance with the findings in internal audits or prevalent conceptions of the personnel
Social interaction and activities	E.g., (1) the extent to which safety-conscious behavior and uncertainty expression is socially accepted and supported, (2) the extent to which the gap between work as prescribed and work as actually done is known and monitored in the organization and (3) the extent to which the personnel perceive that they have to make tradeoffs between safety and economy in daily work
Technology	E.g., (1) continuous measures of the current condition of systems, components and structures and (2) percentage of safety–critical equipment that fail inspection/test
Environmental variability	E.g., (1) extreme weather phenomena for process plants and (2) age distribution of the population for healthcare organizations

Figure 31: Examples of monitor indicators