Ivonne Andrade Herrera

Proactive safety performance indicators

Resilience engineering perspective on safety management

Thesis for the degree of Philosophiae Doctor

Trondheim, February 2012

Norwegian University of Science and Technology
Faculty of Engineering Science and Technology
Department of Production and Quality Engineering

NTNU – Trondheim
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Engineering Science and Technology
Department of Production and Quality Engineering

ISBN 978-82-471-3589-1 (printed ver.)
ISBN 978-82-471-3590-7 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2012:151

Printed by NTNU-trykk

Abstract

The objective of the thesis is to present studies addressing the question: *How do we identify that a system drifts or experiences sudden changes in the safety space?* Hence, the studies in this thesis focus on the identification of safety performance indicators. This identification is influenced by the way safety is understood. If safety is seen as the absence of accidents and failures, then indicators refer to failures, malfunctions or deviations. Safety critical organizations like airlines, air navigation service providers or oil companies have implemented numerous improvements by using such lagging indicators. These lagging indicators by their nature provide information after the fact. Nevertheless, today's systems and organizations must be able to function in rapidly changing environments in which there is a great deal of uncertainty. The Resilience Engineering perspective does not see safety as the absence of failures but as something that the organization or the socio-technical system does. Thus successes and failures are related to the capability of the system or organization to adjust and continue operations in the presence of continuous changes and operational constrains. Given such a perspective, finding indicators that allow an organization to act before something happens, i.e. to be leading, rather than reactive and lagging, is a main challenge.

In order to address this challenge, the work starts with a systematic review of existing methods that could be applied for the identification of safety indicators. The thesis explores methods accounting for the monitoring of failures, deviations and everyday performance. In everyday operation, there are many successes i.e. flights arrive on schedule without significant problems. Hence, the empirical part of the work deals with the understanding the deficiencies and strengths of specific incidents and daily operations. Established and relatively new methods are tested. Each method represents a different perspective on safety having an influence on the identification of indicators. Using a multidisciplinary analysis that combines methods with different perspectives provides broad understanding. The studies document the application of the following methods: 1) Triangulation using a list of outcome and activity indicators, 2) Sequentially Timed Events Plotting (STEP method), 3) storytelling and 4) the Functional Resonance Analysis Method (FRAM method). Findings from the studies and their limitations have been presented and discussed with industry and the research community.

It is essential to understand how a system operates under operational and financial constraints in order to identify that the system drifts or experiences sudden changes. Therefore, the main argument in this thesis is the identification of indicators related to failures and everyday successful operations. For that purpose, it is reasonable to distinguish from among three types of performance indicators: 1) lagging indicators, which refer to what has occurred in the past; 2) current indicators, which refer to what is occurring now; and 3) leading indicators, which refer to what may occur in the future. While, there is considerable data available for lagging indicators, a balanced composition of lagging, current and leading indicators is needed. Hence, the thesis represents a step forwards for the identification of current and leading indicators. The test of the methods provides a tool box for industry. The work also contributes to develop the FRAM method. While most of the subjects are related to aviation some ideas and methods presented in this thesis are useful to other safety critical organizations such as industry in the nuclear, oil and gas or railway sectors.

Acknowledgments

This thesis could not have been possible without support from my family including my in-laws and friends. Special thanks to my husband Snorre and my children Isak and Ingeborg for their love and support during these four years.

I am very grateful to my main supervisor Professor Jørn Vatn at the Department of Production and Quality Engineering and co-supervisor Professor Jan Hovden at the Department of Industrial Economy and Technology Management (both at the Norwegian University of Science and Technology - NTNU) for sharing their knowledge, experience and for their stimulating discussions. I appreciate their open minds and critical awareness and most of all the freedom I had to explore and learn. I want to extend special gratitude to Professor Erik Hollnagel at the Department of Industrial Safety – MINES ParisTech in Sophia Antipolis, France for his challenging discussions. All three professors represent different perspectives which provided me with great help and raised some problems. I am indebted to Professor Erik Hollnagel for facilitating my stay, discussions and good time with his group. I have been very fortunate to be a visiting scientist at the MINES ParisTech, Industrial Safety Chair in Sophia Antipolis, France. I would like to acknowledge fruitful discussions with Luigi Macchi, Erik Rigaud, Denis Besnard, Damian Fabre and Eduardo Runte.

The thesis was financed by Trygg Insurance Company in Norway. The six months abroad at MINES ParisTech was financed by the Fund of the Norwegian Institute of Technology and Det Norske Veritas. I am truly thankful for this financial support.

I want to thank my co-authors for their contributions and support. Regarding the studies presented in this thesis, I would like to acknowledge the openness and support from the aviation community, in particular Håkan Bengtson and Fredrik Strand (Scandinavian Airlines), Trond Lamark (Bristow), Glenn Christenssen (CHC), Magnus Jerpstad (AVINOR) and Erik Hamremoen (Statoil). Stewart Clark's (NTNU) assistance in editing the thesis and papers is highly appreciated.

I would like to thank my colleagues for sharing their knowledge, support, discussions and comments on papers. In particular, I would like to mention and acknowledge Erik Jersin, Per Hokstad and Camilla K. Tveiten. My new colleagues at the Department of Software Engineering, Safety and Security at SINTEF and my old colleagues Eirik Albrechtsen, Solfrid Håbrekke for creating a stimulating work environment. Coming from industry to research, I thrived on the challenging books and literature that Ragnar Rosness introduced for me. This new world inspired some of the work addressed in this thesis.

Finally, the work presented in this thesis is result from interaction with other researchers that have influenced my work. Therefore last but not least, I would like to acknowledge John Wreathall, Teemu Reiman and Pia Oedewald, Arthur Dijkstra, Marvin Rausand and Urban Kjéllen. I also want to thank Andrew Hopkins for his positive response to my mails and coming to Norway to discuss his work on safety performance indicators.

List of publications

Systematization of existing concepts and theories:

- **Paper I (Journal paper):**
  Hovden, J., Albrechtsen, E., Herrera, I., A. (2010). Is there a need for new theories, models and approaches to occupational accident prevention? *Safety Science*, vol 48, pp. 950-956

- **Paper II (Journal paper):**
  Øien, K., Utne, I., B., and Herrera, I.A. (2011). Building Safety Indicators, Part 1 - Theoretical foundation. *Safety Science*, vol. 49, pp. 148-161

- **Paper III (Conference paper):**
  Herrera, I.A., Hovden, J. (2008). Leading indicators applied to maintenance in the framework of resilience engineering: A conceptual approach. *Proceedings of the Resilience Engineering Symposium III*, Juan Les Pins, France

Empirical study: looking at the past, study of an aviation incident

- **Paper IV (Journal paper):**
  Herrera, I. A., Woltjer, R. (2010). Comparing a multi-linear (STEP) and systemic (FRAM). Reliability Engineering and System Safety, vol. 95 pp. 1269-1275

Empirical studies: looking at the past and present

- **Paper V (Journal paper):**
  Herrera, I., A., Nordskag, A.O., Myhre, G., Halvorsen, K. (2009). ”Aviation safety and maintenance under major organizational changes, investigating non-existing accidents”. Journal of Accident Analysis and Prevention, vol. 41 pp. 1155-1163

- **Paper VI (Conference paper):**
  Herrera, I.A., Tinmannsvik, R., K. (2006). Key elements to avoid drifting out of the safety space. *Proceedings of the Resilience Engineering Symposium II*, Juan Les Pins, France

Empirical studies: looking at the present

- **Paper VII (Conference paper):**
  Herrera I.A., Hollnagel, E., Håbrekke, S. (2010). Proposing safety performance indicators for helicopter offshore on the Norwegian Continental Shelf. *Proceedings of the 10th International Probabilistic Safety Assessment and Management Conference (PSAM)*, Seattle, USA

- **Paper VIII (Conference paper)**
  Herrera, I.A., Forseth U., Hokstad, P., Håbrekke, S., Kråkenes, K. (2010). Approaches to elaborate on the safety of helicopter operations. *Working on safety* 2010, Røros, Norway.

Table of Contents

List of Figures

List of Tables

PART I – Summarizing papers contribution

1 Introduction

1.1 Motivation

The aviation safety records show a stable accident rate while there is a concurrent increase in the number of passengers. Efforts were focused on the prevention of Controlled Flight into Terrain (CFIT) type of accidents. Analysis of accident data confirms that these efforts have been successful. The International Civil Aviation Organization's (ICAO) working paper (2006) emphasizes that without a new safety breakthrough, it is likely that the number of non-CFIT accidents will increase in line with the increase in traffic. One ICAO's recommendation is to have a more proactive approach to prevent accidents. The proactive approach includes a scientific approach to risk assessment, human factors and development of the means of collecting and analysing data. By the beginning of 2011, ICAO launched a Continuous Monitoring Approach (CMA) as a more proactive framework to monitor safety oversight capabilities frequent basis incorporating the analysis of safety risk factors. This new initiative illustrates how the aviation community acknowledges the need for continuous monitoring rather than one-time snapshot. Every element of the Air Transport System (ATS) is interdependent with many points of interactions and extensive use of advanced technology. Therefore, approaches for monitoring should be tailored to the level of complexity that currently exists in aviation.

This thesis is motivated from my professional background and identified needs mostly in aviation and but also from the oil and gas industry. I am educated as an electrical engineer with a master's degree in aeronautical maintenance and production. I take advantage of nine years' experience from engineering and maintenance work in different airlines such Avianca (Colombia), Air France, Braathens (Norway), two years of project management for Air Traffic Management control system deliveries to international airports like Kuala Lumpur Airport and Madrid-Barajas Airport and eight years of safety research for the aviation and oil and gas industries. I am also invited as independent expert acting as evaluator or as reviewer for the evaluation of research proposals and initiatives of the 6th and 7th Framework Programmes for research, technological development and demonstration of aeronautical activities by the European Commission (EC).

Since, I started my engineering work in aviation most of the modifications and repairs were performed to improve performance and safety. These improvements are usually responses to address accidents, failures or to enhance performance. Examples are the traffic collision avoidance system to reduce mid-air collision between aircraft or the modification to reduce vertical separation minimum between aircraft to provide additional cruising levels. On the positive side, these developments are seen as technical and operational improvements. They demand a significant change in the regulations, the procedures and practices of airlines and air navigation service providers. On the negative side, these changes increase dependencies and complexities which may lead to a new type of accidents i.e. the Überlingen mid-air collision accident. In parallel, to technical and operational improvements, the aviation industry experiences concurrent organizational changes, fierce competition and economic pressure. The aviation industry is subjected to high safety standards, to be more efficient and to be cost effective. There is concern about the reduction in safety margins and its safety impact.

In the period between 2004 and 2005, I was involved on a study carried out by the Norwegian Accident Investigation Board to analyse the relation of concurrent organizational changes and safety (AIBN, 2005). A need was identified in relation to the indicators used by the industry which are mainly lagging ones. Real-time performance monitoring is used for technical equipment but there is a need to have better tools to address performance monitoring of the whole socio-technical system. The introduction of CMA by ICAO represents a step forward toward more real-time performance monitoring. The CMA concept is very recent and the motivation of this thesis is to look for ways to improve the monitoring of operations to be better prepared and more resilient. Hence, by using a combination of practical experience and theoretical knowledge the thesis contributes by making a step forward towards a more proactive approach to safety in the aviation industry. While most of the subjects are related to aviation some ideas and methods presented in this document are useful for other safety critical organizations such industry in the nuclear, oil and gas or railway sectors.

1.2 Objectives and research questions

The objective of this PhD work is to discuss and perform studies to answer the main question: *How do we identify that a system drifts or experiences sudden changes in the safety space?* The main concern of this PhD work is the reduction of safety margins and its consequences. The study focuses on the identification of safety performance indicators from accidents, incidents, near misses and everyday operations (meaning operations when nothing goes wrong). Safety critical systems or organizations like aviation organizations are studied looking into the interaction between people as individuals or organizations and modern technical systems. Several researchers argue that these systems cannot solely be described by considering independent components and failures (Hollnagel, 2004, 2009; Reason, 2008; Hitchins, 2004; Rigaud and Guarnieri, 2006). Therefore, this work needs to use various methods to highlight factors, a combination of factors or identify patterns that may affect safety.

The aims are to explore methods for the identification of safety performance indicators. The methods should be able to

- monitor and identify changes affecting safety
- inform personnel of the threats and opportunities associated with their work
- serve as a tool for a safety analyst to identify safety improvements for operations in a specific context

Sub-questions and specific objectives are used to answer the main question. The research is based on different studies addressing the sub-questions. The sub-questions and their specific objectives are:

- What do we mean by safety performance indicators and related terms like leading and lagging indicators? (Papers III & VI). The specific objective is to classify current knowledge in relation to indicators that could be applied in the identification of drift or sudden changes.

- Do we need new methods for the identification of safety performance indicators? (Papers II & III) The specific objective is to understand the need for new methods to monitor safety performance.

- Which methods can be applied to identify lagging and leading indicators in the perspective of Resilience Engineering? (Papers I, IV, V & VIII). The specific objective is to explore and test different methods to identify safety performance indicators.

- What are the relevant safety performance indicators that could be applied to aviation and are they accountable to safety? (Paper VII) The specific objective is to explore which indicators may be identified for safety monitoring.

- How is a specific operation carried out in the real world and what implications for indicators and models are available to improve the monitoring of safety? (Paper VII). The specific objective is to explore which indicators may be identified for the monitoring of normal operations.

The sub-questions are discussed in this thesis through studies documented in the papers and this summarizing part.

1.3 Scope

Several methods have been developed in the areas of accident analysis, risk analysis and safety management. The scope of the work is to systematize and apply existing knowledge in the development of a method that could be applied to propose safety indicators addressing the main question. The work involves learning from specific normal operations, incidents and accidents. Environmental risks and security are not considered. Decision-making processes and mechanisms for the utilization of safety indicators are not considered in detail.

The function of the indicators is in terms of drift or sudden changes having a safety impact. The thesis concentrates on early detection and anticipation of expected and unexpected events. It attempts to identify both the strengths and weaknesses of specific operations through case studies. The work uses both qualitative and quantitative methods. The challenge is to combine methods rather than to unify them into a single method. Although the results are based on cases studies, generalizations and limitations are provided.

1.4 Structure of the thesis

The thesis consists of two parts as illustrated in Figure 1.1.

Figure 1.1 Thesis structure

The function of Part I is to group the papers in order to provide an overview of the contribution from this thesis. Chapter 1 considers the motivation for the thesis, its objectives and delimitations. Then, Chapter 2 turns to the description of the theoretical framework of reference. It includes recent developments in Resilience Engineering and indicators. Readers who wish to read detail information about indicators and its evolution are referred to Chapter 2; otherwise it is possible to read the last section of this chapter which summarizes the theoretical framework of reference. After this, Chapter 3 introduces the research approach, methods and criteria for evaluating the quality of the data. Chapter 4 describes briefly the application of methods by providing an overview of the papers. This is followed by Chapter 5 which discusses the contribution from the papers, critiques and comments given during the development of the studies. Finally, Chapter 6 presents the main conclusions and the recommendations for further research.

Part II consists of research papers that have been published in four international journals and four international conference proceedings. The contribution of the papers can be related to the ICAO Safety Management System (SMS, ICAO, 2009) approach adding the Resilience Engineering perspective. ICAO´s SMS proposes that navigation aids can be captured by reactive, proactive and predictive methods.

- Reactive methods relate to events with considerable damaging consequences. This monitoring uses the notion of waiting until "something breaks to fix it". Indicators related to accidents and serious incidents are examples of reactive navigation aids.
- Proactive methods consider less serious events with little or no damaging consequences. This monitoring uses the notion that system failures can be minimized before the system fails. Examples of indicators are related to mandatory and voluntary reporting systems, safety audits and safety surveys.
- Predictive methods are related to routine operational data captured in real time. They use the notion of trying to find trouble, not just waiting for it to occur. Therefore, indicators are related to potential problems, emerging safety risks from a variety of sources.

Note: ICAO SMSs strategies regarding safety improvements focus on avoiding that something goes wrong. Resilience Engineering addresses socio-technical systems ability and capability to adjust and to continue operations in presence of continuous disturbances. So, in this thesis predictive and proactive monitoring relate to the way the system adjusts to continue operation. These adjustments are always approximately and can result in failures or successes. Then, proactive methods focus on indication related to the current state while the predictive method focuses on the potential for future problems or opportunities.

Part II relates to ICAO SMS as follows:

- Concepts, theories and methods on indicators are presented on Papers I, II and III. These concepts are relevant proactive and predictive methods.
- Looking at the past: A reactive method based on Resilience Engineering perspective is presented in Paper IV.
- Looking at the present: Regarding proactive methods, an empirical study is discussed on Papers V and VI. These papers present an analysis of maintenance activities and organizational change for airlines and helicopter operators. Indicators are proposed in the papers.
- Identifying aspects that may become relevant in the future: The helicopter empirical study (HSS-3) includes observations of how the work is carried out in the real world from a Resilience Engineering perspective. The development and application of predictive methods accounts for everyday operation. The analysis of daily helicopter landing on helideck operations as a basis to propose indicators are presented on Papers VII and Paper VIII. Papers VII and VIII also present a combination of different methods and perspectives to propose indicators.

2 Theoretical framework

2.1 Introduction

This chapter summarizes the theoretical background of the thesis. This chapter complements and consolidates the limited information on the development of indicators provided in the papers. The thesis has been an iterative exercise through studies. Therefore, the theoretical framework presented in this chapter is not the starting point of the thesis; it rather provides an update on current developments including lessons learned from the studies. Section 2.2 presents the indicator concept, categories and characteristics. Section 2.3 illustrates the evolution on safety thinking and its implications for the selection of indicators. Section 2.4 describes different approaches for the identification of indicators. Section 2.5 introduces Resilience Engineering as a perspective for safety management. Section 2.6 describes recent studies on indicators inspired by Resilience Engineering. Section 2.7 discusses reviewed literature and the challenges addressed in the thesis.

2.2 The indicator concept

Indicator definitions, classification and characteristics have some assumptions. Indicators are often related to underlying models. These model(s) and assumptions have an impact on which factors are considered important for the monitoring of safety performance. Moreover, they also imply that some factors are left behind and not considered. Today´s socio-technical systems are complex, dynamic with many interrelations and dependencies. Therefore, it is necessary to be aware of which type of information is taken into account and also which information is left behind. The definition selected is influenced by the purpose of the indicators e.g. to allow management to be more proactive in terms of monitoring the status of the barriers or to identify emergent issues that may become relevant in the future. The characteristics are useful to operationalize indicators. These characteristics help to assess the quality and suitability of indicators. This section presents indicator definitions, classifications and characteristics to justify the choices I made in relation to support anticipation and action before something happens.

The safety science debate on performance indicators highlights the diversity of understanding and the genuine confusion in this area (Hale, 2009a; Hopkins, 2009b; Allford, 2009). Further work is needed to clarify the concept and its application[1]. The term indicator has been adopted and used in several ways by the safety community, which means that there are many definitions.

The etymology of the word indicator is ‘one who points out.’ It is related to Latin indicare, ‘to point out.’ Indicators are widely use in economic systems for forecasting and analysis. ICAO SMS (2009) defines safety indicators as “*parameters that characterize or typify the level of safety of a system,* the level of safety “*an emergent property which represents the quality of the system, safety wise*” and a safety indicator value as “*the quantification of a safety indicator”*. The document suggests that safety performance can be defined by a combination of quantitative and qualitative safety indicators. It also expresses that the objective should only be a quantitative measure. The problem with this objective is that it implies analytical sacrifices e.g. the number of inspections or number of recurrent training does not say anything about the quality of the training or the inspection. My opinion is that

---

[1] The main message from the papers and lessons learned from discussions with Erik Hollnagel, Luigi Macchi, Teemu Reiman and John Wreathall are summarized in this section. For more detailed information

a combination of quantitative and qualitative measures enables a deeper understanding of the situation analysed. This is based on experience and documented in Paper V.

Wreathall (2009) defines indicators as "*proxy measures for items identified as important in the underlying model(s) of safety*". I recommend this definition because it explicitly expresses the reasoning behind the selection of indicators. It illustrates how definitions and uses are often based on a specific understanding on safety. The relation between safety understanding, underlying models, methods and safety indicators is shown in Figure 2.1. Indicators are often based on theories, models and methods that describe, in the most accurate possible way, the reality of what may take place and what has taken place. The models and the method influence the kinds of data that are gathered and the way these data are analysed. It is also remarked that once an indicator set is proposed, this alone does not improve safety. The indicator is an input to the safety management and decision-making processes. Indicators also bring attention to specific issues and shadow other issues. A side effect from managing based on indicators is that there is too much effort on the improving the indicator value and too little attention on whether the measure actually contributes to improving safety in a sustaining way. Therefore, Figure 2.1 shows the need to revise the indicator set to assess whether the indicator set is still valid or will identify other areas where improvements are needed.

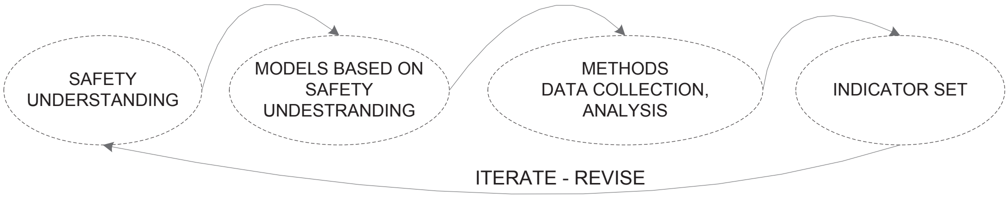


Figure 2.1 Safety understanding and indicators

The main purposes of safety performance indicators are: 1) to monitor the level of safety in a system (irrespective of whether this is a department, a site, or an industry), 2) to decide, where and how to take action, and 3) to motivate those in a position to take necessary action to actually do so.

Wreathall (2009) states "*Leading? Lagging? Whatever!*" and Hopkins (2009) argues that the distinction is not clear and might not be important. I will agree with the majority opinion in the safety science debate on indicators regarding the distinction as important. Leading indicators address the need to predict and act before an unwanted event (Hale, 2009b). Kjellén (2009) argues that the term leading indicators is borrowed from the field of economy without taking the full consequences of this change. The aviation industry has performed several improvements by using lagging information (EASA, Annual Safety Review, 2010). For this reason, the system is more vulnerable towards new or emergent types of disturbances. The identification of changes that might have an impact on future operations is essential. Thus, the need to have a balanced composition of different types of indicators and not just lagging indicators becomes important.

The literature gives different types of safety performance indicators (adapted from Reiman and Pietikäinen, 2010):

- lagging versus leading (Step-Change in Safety, 2001; HSE, 2006, EUROCONTROL, 2009)
- outcome versus activity based indicators (Kjellén, 2000; OECD,2005; ATSB, 2005)

- output versus input indicators (Van Steen, 1996)
- feedback, monitor and drive indicators (Reiman and Pietikäinen, 2010)
- process versus personnel indicators (Hopkins, 2008, 2009a)
- incident and barrier indicators (Vinnem, 2010)
- technical versus human factors indicators (DOE, 2005; Shappell and Wiegmann, 2001)
- lagging, current and leading indicators (applied in this thesis)

Regarding the classification, the first four categories are closely related. The definitions of lagging, outcome, output, incident and feedback indicators are similar. These indicators are related to the past performance of the system e.g. accidents, near misses, incidents. Activity and drive indicators are similar in that they measure efforts to prevent accidents. Monitor indicators and current indicators relate to the current state of the system. Lagging, current and leading are very similar to feedback, monitor and drive indicators.

The other types of indicators given above are classified from a different perspective than the leading/lagging notion. For instance when Hopkins differentiates between process and personnel indicators, it is to emphasize on the different failure mechanisms the indicator measures. It is crucial to understand that an indicator for occupational safety will most likely give no indication on process safety level and vice versa. Vinnem (2010) differentiates between incident and barrier indicators, which are both process safety indicators where one is measuring some kind of event and the other is an indicator reflecting a barrier's performance or condition. Both types may be both leading or lagging based on the definition used in the thesis. Finally, technical versus human factor indicators are evident where Shapell and Wiegmann (2001) separate indicators reflecting technical system condition and human and organizational influencing factors on safety.

Process indicators and barriers indicators are related to the status of safety barriers according to the defence-in-depth principle. In my view, process and barriers indicators are a good example of how the model influences the selection of indicators. In this case the model is based a barrier perspective where accidents are related to uncontrolled transfer of energy (Gibson, 1961; Haddon, 1970, 1973). The system is improved by placing barriers between the energy source and the object to be protected. In conclusion indicators from modelling based on this perspective will say something about the status of process (energy source) and the barriers.

The thesis focuses on early detection and anticipation of expected and unexpected events. In, for example, aviation there is extensive experience with the use of lagging indicators (e.g. accident rates, EASA, 2010), thus there is a need to recognize and use a balanced set of indicators addressing the past, present and future performance (reactive, proactive and predictive). For these purposes, it is reasonable to distinguish between three types of performance indicators: 1) lagging indicators, which refer to "after the fact" events that have happened or to system states in the past, 2) current indicators, which refer to what happens now or to system states in the present, and 3) leading indicators, which refer to what may happen or to possible system states in the future. Recent developments in aviation and other industries highlight the need to more active use of leading indicators. To align to aviation developments, early detection and anticipation, I choose to differentiate between leading and lagging. For future developments I would recommend a more granular use of lagging, current and leading indicators. This is because not all indicators related to present performance might be relevant for future performance. Leading indicators are relevant when the organization is moving towards its limits of operation focussing on changes that affect the ability of the system to continue safe operations.

The knowledge about the system that is gained from the performance indicators is used to decide what to do, how to do it, and when to do it. In practice all indicators have registered past or immediate past information since it is practically impossible to have measurements of future events. Lagging indicators may be data that were registered or collected in the past of after the fact events or conditions, and used to understand what has happened. These indicators can be used to improve performance after an unwanted event has happened and provide experience to improve the future functioning of the system. In most cases it is not sufficient to wait until an unwanted event has happened to improve performance. In some cases past performance is no longer relevant since the system or its operational environment has changed. Hence, it is useful to use current indicators related to the actual state of the system. These indicators show how the system actually operates under various constraints. Leading indicators are based on information pointing to possible future states. Often more indicators are needed to give a confident forecast. Interpretation and ability to see things in their proper perspective is consequently required to fully enable the data for decision support. Usually, leading indicators are not as easily acted upon, because the casual link can often only be established after the fact. Accident and incident analyses show that it is usually not single precursor events, but particular patterns of events that lead to negative safety outcomes (Grote, 2009). Gaining more knowledge about these patterns and building competence in companies and regulatory bodies for recognizing relevant patterns are crucial. Ale (2009) seems to support this argument by pointing out that many accidents have happened, not because of parameters being outside the design envelope, but the probability of coincidence is considered extremely rare. High Reliability Organizations are interested in single early events that provide weak signals of something being wrong (Hopkins, 2009b).

In this thesis the main purpose of the indicators is to support monitoring and more specifically anticipate and support action before something happens. I use indicator characteristics to operationalize indicators assessing their quality and suitability. In the literature and in the industry there many characteristics of safety performance indicators that can be used to assess indicator quality and suitability (Tarrants, 1980; OECD, 2005; TGRE, 2004; Kjellén, 2000; Wreathall, 2006, 2007; Webb, 2009). Yet few, if any, of the reviewed documents discuss the reasons behind the selection of the characteristics. A good indicator is meaningful; this characteristic is agreed upon in the Safety Science (2009) debate on indicators. Even though they must be meaningful it is important that indicators are easy to use. Unfortunately indicators are often selected because they are simple rather than inherently meaningful. In order to avoid this it is necessary to find a proper balance between easy to use and meaningfulness. There is not a single measure that will meet all the characteristics. I agree with Kjellén (2000) about a combination of measures can provide a reasonable compromise. Based on literature review, interdisciplinary discussion and lessons learned during this thesis work, the recommended characteristics of indicators are shown in Table 2.1.

Table 2.1 Criteria for safety indicators

Criterion	Definitions
Meaningful	Indicators are relevant to production and safety and can be used to address what is happening to the system in a specific context. Indicators provide information which guides future actions.
Sensitive	Indicators provide a clear indication of changes over a reasonable period of time.
Reliable	Indicators lead to the same interpretations when used by different people for the same situation. The interpretations are related to the system and its operational context.

Criterion	Definitions
Measurable	The values of indicators can be rendered in a concise manner, either quantitatively or qualitatively.
Verifiable	It is possible to confirm the correctness of the value or description of the indicators.
Inter-subjective	Indicators are understood in the same manner by different people, either from the same technical community or from society at large.
Operational	The indicators can be used to support concrete actions within the operational context.
Affordable	The cost of obtaining and using the measures is affordable vis-à-vis the benefits.

So in searching for safety performance indicators, we need to be aware of definitions, characteristics, models and their implications for the identification and selection of indicators. There is no agreement on the definition and use of indicators. Therefore, the purpose of the indicators should guide the choices. The purpose of the thesis is anticipation and early detection. In this case, as justified in this section, it is appropriate to adopt the differentiation between leading and lagging. The dimensions in characterizing a good indicator need to be related to the purpose of the indicators (Harms-Ringdahl, 2009). A set of relevant characteristics has been identified to assess the quality and suitability of indicators. In particular, sensitive characteristics are important since leading indicators should provide a clear indication of changes over a reasonable time. The characteristics proposed represent a pragmatic solution to operationalization of the indicators.

A method and its associated model guiding the selection of relevant indicators are needed. It should aim is to understand everyday performance and emergent patterns related to expected and unexpected outcomes (Wreathall, 2009; Woods, 2009). Consequently, it is necessary to look into the evolution of safety thinking and its influence for the development of models, methods and the selection of indicators. This theme is the subject of the next section.

2.3 Developments in safety management and indicators

The view of an accident influences developments in the understanding of safety. The way safety is understood has a parallel in the evolution of Safety Management Systems (SMS, Hale, Baram and Hovden, 1998; Hale and Hovden, 1998; Amalberti, 2001; ICAO, 2009). Consequently safety understanding and evolution in safety management strongly influence the selection and interpretation of indicators. The evolution is represented by “ages” in safety management. To summarize, Figure 2.2 shows some of the most influential accidents and developments in safety understanding. The accidents pinpoint important aspects to monitor and improve safety. They inspired different perspectives in safety thinking (Rosness et al., 2010). The “ages” in safety management continue their evolution are still relevant, and influence each other.

Figure 2.2 Accidents influencing evolution in safety management

In the first age of safety, the age of technology (Figure 2.2 starting from 1931), safety assessment concerns were mainly related to technological failures and their potential to trigger accidents. The domino model explains an accident as a combination of series occurring in a fixed order (Heinrich, 1931; Reason, 2008). Safety assessment methods were developed to search for causes of accidents by applying linear thinking about cause-relationships. A recognized technique inspired by linear thinking is the Fault Tree Analysis (FTA). This was a technique developed for the evaluation of the Minuteman Launch Control (circa. 1961), used for commercial aviation (circa 1966), adopted by nuclear power plant (circa 1971). Today, this technique is widely used by other industries. This technique has inspired other advance modelling techniques and associated indicators (discussed in more detailed in Section 2.5). In this age, human factor aspects were related to design and function allocation between humans and technical equipment in air traffic control (for example Miller, 1953; Fitts, 1951). In this period, the energy and barrier perspective accidents was proposed. It addresses unwanted energy release. Counter measure strategies are related to prevent, modify or mitigate the source of energy inspired by Haddon (1970, 1973). In relation to safety performance, the energy and barrier perspective highlights the importance of the monitoring of the quality and effectiveness of the safety barriers e.g. barrier indicators as proposed by Vinnem (2010). Today the barriers concept is extended from the purely technical to include human and organizational aspects.

The age of technology period was characterized by safety improvements related to the establishment of technical barriers to prevent accidents and technical failures. Safety performance indicators were identified from databases for incident and accident reporting systems (this approach still remains to a large extent in aviation). Data collection and indicators were related to technical failures to improve reliability and ensure regularity. Consequently, technological improvements led to a decline in accident frequency related to technical problems. Today, data collection and indicators are related to technical failures and real-time monitoring of the status of technical systems.

The second age, human factors which started in the 1970s, safety understanding and responses to major industrial accidents are attributed to human error. Figure 2.2 shows accidents such as Flixborough (1974), Three Mile Island, Bhopal (1979), Challenger (1986), Chernobyl (1986), the Herald of Free Enterprise (1987) and the King’s Cross Underground fire (1987) influence evolution in safety management in this period. Research is focused on human error using psychological and psycho-sociological knowledge on the occurrence, typology and mechanisms of human error (Swain and Guttmann, 1983; Reason, 1990; Maurino et al., 1995; Shappell and Wiegmann, 2000).

The overall safety of the system considers the patent errors of the front line actors (sharp-end) and latent failures generated by the design and organization (blunt-end, Swiss Cheese Metaphor, Reason, 1990; Reason, Hollnagel and Paries, 2006). The development of safety assessment methods evolved towards the adoption of multi-cause linear thinking, also expressed by epidemiological models (as in Reason’s Swiss Cheese model). Avoiding human error became the primary objective of Safety Management Systems and training was often the identified barrier. Practices such as Crew Resource Management (CRM) and Line Oriented Flight Training (LOFT) were introduced to identify, prevent and reduce the consequences of human errors.

Safety performance indicators are derived from taxonomies such as Human Factors Analysis and Classification Systems (HFACS, Shappell and Wiegmann, 2000; based on Reason’s Swiss cheese model mark I; Reason, Hollnagel and Paries, 2006). It is pointed out that the analysis of the information stored in databases is limited by the lack of information about the context in which the event has occurred (Gosling, 1998). Typically for aviation maintenance the major adverse influence of maintenance performance referred as "the dirty dozen" were: 1) lack of communication, 2) complacency, 3) lack of knowledge, 4) distraction, 5) lack of teamwork, 6) fatigue, 7) lack of resources, 8) pressure, 9) lack of assertiveness, 10) stress, 11) lack of awareness and 12) unsafe norms (CAA UK, 2003; Pantakar and Taylor, 2004). Safety performance indicators may be derived from this taxonomy. This rationale is found very often in aviation with a basic belief that solutions for future deviations use information on deviations, errors and violations.

Amalberti (2001a,b) argues that knowledge on human performance improved in the 1990s. He also argues that the error and incident strategy is inadequate for safety critical systems or ultra-safe systems such as aviation. It is essential to acknowledge that violations are symptoms of adaptations to cope with pressures and not the loss of control. Then a safety solution is not about suppressing violations but controlling them. Consequently, it is necessary to have means to identify safety indicators associated with the control of these adaptations. I have identified two major trends related to human factors in aviation. One trend is related to prevent "errors" so changes in design, operation and monitoring aim to eliminate or control factors associated to these errors. The second trend is to focus on the understanding of the adaptations needed to cope with pressures, disturbances and continue operations.

From the late 1980s (the age of organization and safety culture), the role of organizations and organizational culture is significant (as shown in Figure 2.2). A pioneer perspective in this thinking is the man-made disasters theory. This theory presents accidents as the result of a breakdown in the flow and interpretation of information leading to a disaster (Turner, 1978). It introduces the concept of "incubation period" where a *"chain of errors or several chains of errors developed unnoticed*" (Turner, 1978). Indicators based on this perspective are related to the ability of the organization to identify and follow-up signs of danger. Pigdeon and O'Leary (2000) suggest using (1) near-misses and use "what if" they become an accident, (2) consider worst cases scenarios, (3) identify an incubation period related to high ambiguity and uncertainty, and (4) ability to step outside and see the significance of hazards and their consequences.

The safety culture triggered by the Chernobyl accident in 1986 (International Atomic Energy Agency, IAEA, 1986), the Normal Accident Theory (NAT) triggered by the Three Mile Accident (Perrow, 1984) and High Reliability Organization (HRO) theory (LaPorte and Consolini, 1991) influenced the development to promote the safety culture and evaluate organizational factors (Reiman and Oedewald, 2009).

The theory of Normal Accidents Theory proposed "system accidents involve the unanticipated interaction of several latent and active failures in a complex system." (Perrow, 1984). Perrow (2007) proposes as preventive measure to shrink the targets and propose an example "the network of small firms". Monitoring strategies are related to interactive complexity and tightness of the coupling. In my view, indicators could be related to the control structure, technology and their interactions.

The HRO theory enhanced the study of organizations that successfully handle complex technology. This theory can be seen as a response to NAT. HRO studies organizations that handled successfully complex technologies. HROs reframe the term 'mindful' organizations (Weick and Sutcliffe, 2007). This type of organization considers continuous anticipation and containment. Anticipation is related to the ability to become aware of unexpected events through preoccupation with failures, reluctance to simplify, and sensitivity to operations. Containment involves resilience and the ability to deference to expertise. Deference to expertise relates to the migration of decision making to the levels where people come together to solve a problem. The monitoring of performance based in this perspective could be performed using questionnaires to assess the mindfulness of the organization. Surveys are developed to assess the safety culture of an organization. These surveys may collect quantitative information from questionnaires and qualitative data from interviews.

Among advances in the age of organization are those related to the development of management frameworks. For example in 1997, a framework based on Structured Analysis and Design Technique (SADT) was introduced to address the dynamics of safety management as a process (Hale, Heming, Carthey and Kirwan, 1997). On the one hand this framework links risks and direct preventive and control measures. On the other hand, it presents the organizational and management processes which support preventive measures. Its aim is to illustrate how organizational factors can be added or improved to ameliorate the performance of the SMS.

By the end of the 1990s, accidents started to be addressed as the outcome of normal system functioning rather than out-of-the-ordinary events (e.g. Drift into failure, Rasmussen, 1997; Normal deviations theory, Vaughan, 1996). Accidents arise due to the interaction between human, machines and the environment that cannot be explained by simple chains of effects and linearity. Decisions affecting safety management are made by different actors ranging from a political system to individual operators in different contexts and even different technical systems. There is a need to understand the consequences of the systematic migration of a socio-technical system towards augmented complexity and performance (Rasmussen, 1997). This “conflicting objectives” perspective includes models on migration towards the boundary of operation and distributed decision-making. The migration is also named drift into a failure. Vaughan (1996) use the term “normalisation of deviance” referring to the repeated handling of anomalies under production pressure leading to a change in culture. This culture changes the definition of anomalies into acceptable risk changing basic assumptions concerning what is normal and acceptable. Snook (2000) introduces the term practical drift. This term is related to a dimension of the pattern of tight and loose couplings and a dimension related to rule compliance versus local adaptations. Dekker (2011) argues about five concepts characterizing drift. These concepts are scarcity and competition, decrementalism or small steps, sensitive dependence to initial condition, unruly technology and contribution of the protective structure.

Control theory is proposed as means to detect and manage migration towards states of higher risks. Based on this theory, accidents are the result of inadequate control. Efforts in the organizations are related to local adaptations to optimize performance. The monitoring of performance could be based on feedback information about the significance of these adaptations across different levels in the organization from operations to management.

A response takes place in research in terms of methods to represent organizational factors and a useful overview of organizational and management frameworks is provided by Øien (2001). Other developments include Man Technology and Organization (MTO) concept focusing on humans and organizational factors and its relation to nuclear safety (Rollenhagen and Andersson, 1999). In the early 1990s, Managing Engineering and Safety Health (MESH) was an example of identification and assessment of local and organizational factors. Collectively, these metrics were designed to give an indication of the safety of the system (Reason, 1997; Reason and Hobbs, 2003). To a large extent, the organizational indicators in aviation are identified from accident or incident analysis.

In the beginning of 2000 the age of complexity[2], theories and models started to acknowledge, more or less explicitly, the constant occurrence of trade-offs, adaptations, adjustments and deviations from procedures. The same theories and models highlighted how non-prescribed behaviour might become acceptable and how it constitutes the normal context for successful, functioning of the organization until something goes wrong. Systems today present changes that challenge the established accident models, accident prevention and risk techniques calling for new paradigms (Dekker, 2005; Leveson, 2004; Woods, 2003; Rasmussen and Svenung, 2000). These changes are the fast pace of technological change, the change in management structures, the changing nature of accidents, new types of hazards, decreasing tolerance for single accidents, increasing complexity, integration and coupling of systems, additional complex relationships between people and automation, changing regulatory and public views of safety.

The age of complexity emphasizes that the functioning of the system cannot merely be explained by the aggregation of factors, but it has to be understood as an emergent phenomenon, where successes and failures are related, respectively, to the ability and inability to anticipate and recognize risks and critical situations, and to take appropriate actions (Hollnagel, 2004). Explaining operational success and normal work therefore becomes as relevant as explaining failures and accidents, but, as Dekker (2006) points out, Safety Management Systems chronically lack of theories and models to do so. Efforts have been used in the development of management tools like Line Operations Safety Audit (LOSA; Helmreich et al., 2003) and Normal Operations Safety Survey (NOSS) but, still, they are mainly used to manage errors and threats for flight operations and air traffic management, rather than to understand normal work.

An interesting development that is proposed in the nuclear industry regards a change of perspective to the organizational potential for safety. Reiman (2010) proposes a positive group of elements corresponding to the “dirty dozen”. They are: 1) clear communication, 2) self-criticism and reflection, 3) adequate task and safety knowledge, 4) good task and work design, 5) functioning teamwork and cooperation, 6) vigilance and energy, 7) flexible organization and slack resources, 8) social permission to carry work thoroughly, 9) assertive attitude to safety issues, 10) motivation and mental resources, 11) situation awareness and 12) norms supporting safety. A new perspective is added to the negative: accounting to measure the status of important elements. These kinds of indicators enable the monitoring and developing of a system safety (Reiman and Pietikäinen, 2010).

The Resilience Engineering (Woods, 2003; Hollnagel et al., 2006) perspective builds on previous understanding, perspectives and elaborates safety understanding related to today´s complex socio-technical systems. Here the focus is concerned with theories and tools to “*create foresight about the changing patterns of risk before failure and harm occurs*” (Woods, 2003; see Section 2.4 for further details). This perspective relates to the coping ability to handle expected and unexpected situations to continue operations. Indicators using the resilience perspective are related to this coping ability. A systems approach is advocated and themes such as influence of the context of operation, systemic, emergence, non-linearity and dynamics are relevant. Example of indicators in this case are markers of exceeding demands e.g. resources consumed to address a disturbance or indicators related to the ability to change to a new mode of operation (Woods and Branlat, 2011)

The developments in safety management provide a rich repertoire of perspectives and views. It is a great asset to safety practitioners to position themselves to bring out new

---

[2] Different names are proposed : Inter-organizational (Wilpert and Fahlbruch); Adaptive age (Borys et al., 2009), Age of complexity (Woods, 2010), Systemic (Herrera et al., 2010)

messages from divergent opinions that may respond more effectively (Rosness et al., 2004). Indicators based on different perspectives will direct attention to specific areas that need improvement (as shown in Table 2.2). In domino models leading indicators consist of single elements. When they fail, these may subsequently lead to catastrophic failures. Indicators in the Swiss cheese model monitor the performance of safety barriers. The resilience perspective proposes a systemic model looking into the dynamics of safety. Systemic models emphasize that the functioning of the system cannot merely be explained by the aggregation of factors, but also have to be understood as an emergent phenomenon. Indicators are needed to monitor the variability of normal performance (Hollnagel, 2004) or the adaptive capacity of the system (Woods, 2011) calling for the development of methods that, in accordance with the resilience engineering principles, look at safety in other ways than solely on hindsight and error tabulation.

Table 2.2 Ages of safety management and safety performance indicators

Ages of safety management	**Safety understanding**	**Indicators**
Technology	From monitoring technical failures to monitoring the continuous health monitoring	Indicators related to technical failures and real-time technical status. Pilots or technical reports related to component failures, maintenance findings and in-flight engine shutdown (Kinnison, 2004).
Human factors	From human errors to human performance	Indicators related to human errors poor judgement, fatigue, situation unawareness (Maurino, et al., 1995)
Organization and safety culture	Factors related to human, technology and organization	Indicators related to organizational failures training-competence, planning-coordination, maintenance programme, design (Øien, 2001), safety management systems audits, safety culture evaluation, organizational culture studies (Reiman and Oedewald, 2010)
Complexity, systemic models	Safety cannot be reduced to independent factors = dynamic on-going non event	Indicators considering the system as a whole, looking into its variability positive and negative. Rather than a single indicator, this requires the interpretation of a set of indicators related to specific context. Questionnaires related to resilience abilities, a set of indicators represented by the Resilience Analysis Grid (RAG, Hollnagel, 2011).

2.4 Resilience engineering perspective for safety management

A dictionary definition of resilience is “the ability that a person or institution has to recover quickly from a setback or misfortune”, or “the quality that something has of being strong and not damaged easily, for example by being hit, stretched, or squeezed”.

In safety literature, resilience is considered in terms of the capacity of a system or organization as a whole to simply “bounce back” (Wildavsky, 1988). He characterizes resilience as the “capacity to cope with unanticipated dangers after they have become manifest” (Wildavsky (1988). A distinction between passive and active resilience has also been proposed. Passive resilience defined as “the mere ability to bounce back without breaking”. Active resilience defined as “a deliberate effort to become better able to cope with surprise” (Lovins and Lovins, 1982 quoted in Wildavsky, 1988). Foster (1993) defined resilience as “an ability to accommodate change without a catastrophic failure”, or” the ability to absorb shock gracefully”. Resilience has also been defined as “the properties of an organization to make it more resistant to its operational hazards” (Reason and Hobbs, 2003). Rosness et al. (2004) adapted a similar definition of resilience as “the capacity of an organization to accommodate failures and disturbances without producing serious accidents”.

It also seems that for some researchers the core of resilience is to be prepared for surprises. Weick and Sutclife (2001, 2007) describe five types of surprises. The first is the one that nobody was expecting, there was not a hint that it was coming. The second is recognized but the expectation is in the wrong direction. The third and fourth surprises are expected, but the timing is wrong. This means, either the surprises came too early, too late or had an unexpected duration. Finally, the fifth surprise is expected but the amplitude was not foreseen.

In addition, from the Resilience Engineering literature, there are new definitions that also implicitly revise the premises and gradually include new aspects, in their attempt to capture what it takes to accomplish resilience:

- Hale and Heijer (2006) suggest that resilience concerns ”the characteristic of managing the organization's activities to anticipate and circumvent threats to its existence and primary goals”

- Leveson et al. (2006) understand resilience as the capability of a system to prevent or adapt to changing conditions in order to preserve its control over a system property.

- Hollnagel (2011) proposes resilience as “the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions”.

Resilience Engineering searches for ways to enhance the ability of the organizations to monitor and revise risk models, to create processes that are robust and flexible, and to use resources proactively in face of disruptions or ongoing production and economic pressure (Woods et al., 2010). It aims to develop theories, methods and tools to proactively manage the ability of organizations to function effectively and safely.

The literature identifies four essential capabilities in a resilient system (Hollnagel, 2009):

- Learning from experience requires actual events, not only data in databases. This requires selecting what to learn and how the learning is reflected in the organization, i.e. what is reflected in changes in procedures and practices. This ability is related to coping with the *factual*.
- Responding to regular and irregular threats in a robust and flexible manner, corresponding to the reactive part of safety management. The system is designed to provide a limited range of responses. There is still a necessity to adjust responses in a flexible way to unexpected demands. This ability enables coping with the *actual*.
- Monitoring in a flexible way means that the system's own performance and external conditions focus on what it is essential to the operation. This includes internal monitoring as well as monitoring the external conditions that may affect the operation. This will make it possible to identify what could be *critical* in the near future.
- Anticipate threats and opportunities. It is required to go beyond risk analysis and have the requisite imagination to see what may happen, and see key aspects of the future (Westrum, 1993). It is not only about identifying single events, but how parts may interact and affect each other. This ability addresses how to deal with the irregular events, possibly even unexpected events thereby allowing the organization to cope with the *potential*.

Safety management is regarded as a control problem where accidents are the result of interactions that violate safety constraints on design or operation (Rasmussen, 1997; Rasmussen and Svenung, 2000; Leveson, 2004). Economy, workload and practice are example of pressures that influence system performance as illustrated in Figure 2.3. One possibility is that the system moves towards unacceptable boundaries. This vulnerability can be manifested by an accident, an incident or an interruption of operations (e.g. Alaska accident, NTSB, 2002). Defences may be eroded when facing production pressure and the system drifts towards failure as in the Columbia shuttle accident (Woods, 2003). Dekker (2004) argues that the process of erosion of drift towards the margins cannot be captured by static models. Thus it is necessary to also address the dynamics of the system and understand how the system copes with these influences. Control theory may be combined with the idea of the safety space using a combination of proactive and reactive navigational aids (Maurino et al., 1995). Indicators can be regarded as navigational aids for safety management.

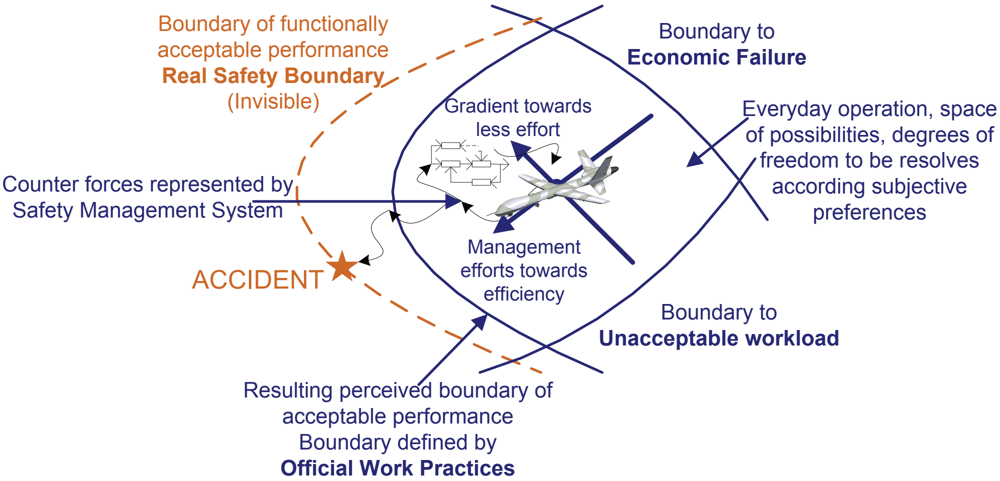


Figure 2.3 Migration towards the boundaries of acceptable performance
(Adapted from Rasmussen, 1997)

Hollnagel (2008a) contrasts Resilience Engineering with (traditional) Safety Management Systems (SMS) by using a control theory (Figure 2.4). According to this, an SMS is both:

- Reactive in the sense in which improvements are based on correction of failures.
- Proactive in the sense of closed-loop feedback control. A reaction is triggered by the difference between the actual state and the target state. However, the difference can arise because of unanticipated internal and external disturbances. Therefore, the response may not be appropriate.
- Predictive in the sense of feed-forward control. Here the focus is on leading the system from an actual state to a desired future state. The model of the process is developed describing the process and the possible disturbances from the environment. The reaction is designed to react to anticipated disturbances, rather than actual ones. As all models are approximations, it is easy to fail in relation to disturbances that are not foreseen. Hence, the limitation in this approach is that the response will not take place if disturbances have not been taken into account in the model. Feed-forward control may also result in unnecessary responses.

However, a successful, proactive SMS may experience a fundamental regulator paradox (Weinberg & Weinberg, 1979; Van Steen, 1996) of cybernetic theory (Ashby, 1956, 1981;

Beer, 1985). That is, if the number of events to correct against drops significantly because of the absence of "errors" and incidents, the process may be uncontrollable in the face of a sudden disturbance. The lack of information may be misinterpreted to mean that the process is under control, while reality is the opposite. Resilience Engineering does not see safety as an absence of accidents but as a capability of the system to adjust and cope with current conditions. Based on a control theory perspective, it makes more sense to use a definition of safety in relation to production such that the output (the amount of essential information) increases when safety improves. The "target" of the control loop should not be to avoid or get away from something, but rather to achieve or get closer to something.

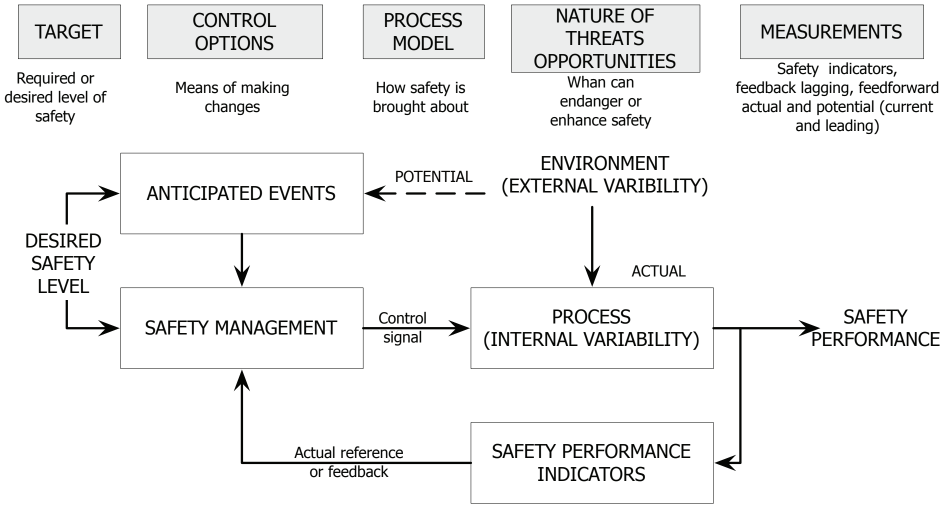


Figure 2.4 Safety Management as feedback and feed forward control
(Adapted from Hollnagel, 2008a)

Resilience Engineering is about increasing the ability of the organization to make correct adjustments. Within the paradigm of the control loop, Resilience Engineering pursues its target by a combination of feedback and feed-forward control, in a scheme where:

- The controller is the "SMS" including the means of intervention.
- The process model is the description of how safety is produced, and what is required to detect significant changes to be able to select and support appropriate adjustments.
- Disturbances can represent both threats and opportunities, and may originate from internal or external variability.
- The "SMS" should be able to handle the regular threats as well as the opportunities

The output is related to safety and productivity performance that is monitored by performance indicators. In addition to lagging indicators, current and leading indicators are needed that provide information about what the state may be in the future. Current indicators are also indicators of the immediate past than of the present, although the delay may be so small that it can be disregarded, unlike the delay associated with lagging indicators.

Table 2.3 Using indicators in controlling a socio-technical system

	Lagging indicators	**Current indicators**	**Leading indicators**
Targeting – defining targets for the system, what should be achieved.	Targets can be defined from an analysis of past performance.	Current indicators can be derived from leading indicators (as targets for	Leading indicators can be derived from targets and company objectives.

	Lagging indicators	Current indicators	Leading indicators
		performance)	
Monitoring – keeping an eye on how a system performs, its general "state of health", use of resources, etc.	Monitoring uses lagging indicators to adjust functioning, following disturbances.	Monitoring uses current indicators to adjust performance during operation.	Monitoring uses leading indicators to anticipate developments, potential problems or opportunities.
Regulating – the actual control of the system to meet permanent or temporary targets, including the moment-to-moment control of individual functions.	Lagging indicators provide feedback for tracking.	Short-term regulating makes use of current indicators.	Leading indicators provide feed-forward information for response. Actions required to anticipate changes.
Examples	Information from aggregated data to illustrate trends	Production rates, available resources	Slack of resources to cope with situations. Limited time available More resources required to cope to specific situation Weak signals when production pressures. For technical systems, non-destructive test are used to monitor the actual state of structures to plan further actions

The "control theory interpretation" of Resilience Engineering hence advocates an understanding of normal, productive behaviour. It expands the focus of analysis of proactive Safety Management Systems including data about both what goes wrong and what goes right (Hollnagel, 2009). The former has to be used to avoid a re-occurrence of similar events, whereas the latter must check the "vital signs" of the organization and identify areas for continuous improvement of the core business process (Reason, 2008). In relation to indicators, Resilience Engineering aims to provide a better understating of the functioning of organizations to improve their ability to anticipate adverse and beneficial conditions and to act effectively. In order to be effective, proactive Safety Management Systems need to do the following (Reason, 2008; Weick, 2001, 2009; Dekker, 2005; Reiman and Oedewald, 2009):

- See safety as a dynamic process and emergent phenomenon;
- Help organizations in balancing production pressure and protection needs;
- Recognize the combined contribution to safety provided by technical systems, people and organizations;
- Be sensitive to the creation of opportunities and not only the presence of deficiencies;
- Be focused on organizational processes and the influence of the context of operations and inter-organizational aspects.

Figure 2.3 shows the necessity to expand the focus of the analysis to show information that is otherwise obscure or critical, we must look for challenges and how the socio-technical system copes with constant trade-offs. For example, in the context of a nuclear power plant local adjustments and rearrangements of rules, including sometimes rule violations are necessary for the organization to achieve its goal (Bourrier, 1996; Perin, 2005). McDonald (2006) and Pettersen (2008) provide similar examples from the aviation industry. Another issue is that it is necessary to have indicators addressing past, current and future performance. The aim of monitoring is to capture the interactions and adaptations that make the system work. As a consequence, it provides better knowledge of everyday performance to identify future problems and opportunities. Bourrier (2002) supports this argument by stating that "*the study of normal operation helps us improve our level of understanding of complex organisations, because it focuses on the duality of organisational life: the dark side and the bright side, always tightly coupled*". Therefore, shifting the level of abstraction to include the analysis of "what goes right" compensates for the drawbacks of relying solely on negative outcomes and linear cause-effect thinking. The analysis of normal operations when nothing goes wrong triggers the analysis of the unintended consequences of adaptation.

Thus Resilience Engineering addresses complex socio-technical systems. In aviation, these systems are complex, subject to many interdependencies, operational and economic constraints. Systemic methods inspired by Resilience Engineering address safety as emergent phenomena. Some methods, such as the Functional Resonance Analysis Method (FRAM, Hollnagel, 2004), have been proposed for modelling socio-technical system functioning which describes the variability of everyday performance. The resulting FRAM model helps to identify indicators whether they are current, leading or lagging (Paper VII). Another study proposes the identification of indicators related to the activities that take place, as well as the abilities, skills and the organizational potential for safety (Reiman and Pietikäinen, 2010). Further work on these cases and the use of indicators is required to refine the monitoring of the variability of everyday performance. The Functional Resonance Analysis Method has been explored in Papers IV and VII.

2.5 Models, questionnaires and indicators

The identification of indicators is often based on models or questionnaires. There is a strong tradition of modelling safety management systems. Models and methods have been developed to take into account organizational, human as well as technical aspects. Questionnaires use key concepts of a safety perspective. This section briefly discusses some selected methods and modelling approaches.

The aviation, chemical, nuclear and railway industries present examples of methods and models. The Integrated Risk (I-RISK) approach aims to produce a probability of major hazard occurrence weighted by human and organizational factors. This approach integrates the safety management system into the quantification of risk for chemical industries. I-Risk uses SADT for modelling the management of the system (Bellamy et al. 1999, Papazoglou et al., 2003). Accidental Risk Assessment Methodology for Industries (ARAMIS, Andersen et al., 2004) aims to develop a risk assessment methodology to evaluate the risk level considering prevention tools implemented by the operators. It integrates structural aspects of management and aspects of safety culture. A variety of tools are used to measure the structural aspects and culture. The tools include questionnaires; interviews and auditing that are integrated in the modelling. The objective of the Causal Model for Air Transport Safety as its name indicates is to produce a fully operational causal model that represents causes for air transport accidents and safeguards in place (CATS, Ale et al., 2006; 2008). Advance modelling techniques uses a combination of 33 generic accident scenarios represented by an Event Sequence Diagram (ESD) combined with a Fault Tree for each event, including the human performance represented by Bayesian Belief Network (BBN). It considers factors that influence error probability e.g. flight crew error probability. Other main initiatives are proposed to the industry: These modelling initiatives include the Federal Aviation Authority Hybrid Causal Logic (HCI) model, the EUROCONTROL Integrated Risk Picture (IRP) and the Norwegian Risk

Influence Model (RIF) applied to helicopter operations on the Norwegian Continental Shelf (SINTEF, 1990, 1999, 2010). Traffic Organization and Perturbation AnalyZer (TOPAZ) use agent-based Dynamic Risk Modelling (DRM) combining Petri-Nets and Monte Carlo simulation for collision risk assessment. TOPAZ is under continuous developments and recent updates include safety culture assessment. It is argued that this approach provides a better understanding of problems and organizational behaviour (Stroeve et al., 2011). Socio Technical Risk Analysis (SoTeRiA) aims to extend Probabilistic Risk Assessment modelling to include the effects of organizational factors. The method combines FT, ESD, BBN and System Dynamics (SD), the models are integrated using HCI. The modelling produces total system risk, failure probabilities and status on risk influencing factors (Mohaghegh et al., 2009, 2010). Examples from the Norwegian oil and gas industry include the Organizational Influence Model (ORIM) (Øien, 2001), Barrier and Operational Risk Analysis (BORA) (Aven et al., 2005; Sklet et al., 2006) and Operational Conditional Safety (OTS, Vinem et al., 2007, 2008). The oil and gas industry modelling examples propose organizational factors that have an influence on major risk. Earlier modelling from the railway sector includes management and organizational influence factors arising from human error (Model of Accident Causation, MACHINE, Embrey, 1992), An example from the nuclear industry is the modelling work processes and incorporating organizational factors into risk assessment (Work Process Analysis Model, WPAM, Davoudian et al., 1994a; Davoudian et al., 1994b). One of main contributions from these analyses is that they aim to monitor risk addressing different aspects of operation or organization.

The accident model called STAMP (Systems-Theoretic Accident Modelling and Process) was proposed by Nancy Leveson (2002). The model is based on systems control theory. The motivation to the development of the model was the need to improve the analysis of systems containing software and complex human decision making; the organizational and managerial aspects of systems, and the adaptation of systems over time (migration towards hazardous states). The STAMP hazard analysis (STPA) method has as its main purposes (1) the identification of the system hazards and the safety constraints necessary to ensure acceptable risk, and (2) accumulation of information about how these constraints could be violated, to be used for eliminating, reducing, and controlling hazards in the system design and operations. The outcome STPA analysis is information for hardware, software, and human components of the system to:

- guide the test and verification procedures (or training for humans)
- change the overall system design to provide protection against the error
- add fault tolerant features to the component itself to protect against the identified hazardous
- guide a fault tolerant design process

Some comparisons between STPA and traditional techniques such as FTA and hazard and operability study (HAZOP) have been performed (Leveson, 2011). The comparison documented that STPA allowed identification of new factors that were not identified by traditional techniques. The analysis supported the identification of risk resulting from the integration of different system elements. The analysis allowed the identification of recommendations and modification of the systems analysed. Traditional techniques were developed for simple electro-mechanic system. The application of new methods such as STPA demonstrated the need of new methods and theories for complex, human and software intensive systems. More developments are expected from STPA.

Most of the risk analyses are triggered when a deviation occurs and take into account failure probabilities, while a systemic approach based on resilience engineering, is used for an understanding of variability of a performance that is not considered in the above risk analysis it does not necessarily represent a failure in the system (Herrera et al., 2010).

Regarding the questionnaires, Section 2.2 mentions Weick and Sutcliffe (2007) to assess the mindfulness of an organization. There are many questionnaires developed to evaluate performance. Paper V discusses the utilization of a specific questionnaire. Hollnagel (2010) proposes the Resilience Analysis Grid (RAG) as a basis to develop an organization profile in relation to the resilience capabilities. For each capability a set of questions is proposed. Both approaches provide a snapshot of the organization in relation to specific resilience characteristics. Therefore, it is recommended to repeat the utilization of the tools periodically.

The Structured Analysis and Design Technique (SADT) is proposed as an alternative to capture the dynamics and flows that take place in a socio-technical system. This analysis is a functional modelling approach like FRAM. Therefore, it is interesting to reflect and compare SADT and FRAM. I identify the following differences and similarities:

- Purpose of the analysis: The SADT framework provides a statement about how the SMS should be structured and work. It is proposed that SADT could be used for accident analysis tracing to particular failures of the SMS, training, SMS assessment and SMS design. FRAM can be used for retrospective analysis e.g. analysis of incidents, risk analysis or analysis of everyday operation.
- Structure: One difference is that SADT has hierarchical levels such as system structure (S), planning, organization and procedures (P) and executing (E). So the top influences what happens at the bottom. In FRAM functions have no levels. The functions can be of types such as organizational, human or technical.
- The three levels: knowledge-based, rule based, and skill- based proposed by Rasmussen can be related to the different levels in the SADT model. Decision making corresponds to the SADT levels. In FRAM decisions are translated into operational or organizational function.
- Links: SADT fixed links and feedback loops from one level to the level above. The FRAM model does not have fixed links. The links in FRAM are materialized at different times and are explicit through instantiations of the model.
- Model: The SADT model consists of the activities linked while the FRAM model consists of the functions without links.
- Activities and functions: An activity in SADT is equivalent to a FRAM function with some differences. The activity has four aspects: inputs (1) which are transformed or used to produce outputs (2); the use of resources (3); and this activity is under certain controls (4). A FRAM function is characterized by six basic aspects: Input (I, that the function uses or transforms), Output (O, that the function produces), Preconditions (P, conditions that must be fulfilled to perform a function), Resources (R, that the function needs or consumes), Time (T, that affects time availability), and Control (C, that supervises or adjusts the function).
- Results from analysis: The analysis based on SADT provides improvement measure in relation to SMS improvements. A FRAM analysis provides recommendations related to managing the performance variability. These recommendations could be related to barriers for unwanted variability, recommendations that support desired variability or propose indicators to monitor variability.

The FRAM analysis is based on Resilience Engineering perspective which is the perspective selected for the thesis. This method is under development and the thesis could contribute to this process. SADT is seen as an audit tool to structure safety management systems. It is argued that FRAM could support the identification of indicators (Hollnagel, 2004). The scope of the thesis is to advance knowledge in the development of a method that could be applied to monitor safety performance. Rather than providing a snapshot, it is necessary to capture the dynamics of the system. Among the reasons for selecting FRAM are its suitability for modelling dynamic scenarios, its use of operational experience, it covers safety as well as operational aspects, it takes into account technical, human and organizational performance and it considers dependencies. Finally, it looks at the system as an open system influenced by its context of operation.

2.6 Resilience engineering perspective and indicators

Since Resilience Engineering is a relatively new area in safety management, no established way to think about resilience has been agreed. Rather, we are faced with a variety of understandings. These understandings vary from only focusing on human performance, organizations or a systems approach. Resilience Engineering accounts for the context of operation and advantages that interdependencies represent to continue operations. The point of interest is a systems view addressing complex socio-technical systems, interactions and interdependencies between humans, organizations and technology. There is no an established and mature way to determine indicators. Nevertheless, there are concepts and principles and examples about the development of approaches for the identification and use of indicators. This section reviews selected studies relevant to indicators using the Resilience Engineering perspective. Studies addressing only human performance, organization and/or technology as independent areas are not considered. Resilience Engineering borrows ideas from different domains such as ecology, sociology, materials properties, neuroscience and control theory. These areas address similar challenges and present interesting alternatives. Examples of how these other areas relate to indicators and resilience engineering are discussed.

Resilience Engineering addresses the ability to recognize and adapt to unanticipated disturbances (Woods, 2006). Then, the understanding on how the system anticipates and adapts in reality to different kind of disturbances is relevant to monitor and manage resilience. Woods argues that "*it is possible to measure the potential for resilience than resilience per se*". Another point is that the perspective influences the focus on some aspects and at the same time obscures other aspects. Shifting between different perspectives is needed to have a broad repertoire of possible solutions.

Woods relates resilience to adaptive capacity. In his view, improving safety is sound adaptive capacity. Systems designs are based on different models and these models have limitations. Systems operate under a design envelope, when the design envelope is challenged due to conditions that were not considered, the system is pushed to its limits of operation. It is my understanding that if system adapts and copes with these surprises the system is resilient, on the other hand if the system cannot adapt the system is brittle. Due to the fact that adaptive capacity is finite, there is a variety of ways to analyse the system and see how it reveals sources of resilience and brittleness. Hence, indicators can be related to the measurement of adaptive capacity. These measurements aim to identify areas where interventions are needed to improve the adaptive capacity. The adaptive capacity relates to what the system adapts to (or not) to and how the adaptation is realized.

Inspiring by ideas from ecology, it is possible to observe that changes in adaptive capacity may originate changes in adaptive cycles from equilibrium, emergence, growth, and maturity. In relation to indicators is necessary to identify changes and tipping points such as transitions in the adaptive cycles and see how indicators changed over time to assess if there are risks or opportunities.

Table 2.4 presents an overview regarding the properties of resilient systems and indicators found in the literature. The first columns list key properties of resilient systems (Woods, 2006). The other two columns present some indicators related to these characteristics. Anders et al. (2006) use these properties as a means to investigate resilience and classes of adaptive challenge in healthcare and Mendoça (2008) proposes candidates for indicators in telecommunications.

Table 2.4 Resilience engineering properties as basis for indicators

Properties of resilient systems Woods(2006)		**Resilience in action (Anders et al. 2006)**	**Indicators as candidate measures for factors contributing to resilience (Mendoça 2008)**
Properties	**Description**		
Buffering capacity	It relates to the size or kind of disruptions that the system can absorb and adapt maintaining safety and effective production without a fundamental failure.	The notion of buffering changes as the scenario evolves. In the example an indicator may be related to the capacity of the trauma unit in relation to patient needs.	No example is identified
Flexibility	It addresses the ability of the system to restructure in response to external changes and pressures.	Reconfiguration of the system. Utilization of resources from other units.	Development of new procedures. Recognition of unplanned-for-contingencies.
Margin	It indicates how closely the system operates to its boundary of performance. It should be noticed that the boundary of performance is not fixed but also varies as the system and context of operation vary.	Deployment of resources in terms of distance to the margin, availability and timeliness.	Resource utilization, network load, network stability
Tolerance	It relates to the way of the system operates closely to its boundary. The system could degrades gracefully, or collapse.	Sacrifice decisions and use additional resources to maintain control.	There is a challenge to elaborate descriptions of organizational behaviour at process level. Indicators may be identified doing pre and post event comparison of communication and decision-making process at the individual, group and organizational levels.
Cross-scale interactions	It relates to the influence of the context to local adaptations, and how local adaptation has an impact on more global, strategic goals.	Coordination with different units.	No example is available

For his studies, Mendoça uses a combination of data sources such as interviews, questionnaires, after action reports, meeting notes and drawings. These studies suggest triangulation of observation using both quantitative and qualitative methods. He concludes that there is need for more comprehensive range of observation techniques and analytic methods are needed to provide information on how to engineer resilience in power and telecommunications infrastructures. Anders et al. (2006) use the five resilient properties as a framework for types of adaptive capacity. This study illustrates how a balance of these properties can be perceived and adapted in advance, hence allowing the management of the adaptive capacity in phase of a potential collapse.

Another way to characterize and measure the resilience of a system is to use the stress-strain analogy (Woods, Wreathall and Anders, 2006; Woods and Wreathall, 2008). It borrows ideas from world of engineering showing the relationship between different stresses and how the structure stretches in response. In the analogy a stress-strain plot illustrates a resilient systems adapting to different kinds of demands. The plot shows different regions. The uniform region where the system stretches uniformly to the increasing loaded. The adaptive capacity in this region relates to plans, procedures and practices designed into the system. When changes exceed the adaptations and plans built into the system the system enter to a non-uniform region. In the non-uniform region active steps are essential to maintain safe and effective production. If the demands continue to increase and the system is unable to respond, it either fails or needs restructure to continue absorbing stresses. The point is that an organization can modify its adaptive capacity expanding the range of responses. So, it is necessary to manage transitions between regions and expand responses to adapt. A critical point is to recognize transition between regions. Hence, indicators in this plot are gap-filling adaptations and incidents. These indicators are markers of exceeding demands. They uncover the gap-filling adaptations or adaptations that are not built into the system. Such indicators relate to the type of disruptions or demands that challenge the uniform region. Resilient systems are able to recognize the need to shift as they have mechanisms available to provide the adaptation needed for a specific situation while brittle systems move quickly into the non-uniform region of adaptive capacity toward the failure point. A well-calibrated organization uses indicators to see where there are adaptive shortfalls and demonstrate the ability to enhance this particular aspect of adaptive capacity.

Resilience also refers to how well the system manages transitions between regions. Woods and Wreathall (2008) point out some limits in this analogy. First, demands are mapped onto a single dimension while there are different kinds of demands that can affect the system in different ways. Second, the analogy does not address the design and how it is possible to set-up or to modify a system. This plot relates to the adaptive capacity of the system and the level of calibration rather than the anticipation of future threats and opportunities. Further work using this plot is needed to demonstrate and show the management resources that are needed to make the system resilient in the transition between regions that require different kind of adaptations. Lay (2011) explores this plot on a case study from maintenance of power plants as illustrated in Table 2.5. By looking at general situations and selected unexpected situations, one way to become more resilient consists of noticing potential indicators of risk profile changes. Lay argues that noticing trigger actions helps to reduce loss.

Table 2.5 Indicators and actions using the stress-strain analogy
(Lay, 2011)

Candidate indicators	**Possible solutions to act based on candidate indicators information and interpretation**
Examples of indicators related to the transition to other region: • Multiple issues taking crew attention • Schedule impacts, multiple delays, • Speciality personnel on site longer than anticipated • Sudden need for more people • Higher than usual amount of emergent work • Decline in communication	Experienced personnel design a menu of actions to mitigate changes in the risk profile such as: • Stop and assess the situation • Better organization of the site • Communicate up the chain of commend • Develop resources to provide buffering

Adaptation requires anticipation. Anticipation relates to noticing and accounting for signs where the adaptive capacity degrades, accounting for a potential effect of changes and taking advantage of opportunities. Woods (2011) describes patters in how resilient systems may anticipate that the adaptive capacity is falling. Table 2.6 summarizes patterns related studies and possible indicators identified in the literature. The modelling of the dynamics associated to how the adaptive capacity of a system aims to capture general properties that can be used to understand how specific systems will behave when it encounters signs that the adaptive capacity is falling in relation to the challenges ahead (Alderson and Doyle, 2010). Ideas are inspired from adaptive systems and views of organizations to cope more effectively with the tragedy of commons (Ostrom, E., 1999). This tragedy symbolizes the expected degradation of the environment when many individuals use a scarce resource. Such a common-pool resource is defined as a natural or man-made resource from which it is difficult to exclude or limit users once the resource is provided and once a person's consumption of the resource units makes those units unavailable to others (Ostrom et al., 1994). Based on these ideas, work has started in the development of polycentric control architectures that enable the dynamic management of relationships across diverse and interdependent roles, organizations process and activities (Hofman, 2004; Hofman and Woods, 2011). It is my view that indicators can be related to the following five trade-offs focusing on performance and adaptive capacity to events and changes in the future. Focus and optimization with respect to specific criteria guarantees an increase in brittleness with respect to changes and variation that fell outside these criteria (Doyle 2000). Table 2.6 summarizes the five trade-offs for resilience engineering as basis to develop polycentric architectures. This table also describes the implications for indicators.

Table 2.6 Fundamental trade-offs and implications for indicators

Fundamental trade-off that bound performance of complex adaptive systems (Hofman and Woods, 2011)		**Patterns of anticipation**	**Implications for indicators**
Type	**Description**		
Bounded ecology, optimality - resilience	Gaps in fitness since an adaptive system can never completely adapt to its environment.	Ability to recognize that the adaptive capacity is failing to allow the system to bounce back or degrade gracefully. Of particular interest are potential cascading effects, new connections and interdependencies	Indicators should be related to the fitness of the system with respect to its environment of operations. Noticing when the system needs to work harder to maintain control. Examples of indicators are falling behind the tempo of operations or the inability to change to a new mode of functioning when anomalies or contingencies occur. (Branlat and Woods, 2010)
Bounded rationality, Efficiency-thoroughness	Gaps in procedures, plans, models and finite resources.	Ability to recognize the threat of exhaustive buffers or reserves	Indicators should provide information about when the margin of manoeuvre is expanding or contracting relative to the potential for surprise. This aspect addresses the possibility for adaptations of future demands by providing sufficient reserves. For indicators, critical questions are: should resources be consumed to address a growing disturbance or be built and sustained to constitute critical reserves? What are the signs or indicators suggesting a shift between these two strategies?
Bounded perspicuity,	Gaps arise in perceiving the world from any given	Ability to shift and contrast diverse	Multi-method or modelling supporting shifting and contrast of

Fundamental trade-off that bound performance of complex adaptive systems (Hofman and Woods, 2011)		**Patterns of anticipation**	**Implications for indicators**
Type	**Description**		
Acute-chronic	perspective. Perceiving the world from one perspective determines what it is possible to see. The ability to shift perspectives to reveal what is hidden.	perspectives that go beyond their nominal system position.	perspectives. Thus indicators related to different perspectives should be identified and assessed.
Bounded responsibility Specialist-generalist	Gaps can arise across roles as different parts of a distributed system are differentially responsible for different subsets of goals.	Ability to navigate interdependencies across roles, activities and levels. Ability to recognize when to shift priorities across goal tradeoffs	Challenge is indicators that can identify when goals pursued by different parts of the system might conflict and how progress in one of the systems introduces higher demands on other parts. An essential indicator is related to how the organization manages situations where goals conflict and still coordinate activities.
Bounded effectivity, Distributed-concentrated	Gaps in the balance between global plans and local adaptations to meet a goal within a specific context.	Ability to recognize to shift between concentrated or distributed action.	Polycentric Control Architectures for managing interdependencies at scale. Indicators related to scope and scale of the work are relevant to the identification of potential action.

Currently, in the Norwegian oil and gas industry, there is a project designed to establish a set of indicators as early warnings for the prevention of accidents (Resilience Early Warning Indicator, REWI; Øien et al., 2010). The REWI method is based on the Leading Indicators of Organizational Health (LIOH, EPRI, Wreathall, 2000, 2001) method developed for the nuclear industry and consists of different levels. The LIOH themes are also proposed as relevant to measure organizational resilience (Woods and Wreathall, 2003). Table 2.7 shows the relation between LIOH and REWI developments. It is noted that the REWI model is very similar to a Risk Influence Model replacing the risk factors with contributing success factors. In my view, these two developments are focused on the ability to cope with problems, identify risk/hazards and respond to these risks/hazards. These aspects are very important to preparedness and response. RE also takes into account the understanding of everyday successful operations, adaptations and the ability to continue operations. These aspects are not described in detail in the REWI paper. The quantitative indicators proposed by the REWI method need to be supported by qualitative data, the amount of courses does not say anything about the quality of the training and the suitability of the courses to the equipment. Further developments are required addressing the dynamic environment of operations, emergence and non-linearity.

Table 2.7 Leading Indicators of Organizational Health and Resilience Early Warning Indicators

Leading Indicators of Organizational Health (LIOH; Wreathall, 1999)		**Resilience Early Warning Indicators (Storseth et al., 2009; Øien et al., 2010)**	
Themes	Themes are common or recurring terms identified in organizational models (organizational culture, system breakdown, safety quality and reliability): 1. Management commitment 2. Awareness of safety performance 3. Preparedness for problems	Resilience Attributes	Top level
		Level1: Contributing Success Factors (CSF)	Level 1: • 1. Risk awareness to avoid underestimation of risk • 2. Response capacity to a deviation or incident • 3. Support decisions in case of goal-conflicts to maintain critical

Leading Indicators of Organizational Health (LIOH; Wreathall, 1999)		**Resilience Early Warning Indicators (Storseth et al., 2009; Øien et al., 2010)**	
	4. Flexibility built in for responding to problems 5. Just culture (to promote reporting of errors and failures) 6. Learning culture (to promote fixing of problems) 7. Transparency (visibility of safety performance)		functions given a deviation or incident
		Level 2 CSF	Can be seen as influencing factors having a impact on Level 1 as follows: • 1.2.1 Risk underestimation, 1.2.2 attention, 1.2.3 Response • 2.2.1 Response, 2.2.2 Robustness, 2.2.3 Resourcefulness/rapidity • 3.2.1 Decision support, 3.2.2 Redundancy
General Issues	Theoretical or operational subcomponents of themes. Sets of general issues are proposed e.g. for top management commitment a general issue: • “Management seem to value human performance”	General Issues Level 3	A set of general issues related to each CSF level 2 e.g. for anticipation: • Risk/hazard identification • Learn from own experience, accidents • Learn from other’s experience, accidents
NNP Issues	For the above general issue “Human performance matters are important to senior utility management”	Level 4: Indicators candidates	A set of indicators is proposed for each general issue from level 3 e.g for risk-hazard identification: • Portion of operating personnel taking risk courses last 12 months • Portion of staff taking risk courses last 12 months • Portion of operating personnel informed about risk analyses last 3 months
Leading Indicators	Analytical indicator: Line managers are rewarded for tackling human-performance problem	Not applicable	Not applicable

Many of the developments address the capability to cope with problems, adaptations when the system is stretched towards boundary conditions. One aspect that requires further development is the modelling, analysis and identification of indicators related to everyday operations. These cases are interesting because they are markers of the ability of the organization to continue operations. Everyday operations require adaptations and it is needed to see the safety significance of these adaptations.

2.7 Conclusions

The thesis aims is to identify safety performance indicators associated with drift and supporting anticipation in complex socio-technical systems represented by aviation cases. This aim has implications for the theoretical choices. The way safety is understood strongly influences the selection and interpretation of safety indicators. The evolution of safety understanding influences the selection of indicators. This safety understanding allows the development of today’s complex socio-technical systems. As a consequence we need to analyse performance with tools for today´s type of systems, their characteristics and operational context. I agree with Amalberti´s argument that safety in aviation has a mature approach to perform improvements after failures, at the same time there is a need to understand the numerous adaptive process that makes the system work. Moreover, there is a lack of good theory or methods of the functionality of the organizations (Amalbert, 2001a; ACARE, 2010; McDonald et al., 2012). Resilience Engineering views safety as something that the system *does* and not something that the system *has*. One of the main contributions from the thesis should relate to understanding everyday operation and indicators.

Regarding current concepts, theories and principles, the term indicator and its purpose should be explicitly defined each time that it is used. There is disagreement in the safety community on the usefulness of differentiation between leading and lagging. Recent accident investigations recommendations highlight the need for the active use of leading and lagging indicators (Texas accident, 2005; Deepwater Horizon, 2011; Air France, 2011). I consider the differentiation between lagging and leading important. A more granular differentiation such as leading, current and lagging indicators is proposed. These differentiations encourage the search and use of a combination of indicators that shows how the system actually behaves in the present and potential future states. The challenge in the thesis is to identify factors that are critical or essential to continue operations.

Resilience Engineering as a safety perspective for complex socio-technical systems represents choices that makes the monitoring of performance a challenging issue (adapted from Hoffman, 2004):

- Rather than a static or snapshot of performance, it is necessary to account for dynamics, the effects of changes and continuous monitoring.
- The methods should address interactions and interdependencies within and across organizations, so research needs to go across different organizational boundaries
- There are multiple process that happen simultaneously, the choice to study some specific areas will leave other areas “obscure”
- Multiple representations and triangulation are needed because socio-technical systems are multidimensional and each method, data and representation are biased
- Rather than attempt to understand the components, a systems approach is pursued to understand system interactions. The relationships are non-linear requiring analysis on a specific context. The analysis of the system as a “whole” aims to analyse how variability of performance in some functions can be amplified or damped by other functions.

In line with the choices listed above, the Functional Resonance Analysis Method is selected as representative of RE to identify indicators. The FRAM method produces a large amount of candidates for indicators. So, I use the characteristics of a good indicator as criteria for assessing the quality and suitability of the identified indicators. ICAO proposes that the main goal is to have quantitative indicators. I disagree with this approach; the indicators are not absolute measurements and can be either quantitative or qualitative. Quantitative information should be complemented with qualitative data to provide a broader understanding of its significance Depending on when they are sampled in the different case studies; the indicators can be lagging or leading indicators, respectively. The thesis will explore and operationalize FRAM in specific case studies to identify indicators related to the variability in everyday performance.

To support production and improve safety, the analysis of everyday performance, incidents, and accidents has two functions. First, they identify factors and conditions that influence successful performance. Second, they can enable proactive and predictive safety strategies to be developed that can be integrated into production targets. Safety indicators deduced from the analysis of everyday performance have the potential to be critical leading indicators of the quality and opportunities of an organization. The interpretation of this type of indicator must support the anticipation of new paths of performance. Leading indicators are closely related to the target, thus these indicators provide information that drives performance improvements. Studies are foreseen to see how the analysis of everyday performance could provide a basis to propose leading indicators.

The methods and modelling studies presented in this section showed that a variety of perspectives provide complementary understanding than is possible from a single perspective. Thus it is necessary to study and apply different perspectives and methods to assess the system and to identify indicators. The thesis explores a combination of established and RE- based methods. Solutions should be developed and proposed with active participation from industry to enable practical implementation and more operational proactive safety management.

Resilience relates to the adaptive capacity of the system to respond to changes. Thus, it implies that improvements on the adaptive capacity and the abilities of the system will improve safety. Other approaches can be considered relevant such as assessing performance in relation to the fundamental trade-offs. The monitoring of safety can be extended to evaluate performance in relation to the five trade-offs (gaps in fitness, procedures, perspectives, across roles, balance between local adaptations and global plans as mentioned in Table 2.1) to ensure the system's capacity to manoeuvre in the trade-off spaces. This approach is particularly interesting because it assesses the system when the system reaches the boundaries of operation and trade-off that are necessary to cope with situations.

3 Approach and methods

3.1 Research approach

Considering the main research question: *How do we identify that a system drifts or experiences sudden changes in the safety space?* calls for a strategy which is both theoretical and empirical. This requires applied research as it is an original investigation to acquire new knowledge directed towards a specific practical aim or objective (OECD, 2007, 2002). The thesis is problem driven and not methodologically driven in the way that it uses the methods that will provide the best help to answer the research question (Flyvbjerg, 2006).

The case study approach is the selected strategy that attempts to combine qualitative and quantitative data (Eisenhardt, 1989, Yin, 2009). Yin (2009) reworks the definition of case studies and proposes a two-fold definition which is applied in the thesis:

> " *A case study is an empirical inquiry that: investigates a contemporary phenomenon in depth with its real-life context, especially when the boundaries between phenomenon and context are not clearly evident. The case study inquiry copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result relies on multiple sources of evidence, with data needing to coverage in a triangulation fashion, and as another result benefits from the prior development of theoretical propositions to guide data collection and analysis."*

The thesis and its papers have been based on the following steps:

- The PhD project plan identified the questions and the studies required to answer the research question.
- Theoretical studies and exploration of relevant research literature were done to systematize existing concepts and developments for the identification of safety performance indicators.
- Empirical industry studies were conducted to explore different research methods for the identification of safety performance indicators.
- The empirical studies are an iterative process with strong interaction with industry. Results from the application of the methods were presented to the participants to verify facts and to produce the final inputs to the findings

Considering that the study is based on the Resilience Engineering perspective, the research explores established and relatively new methods for the identification of indicators. The indicators address the ability to anticipate threats and opportunities. The choice of methods was based on the objectives of the sub-questions, the resources and competences available. An important aspect is the paradigm that each safety method represents. The utilization of a list of activity and outcome indicators is a response to the available knowledge and interaction with the aviation industry. Resilience Engineering advocates the use of systemic methods. In order to learn about multi-linear methods and systemic methods, the Sequentially Timed and Event Plotting (STEP) method and the Functional Resonance Analysis Method (FRAM) were explored. The STEP method is widely used to investigate serious aviation incidents and accidents by the Accident Investigation Board, Norway. By 2004, two systemic methods had been introduced. The Systems-Theoretic Accident Model and Processes (STAMP) (Leveson, 2004) and the Functional Resonance Analysis Method (FRAM) (Hollnagel, 2004). The reason for selecting FRAM is that relatively little work had been performed in the development and application of the method. The modelling is based on RE principles. It is suitable for dynamic scenarios and takes into account human, technical and organizations. Today, it seems that the word "systemic" is popular and many methods claim to be systemic. Systemic methods like FRAM are based on a systemic model. A systemic model emphasizes that the functioning of the system cannot merely be explained by the aggregation of factors, but also has to be understood as an emergent phenomenon. An emergent phenomenon arises out of complex dynamics, relationships and interactions that cannot be attributed to simple causal explanations (Reason, 2008, Hitchins, 2004)

A set of indicators can be derived using models and associated methods. They provide the reasoning behind the selection of indicators and the type of indicators proposed to monitor safety. The model directs the analyst towards what to look for, the way in which data is collected, indicators are identified and interpreted. These models are used to structure the identified data. The studies also present the utilization of a questionnaire (Paper V) and the utilization of the risk influence model (Papers VII and VIII). I did not participate in the elaboration, application and analysis of the questionnaire. I had a more limited role in the application of the modelling of the Risk Influencing Factors. Therefore, these methods are not discussed at the same level of detail as the triangulation of data, STEP, FRAM and storytelling.

Sequentially Timed Events Plotting (STEP)

The STEP method was developed by Hendrik and Benner (1987). The most important basis for STEP is that neither an accident nor its investigation is a single linear chain or sequence of events. In STEP an accident is viewed as a process that has an unexpected outcome. In this view, the processes are dynamic and interact, thus an accident is not a single event but a group of dynamic actions. The objective of accident investigation is the understanding of the events and the interaction between actors and actions. A multi-linear event sequence is implemented in a STEP worksheet. The rows are labelled with the names of the actors on the left side. The columns are labelled with marks across a time line. The worksheet is the basis for the investigation and to keep the events organized and clear. One of the main purposes of the description of the accident process is to identify problems from which lessons can be learned to improve safety. The safety problems are identified by analysing the worksheet to find event sets that constitute the safety problem. The identified safety problems are marked as triangles on the worksheet. These problems are evaluated in terms of severity. They are then assessed as candidates for recommendations.

Functional Resonance Analysis Method (FRAM)

FRAM was introduced by Hollnagel (2004). FRAM can be used for both accident and risk analysis. FRAM is under continuous development. In its present form, FRAM comprises the following steps:

- Step 0 - Define the purpose of the analysis, since FRAM can be used for both accident investigation and safety assessment.

- Step 1 - These functions are necessary (and sufficient) for the intended (correct) performance to be produced (when "things go right"). The result of the second step is the model of the system. Every function can be characterized by six basic aspects: Input (I, that which the function uses or transforms), Output (O, that which the function produces), Preconditions (P, conditions that must be fulfilled to perform a function), Resources (R, that which the function needs or consumes), Time (T, that which affects time availability), and Control (C, that which supervises or adjusts the function). Recently, it is proposed to classify the functions identified by a FRAM safety assessment as either foreground or background functions. Foreground functions represent the focus of the analysis and background functions represent the context in which foreground functions are performed. (Macchi, 2010). The identification of functions is an iterative process between the analyst and the personnel involved in the scenario. Several iterations are necessary to achieve some degree of completeness. The stop rule and completeness are verified by checking that all aspects described for one function are defined in another function.

- Step 2 - Assess and evaluate the potential variability of each function. Everyday performance variability is mapped through observations and discussions with operative and management personnel.
- Step 3 - Identify functional resonance by means of instantiations. An instantiation illustrates aspects and the potential links among the functions in a defined context for specific time intervals (Paper IV). The aim of this step is to determine the possible ways in which the variability from one function could spread in the system and how it may combine with the variability of other functions. This may result in situations where the system loses its capability to safely manage variability. The propagation may be both indirect via the effects that the variability may have on the general conditions or direct via the output from a function.
- Step 4 - Identify effective countermeasures or barriers that can be introduced in the system. In FRAM, prospective countermeasures aim at dampening performance variability in order to maintain the system in a safe state. It is consistent with the principle of Resilience Engineering to also consider measures that can sustain or amplify functional resonance that lead to desired or improved outcomes. Besides recommendations for countermeasures or barriers, FRAM can also be used to specify recommendations for the monitoring of performance and variability in order to be able to detect undesired variability at an early stage. Performance indicators may thus be developed for individual functions and for the couplings among functions (This development is explored and documented in Paper VII).

Risk Influence Modelling

This method is discussed in the papers and my co-authors were responsible for its application and results. I had limited involvement in this work. The risk influence model has been used to illustrate and communicate the influences on risk, to provide estimates of the change in risk level for a certain period and to predict risk improvement potential by implementing risk reducing measures (Hokstad et al., 2001). The model is based on a number of Risk Influencing Factors (RIFs) arranged in influence diagrams. A RIF is defined as a set of conditions that influence risk, either positively or negatively. RIFs are split into two categories: risk frequency influencing factors and risk consequence influencing factors. These categories are organized into three levels. Operational RIFs (Level 1) are risk influencing factors related to activities that directly influence risk and are necessary to provide safe helicopter operations on a day-to-day basis. Organizational RIFs (Level 2) are defined as risk influencing factors related to the organizational basis, support and control of running activities in the transport of helicopters. Regulatory and customer-related RIFs (Level 3) are defined as risk influencing factors related to requirements and control activities from international organizations, authorities and customers. First, a model is produced to indicate the influence of those RIFs which contribute to a specific accident/incident category. Second, the quantitative model is updated to perform predictions, while estimates of parameters are obtained from an analysis of accidents and incidents supplemented by expert judgements (personnel from helicopter operators, air traffic services and authorities). The combination of statistics and expert judgement enable assessment to be made of the relative contribution to risk from the identified RIF. The expert judgement sessions were carefully planned and performed under controlled conditions following a consistent framework. Several expert judgements were performed aiming at some degree of convergence. This knowledge is used to determine areas in which risk reducing measures are effective.

3.2 Data collection

Multiple approaches have been used for data collection and analysis. The theoretical studies are the result of literature study combined with discussions with other researchers to clarify concepts, methods, models and theories. The empirical studies included gathering company management documentation, interviews, group interviews, workshops and observations in natural and simulated environments. All empirical studies are the result of valuable cooperation between airlines, the helicopter industry, the petroleum industry and the regulators interacting with researchers. This section describes the methods for data gathering in the different studies.

Selection of interviewees, scenarios and topics

Key participants in the industry were interviewed to map technical, operational and organizational aspects. The selection of interviewees was based on consultation in the project team, inputs from industry and my industry knowledge. A strategic group of key participants in industry were interviewed to map technical, operational and organizational developments that had an impact on airlines maintenance activities (PaperV) or on helicopter operations offshore (e.g. pilots, engineers, managers, air traffic controllers, inspectors, national and international experts).

The scenarios were selected in close cooperation with the industry. The incident case was provided by AIBN (Paper IV). For the helicopter papers, RIF modelling started prior to the FRAM modelling. While performing the RIF part of the helicopter study, operations related to helicopter landing and take-off from helideck were considered critical operations. The RIF allowed the identification of potential scenarios. Then, the FRAM helicopter scenario was selected in cooperation with operational personnel.

In relation to topics for storytelling, the first part of the study using RIF and RE modelling and interactions with the industry led to the identification and agreement of four central topics related to changes in internal framework conditions. These topics were:

- Change of decision-making authority/management of resources and its significance for work practices;
- Changes and their significance for maintenance routines;
- Changes in competence and training;
- Changes in cooperation and communication with local management.

Individual interviews

The aim of an interview is to enter another person's perspective (Patton, 2002). The theme to be addressed determined the type of interview. The subject of organizational changes was more delicate to address. These changes in some cases had a direct impact on interviewee's work and feelings. In these cases individual interviews were considered appropriate. In other cases specific subjects needed in-depth discussions and, individual interviews were accomplished. For the interviews, an interview guide was prepared. On some occasions the interviews were recorded. The interviewers took notes and prepared minutes. The notes were subjected to quality checks.

Group interviews and workshops

A limited group of people with specific backgrounds met to provide answers and discuss specific topics. At least two researchers participated in the group interviews. Prior to the interview, the interviewees were informed about the purpose of the interview and the topics to be discussed. If relevant, consent was requested for the interview. The interviewers took notes and consolidated their notes. The notes were subjected to quality check.

Observations
The aim of the observations is to perceive the conditions, constraints and requirements under which operative personnel work (Reiman and Oedewald, 2009). In this case, it was particularly beneficial to have had previous industry knowledge to understand the operations. The observations from these studies are complemented with semi-structured interviews. Notes were taken and the models were complemented with the knowledge gained.

Storytelling
The storytelling perspective permeates a large part of organizational studies. It provides valuable insights into the nature of organizations, the power relations within them and the experiences of their members. Stories reveal how wider organizational issues are viewed, commented upon and worked on by their members. Stories are defined as "*narratives with plots and characters, generating emotion in narrator and audience, through a poetic elaboration of symbolic material. This material may be a product of fantasy or experience, including an experience from earlier narratives. Story plots entail conflicts, predicaments, trials, coincidences, and crises that call for choices, decisions, actions, interactions, whose actual outcomes are often at odds with the characters intentions and purposes*" (Gabriel, 2000). Prior to the interviews for storytelling information about the interview was provided. Moreover, consent was requested from the interviewees. Interviews for storytelling were designed based on a list of specific topics (listed in Section 3.2), as a dialogue in which the interviewees have the opportunity to offer their own stories on these central topics. The interviews were recorded when the interviewees provided their permission. Minutes of the meeting were prepared and submitted to the interviewees to verify the factual correctness.

3.3 Data analysis

Data triangulation and models are applied to analyse data. The analytical approach is a combination of methods. This section presents the application of methods in the different studies.

Data triangulation
Triangulation of data sources and methods increases the accuracy and credibility of findings (Patton, 2002). Paper V presents a combination of analysis of company data, interviews, STEP method, safety indicators and a questionnaire. The selection of indicators is based on Tinmannsvik (2005). This initial list of indicators was adapted to helicopter operations in close cooperation between SINTEF and Scandpower selecting those indicators where information could be gathered from the companies studied. The safety indicators were divided in two categories, which are outcome indicators (reactive) and activity indicators (proactive) (Kjellén, 2000; Papers V and VI). The collected information was analysed by a multidisciplinary team to achieve firm conclusions.

Aviation safety and maintenance under major organizational changes, investigating non-existing accidents (Paper V)
Analysis of documentation, interviews, a list of safety related indicators and a questionnaire were applied for data collection. The information was analysed on several

occasions by a multidisciplinary team in order to achieve as firm conclusions as possible. The organizational changes were mapped through individual and group interviews. The interviewees were selected as being representative for key areas such training, flight operation helicopter and fixed wing, regulators and operators. Prior to the interview, information about the interview was sent to the interviewees and interview guide was prepared. At least two interviewers participated in the interview and took notes. Minutes of meeting were prepared and a STEP diagram was produced to map organizational changes and their influences. After a preliminary version of STEP was prepared, it was presented and discussed in the interviews. Refinements to the model were achieved. Regarding the list of indicators an extensive list with data was requested from the industry. It was found that in some information could not be compared across organizations since the information relevant to one indicator e.g. maintenance costs contained different data. Then, indicators that could be compared across organizations were presented to the AIBN e.g. Minimum Equipment List, Back-log.

Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis (Paper IV)

The Accident Investigation Board, Norway (AIBN) suggested that the authors looked into Norwegian Air Shuttle accident probably due to the fact that contributing factors were not obvious and could not be easily established. The first method for data collection was textual analysis of the AIBN's report complemented by authors' knowledge. I performed most of the STEP analysis. The FRAM analysis and the comparison were performed jointly by both co-authors of this paper (Woltjer, 2009).

A first STEP diagram was prepared based on the incident report and was checked with two experienced accident investigators at SINTEF. The FRAM analysis started with a description of functions, a first model was developed using the incident report and the FRAM Visualizer (FRAM visualizer is a software developed to illustrate FRAM model by the LiU in 2007). The authors focused on the comparison to identify needs. Gaps in information and questions were collected from STEP and FRAM first versions to design interviews. The analysis required an iterative process between the researchers, pilots and air traffic controllers. The models were updated with information gathered through four structured interviews, visit to the Oslo Gardermoen control tower including half a day of landing observations and semi-structured interviews. A workshop with AIBN personnel was organized to discuss the comparison with the participation of aviation, railway, road and maritime investigators. The final version of the model and findings were submitted to operative personnel pilots and air traffic controllers to verify correctness. The comparison was also discussed during the 2nd FRAM workshop with participation of researchers from different industries including aviation. At the end of each meeting, notes were produced and analysed. Consequently, the model and comparison were updated and refined. The data collection and analysis involved a total of about 50 people including air traffic controllers, pilots and accident investigators. The final results are documented in the paper.

Proposing safety performance indicators for helicopter offshore on the Norwegian Continental Shelf (Paper VII)

The first approach was to document a review of indicator literature, statistics provided by the helicopter operators, oil and gas producers and national authorities (Norway and UK) and the Accident Investigation Board, Norway. An internal report was prepared to document this work. The RIF modelling and the literature provided the basis for the selection of a safety critical operation. Landing on helicopter deck is interesting due to the complexity of the operation. On one hand, many incidents involve helicopter deck operations. On the other hand, successful landings occur every day. The operation requires skilled personnel, equipment accuracy and availability, good interaction between pilots and helicopter deck personnel under a harsh environment, time and regularity constraints. Therefore, helicopter landing on helicopter deck was selected as an important case for the analysis of a successful operation. To start the FRAM analysis it is necessary to understand the landing process. So, a first iteration was prepared to identify helicopter landing functions. Individual interviews were designed and conducted with air traffic control, pilot and helicopter deck personnel. The interviews were recorded and verified. The model was prepared and presented to the interviewees in several occasions to verify the correctness and completeness of the model. The identification of indicators was an iterative process using interviews with administrative, technical, operative, helicopter deck personnel. Observations of helicopter landing on a helicopter deck during simulator session and semi-structured interviews were conducted. The observations and discussion with pilots helped to improve modelling and improved the analysts understanding of the context of operations. A preliminary list of indicators was identified and discussed during a workshop and followed by individual interviews with pilots, engineers, training personnel, helicopter deck, air traffic controllers, petroleum representatives and regulators. This was done to assess the indicators against the indicator criteria. Results from interviews were recorded. A final list of indicators has been proposed to the aviation industry.

Approaches to elaborate on the safety of offshore helicopter operations (Paper VIII)
Interviews, group interviews and expert judgement sessions were performed. An interview guide was prepared to map technical, operational, and organizational changes for the period 1999-2008. A total of 12 individual and group interviews were effectuated. The interviews had duration of 3 hours approximately. Minutes of the meeting were prepared and submitted to the interviewees to verify the factual correctness.

In relation to storytelling, as explained in Section 3.2 representatives from management, the operative and technical departments of three helicopter companies participated in the interviews. A total of nine individual interviews were performed each having a duration of two hours in average. The transcripts of the interviews were used in the analysis and to prepare the report. A draft of the report documenting the storytelling was sent to management, operative and technical representatives to verify facts correctness. At the end of the project, recording of the interviews was erased in accordance with the Norwegian regulations for data storage. The storytelling method was used for the analysis of the data collected. In-depth analysis was not planned because of time and resource constraints. The data were collected through individual and group interviews. The material was analysed by a multidisciplinary team. The stories constitute different images of an organization; they are multifaceted and can be interpreted in more than one way. The analysis work is time consuming because it requires interpretation the material several times. Stories related to the topics and main findings were used to illustrate findings.

Regarding risk influence modelling, there were six expert judgement sessions with a total of 35 participants. Each session had duration of two days approximately. The expert sessions focused on different aspects required for the risk influence modelling e.g. technical, organizational, preparedness. A strategic group of pilots, maintenance personnel, managers, inspectors, helicopter deck personnel, air traffic controllers search and rescue interacted together with researchers. The size of the group varied in relation to the addressed themes and to achieve different opinions. Results from the sessions were recorded and distributed to the participants to verify factual correctness. Risk influence modelling was the quantitative method use for the data collected.

Preliminary results from the analyses were discussed with different people in industry and in the research community to assess factual correctness and to test the results. The final findings were documented in the project report (Herrera et al., 2010) and Paper VIII.

3.4 Quality of data

The evaluation of the quality of data uses reliability, validity and generalization as valid criteria for quantitative research whereas credibility, conformability and transferability are valid criteria for qualitative research (Ringdal, 2001; Thagaard, 2003). The case-studies in the thesis represent a combination of quantitative and qualitative methods. These methods are used in a complementary way. Therefore, this section presents the evaluation of data against criteria for both quantitative and qualitative studies.

Reliability and credibility

Reliability is the degree to which methods can be repeated with same results (Yin, 2009). Credibility relates to the rigorous methods, the credibility of the researcher and the appreciation of the qualitative methods as a purposeful sampling (Patton, 2002).

Regarding reliability, it is possible to achieve the same type of results and retest findings of the quantitative methods using statistics, incident and company data (Papers V, VI, VII and VIII). Raw data is stored in the project database. This data was accessed through confidentiality agreements. Therefore, it cannot be disclosed. Nevertheless, the data is stored in regulators' and operators' databases. Once the same quantitative data is gathered the method can be tested. The quantification in the risk influence model includes expert judgements. Several expert judgement sessions were needed to achieve solid results. For the qualitative studies and expert judgements, the context of the interviews changed and data cannot be replicated.

For credibility, the process and methods for data collection and analysis for each study are documented rigorously and described in the papers. All interviews presented in the studies were prepared in advance. For example, a strategy was selected by preparing detailed interview guides (Paper V) or topic guide (Paper VIII) prior to the interview. The data collection has been documented and verified. Moreover, each case study involves several researchers representing different disciplines to collect and analyse to achieve consistent results. These studies are iterative processes between known people in industry and the researchers. The combination of different disciplines and the iterative process enhances the credibility of the studies. It is highlighted that the quality of the information obtained during an interview depends on the interviewers (Patton, 2002). During the interviews, the interviewers tried to avoid influencing the interviewees. Still, the collected information is a result of the data provided during the interview and the interviewers' understanding. In general at least two researchers with different backgrounds participated and transcribed the interviews. In most of the cases, the transcriptions were prepared and consulted with the interviewees to verify that the collected data was correct and to reduce the biases of interviewer's interpretation. The utilization of quantitative and qualitative methods in a complementary way was appreciated. Since, the qualitative methods shed light on aspects that the explored quantitative methods do not address (for example Papers VII, VIII). The qualitative methods also support some of the findings of the quantitative methods. For example, RIF modelling and storytelling enable the identification of the important influence from designer organizations, helicopter operators and maintenance organizations (Paper VIII).

Validity and coverage

External validity refers to the degree to which the findings can be generalized (discussed in the generalization and transferability paragraph). Internal validity seeks to establish a causal relationship while content validity refers to the extent to which the method or the selected indicators provide adequate coverage of the problem studied (Ringdal, 2001; Yin, 2009).

One analytical tactic to address internal validity is using logic models (Yin, 2009). All case studies use different logic models to analyse the data. The final models are the result of several iterations. A first version of the models was prepared and was revised during interviews or seminars. The process was iterated as needed or when a certain level of convergence was achieved. Rather than achieving a complete result after the first iteration, the model is subjected to examination and validation through several iterations. The utilization of models and iterations support the establishment of reasonable validity.

Regarding the coverage of the problem studied, information was collected from several sources. For example, Paper VII covers information from the regulator, the accident investigation board, the helicopter operators, helicopter deck personnel and air navigation services providers. The information was extensive and the study had adequate resources to gather this data. Data triangulation was applied combining different data (quantitative and qualitative) and using different methods (observations and interviews). The combination of methods provides enough evidence for the validity of the research (Silverman, 2006).

Generalization and transferability

Generalization refers to the extent to which research findings from a population can be applied to the population at large. It is normally achieved through statistical sampling. The aim of this sampling is to achieve confidence in the representativeness of the sample. This facilitates inferences to be made about the whole population (Silverman, 2006). Guba and Lincoln (1981) propose the concepts of transferability and fittingness instead of generalization for qualitative studies. These terms are related to the degree of similarity between two contexts. Then, the hypotheses from the original context may be applicable to the receiving context (Lincoln and Guba, 1985; Patton, 2002)

The data are not based on statistical sampling. The findings are influenced by the context of operation. Therefore, they are representative for the companies being part of the studies. Nevertheless, the context may be relevant to other industrial sectors. For example, Paper V concerns concurrent organizational changes. The interviewees were affected by these changes. Therefore, the changes had an impact on the data gathered. It is possible to apply transferability of the applied methods for data gathering and analysis. It is also possible to identify patterns in the operation that are relevant to other operators and lessons that can be learned. For example, Paper V shows the possibility to identify important patterns regarding the balance between outsourcing maintenance activities and keeping the necessary know-how in the company. Paper IV deals with the comparison of two methods. The study reveals that it is possible to use different methods to achieve broader understanding. Hence, the methods may be applied to analyse other events.

4 Summary of papers

ICAO SMS proposes that navigation aids can be captured by reactive, proactive and predictive methods (ICAO, 2009). It is also possible to associate these methods to a feed forward control management strategy. The Resilience Engineering perspective adds the ability to succeed under varying conditions. This perspective includes the understanding of real-time operations and what is essential to continue operations. An overview of papers and their contribution to ICAO's methods is shown in Figure 4.1. The abstract of the appended papers including a brief comment on relevance to their scientific community and industry are presented in this section.

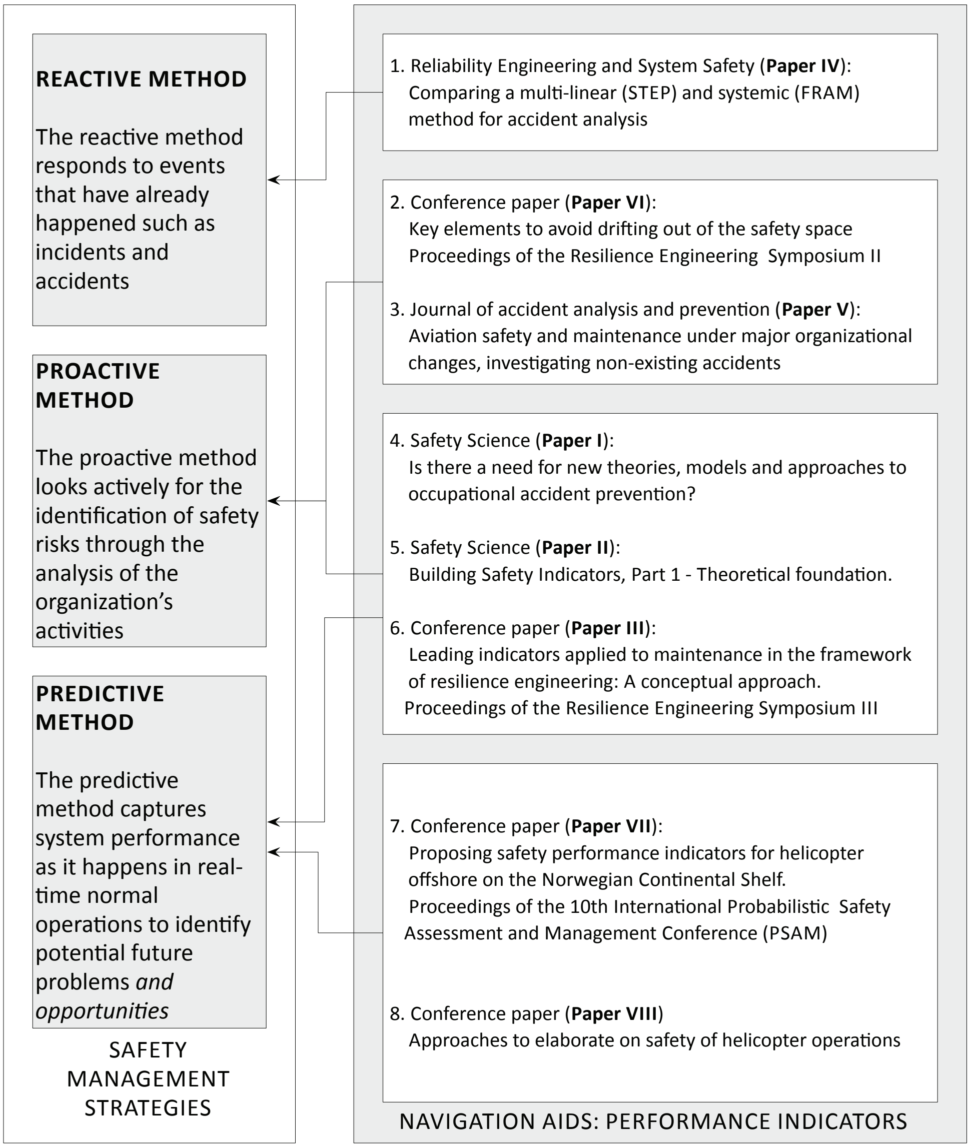


Figure 4.1 ICAO Safety management strategies (2009) and publications contribution

4.1 Paper I: Is there a need for new theories, models and approaches to occupational accident prevention?

The paper discusses occupational accident modelling challenges associated with a changing working life, and asks whether ideas from models developed for high-risk, complex socio-technical systems can be transformed and adapted for use in occupational accident prevention. Are occupational accidents mainly simple component failures or is a systemic approach to the phenomenon of some interest and value?

Methods and models for occupational accidents are inspired by developments addressing major accidents. In the last decade, progress in safety science is identified. Therefore, the intention is not to give clear and finite answers to the questions raised above. The paper invites reflections on the needs and use of accident models in occupational accident prevention.

An interesting aspect of the paper is that it anticipates the impacts of developments of the high-risk socio-technical systems on approaches addressing occupational accidents. It requests a response in the form of future developments to new theories and tools. The paper is an invitation to the research community. For industry, it shows that more methods and tools could be needed to address occupational safety.

Rather than a yes or no answer, it is necessary to consider the system that is analysed. In some cases such as simple systems the need of new models can be considered low. On the other hand, for emerging events which are more difficult to understand e.g. new advanced technologies and changes in context it might be helpful to look into other approaches such as the ones based on storytelling or studying normal processes. The paper identifies the need to explore and develop new tools in areas such as leading indicators, mapping and understanding normal operation, improvements to accident models and accident investigation.

4.2 Paper II: Building Safety indicators: Part 1 – Theoretical foundation

Development of early warning indicators to prevent major accidents – to ‘build safety’ – should rest on a sound theoretical foundation, including basic concepts, main perspectives and past developments, as well as an overview of the present status and ongoing research. In this paper we have established the theoretical basis for the development of indicators used as early warnings of major accidents. The main lessons from this paper are: i) extensive work on indicators has been carried out in the past, and this could and should have been better utilized by industry, e.g., by focusing more on process indicators related to major hazards, and less on personal safety indicators, ii) recent discussions about safety indicators have focused on the distinction between leading and lagging indicators; however, a discussion on terms should not be counterproductive and impede the development of useful indicators that can provide (early) warnings about potential unwanted events, iii) RE might contribute to indicators related to be prepared to handle unwanted events (in the paper these are called positive indicators).

The paper presents a comprehensive literature study regarding developments in safety indicators. It presents a good overview that is relevant for researchers working with the subject of indicators and those in industry who are interested in learning about different developments on indicators.

4.3 Paper III: Leading indicators applied to maintenance in the framework of Resilience Engineering: A conceptual approach

The paper explores the identification of leading indicators in aviation maintenance. Traditionally, improvements in aviation have been based on incident reporting and analyses and learning from failures. This traditional approach measures safety performance based on lagging indicators. However, there is growing concern that this information does not provide the requisite information and insight to prevent future accidents. Resilience has been identified as the ability of the system to adjust prior to, during and after a major mishap. This implies a need to recognize early signals to be able to anticipate and act correctly.

The objective of the paper is to understand leading indicators and the “why” behind the leading indicator as the basis to look forward in monitoring safety performance. With regard to the leading indicator definition applied to aviation and maintenance in the context of resilience, we focus on learning from success and failures and propose: Leading indicators are precursors based on a model of safety implying a significant possibility of a subsequent event that has an impact on safety and performance. Leading indicators can therefore provide information on changes in risk before traditional risk analyses are able to capture this change.

This conceptual paper explores the characteristics of leading indicators based on the resilience engineering perspective. It provides criteria and requirements for the development of leading indicators. The paper is relevant for academia and industry. An aspect that needs further development is the connection between leading indicators and the adaptive ecology perspective.

Some concluding remarks from the paper:
Indicators in the framework of resilience do not replace other approaches to safety performance monitoring, but increase the understanding of everyday performance. Some considerations should be taken into account while developing indicators from a RE perspective. First, special consideration is needed regarding resources for the potential of their use in the future. Second, there is a balance between intra end inter-relation within and across organizations. Third, resilience addresses the dynamics and sustainability to maintain operations. Fourth, it is necessary to have knowledge of present state and indication of transitions between different states of operation. Questions for leading indicators: which factors might push the system to critical thresholds, which kind of indicators are worth monitoring if the system follows a particular trajectory? Finally, while it is impossible to predict all possible scenarios; it is possible to improve system resilience by making sense of unintended interactions and understanding system dynamics to identify possible solutions. The paper summarizes the “why” of leading indicators as a means to look forward and to lead actions for improvement.

4.4 Paper IV: Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis

Accident models and analysis methods affect what accident investigators look for, which contributory factors are found, and which recommendations are issued. The paper contrasts the Sequentially Timed Events Plotting (STEP) method and the Functional Resonance Analysis Method (FRAM) for accident analysis and modelling. The main issue addressed in this paper is the comparison of the established multi-linear method STEP with the new systemic method FRAM. It also considers the insights FRAM provides for accident analysis. Since STEP and FRAM are based on different understandings of the nature of accidents, the comparison of the methods focuses on what we can learn from both methods, how, when, and why to apply them. The main finding is that STEP helps to illustrate what happened, involving which actors at what time, whereas FRAM illustrates the dynamic interactions within socio-technical systems and lets the analyst understand the how and why by describing non-linear dependencies, performance conditions, variability, and their resonance across functions.

The comparison presented in the paper represents an invitation to use different tools for event investigation. It shows by using an example how the method has an implication for the data that is gathered, the analysis and the findings. It represents an invitation to reflect on the selection of the method regarding the system that is analysed. The scientific community is invited to perform further work on systemic methods. Two tools are demonstrated to industry. Their combined application during an analysis provides complementary perspectives and might contribute to a more comprehensive understanding. Work is still needed for FRAM as a more structured approach for generating recommendations in terms of barriers or indicators.

4.5 Paper V: Aviation safety and maintenance under major organizational changes, investigating non-existing accidents

The objective of the paper is to discuss the following questions: Do concurrent organizational changes have a direct impact on aviation maintenance and safety. If so, how can this be measured? These questions were part of the investigation carried out by the Accident Investigation Board, Norway (AIBN). The AIBN investigated whether the Norwegian aviation safety had been affected due to major organizational changes between 2000 and 2004. The main concern was the reduction in safety margins and its consequences.

The paper presents a summary of the techniques used and explains how they were applied in three airlines and by two offshore helicopter operators. The paper also discusses the development of safety related indicators in the aviation industry. In addition, there is a summary of the lessons learned and safety recommendations. The Norwegian Ministry of Transport has required all players in the aviation industry to follow up the findings and recommendations of the AIBN study.

The paper describes part of AIBN study focusing on indicators. In this part three organizations provided inputs. Scandpower reviewed airlines with SINTEF support and the Norwegian Institute for Transport Economics (TØI) has carried out the survey. AIBN gathered all the results and provided the final conclusions.

This is one of the few papers in the literature that discusses concurrent organizational changes, safety and maintenance. The study uses a combination of a list of activity and outcome indicators, STEP method and a questionnaire. Moreover, a concrete list of indicators is presented. A warning is raised in relation to a balanced combination of statistical results and a qualitative analysis. Therefore, one possibility is to complement the proposed list with interviews. The methods and results are of interest to the scientific community and to industry. The industry implemented recommendations from the AIBN study.

4.6 Paper VI: Key elements to avoid drifting out of the safety space

The paper is based on the experience from the development and application of safety performance indicators made in cooperation with the Norwegian and Swedish Civil Aviation Authorities. The following, two main categories of safety performance indicators are discussed: outcome-based indicators (reactive indicators; measuring the outcome/ result after a loss has happened) and activity-based indicators (proactive indicators; measuring efforts to prevent accidents). The paper presents a summary of the application of safety performance indicators to assess the management of safety in aviation maintenance. Lessons learned are presented regarding the utilization of these safety performance indicators. The paper looks critically into how the indicators are used and the conclusions that may be achieved. The paper looks into safety performance indicators and how they can contribute as indicators of resilience.

The paper is an attempt to analyse the results from the AIBN study discussed on Paper I in light of the Resilience Engineering perspective. It shows how the aviation industry has a strong tradition to learn from reactive indicators. Another aspect that it is interesting is the identification of cost reduction having an impact on reducing recurring training. This paper mainly represents an academic exercise.

4.7 Paper VII: Proposing safety performance indicators for helicopter offshore on the Norwegian Continental Shelf

Over last 10 year period there has been just one helicopter accident (with no fatalities) in the Norwegian sector of offshore helicopter operations. In this case, safety monitoring cannot be based on the absence of accidents. The main objective of the paper is to suggest a combination of leading and lagging indicators to monitor safety performance for offshore helicopter operations. An approach is described to identify indicators using different perspectives: a Risk Influence Model, the Functional Resonance Analysis Method (FRAM), and lessons learned from previous studies. The approach uses accident and incident data, as well as normal operations (when nothing goes wrong). The suggested indicators were evaluated through observations and interviews/workshop with helicopter operators, air traffic controllers, helicopter deck operators and regulators. The paper discusses the approach and proposes a set of domain specific safety performance indicators. The work was carried out under the Norwegian Helicopter Safety Study 3 (HSS-3).

The innovative aspect in the paper is the utilization of the Resilience Engineering perspective and monitoring when nothing goes wrong. The combination with traditional methods is also highlighted. While Paper IV explores incident analysis, Paper VII adopts a risk analysis perspective. The paper is relevant to the research community and industry. Main contributions are 1) it addresses the monitoring of everyday performance, 2) illustrates the identification of indicators using FRAM, 3) shows a combination of methods based on different safety perspectives. It is appreciated that Norwegian industry is actively involved in the implementation of the recommendations originated from the HSS-3 study including recommendations for indicators.

4.8 Paper VIII: Approaches to elaborate on the safety of offshore helicopter operations

Accidents during helicopter transportation represent a high risk for workers on offshore installations. This paper discusses the application, achievements and limitations of two approaches, risk influence modelling and storytelling, which are used in a joint industry project known as the Helicopter Safety Study 3 (HSS-3). This study represents collaboration between scientists in the fields of sociology, human factors engineering and scientists in industry. The two methods represent different scientific paradigms, each with their own perspective on understanding safety and risk. The HSS-3 risk model is based on the use of risk influencing factors (RIFs) and accident categories. The HSS-3 risk model represents quite simple influence modelling, and is not designed to account for interdependency between RIFs, and does not describe the influence from the operational context on risk. As a consequence, the storytelling is combined with the risk modelling approach to identify the effects of organizational changes on work practices and safety. The application of these two methods leads to the following question: To what extent do the methods complement each other, and how can they be combined to provide a better understanding of helicopter operations? Shifting the level of abstraction from linearity and decomposition (i.e. RIFs) to storytelling and richer data sets triggers new ways of understanding and provides new knowledge about the “internal life” of the system. It is also noticed that RIF and FRAM modelling in combination with inputs from the industry triggered the identification of the topics analysed using storytelling. The combined use of these approaches provided a better way to identify changes in risk, safety threats and also safety improvement measures. While the identified safety measures and findings are domain specific, the combined use of these two approaches may also be of interest to other industries.

5 Main results and discussion

This chapter summarizes the significance of the research carried out. Section 5.1 presents indicators identified in the studies. Sections 5.2 and 5.3 give the contribution form the thesis in relation to the academia and to the industry. Section 5.4 discusses the research work in relation to its objectives. Finally, Section 5.5 presents an evaluation of thesis work.

5.1 Revealed indicators in the case studies[3]

Papers V and VI suggest a combination of quantitative indicators with qualitative data while Papers VII and VIII include interviews and observations. Papers VII and VIII document results from the HSS-3 study, the study recommended a priority list of ten indicators for the helicopter operators, five to air traffic services and two for helidecks. These indicators reflect a pragmatic selection and need continuous evaluation to identify if new aspects require attention. It is possible to see an evolution in the recommendations from indicators related to specific organizations (Papers V and VI) to include indicators related to the interactions across organizations (Papers VII and VIII) e.g. collaboration and communication pilots – helicopter deck personnel; contract conditions.

The qualitative data complement the information provided by the quantitative indicators. The data suggest the use of interviews and observations as a means to provide better information of what works well. Indicators focus on specific aspects and obscure other aspects, therefore observations are also important to identify changes or new indicators that might have significance.

Finally, Paper VIII argues for the use of storytelling to illustrate the impact of organizational changes. Rather than factors, the stories contain rich information that supports recognition of unwanted effects of changes.

Paper V: Aviation Maintenance under major organizational changes, investigating non-existing accidents & Paper VI Key elements to avoid drifting out of the safety space
Major concurrent organizational changes happened in the Norwegian aviation between 2000 and 2004. Indicators were adapted to maintenance organizations. A subset of indicators important for safety trends are shown in Table 5.1

Table 5.1 Indicators - aviation maintenance and organizational changes (Papers V & VI)

Importance	**Reactive**	**Proactive** **(Note: indicators require qualitative data for their interpretation)**
Important	Accident rate, Deviation rate	Audits: • Number of internal and external audits Training: • Number of continuing training or recurrent training Maintenance programme: • Backlog, hold item list • Minimum Equipment List items
Average	Serious incident	Audits: • Number of deviations identified • Number of dispensations requested Training: • Number of certified personal per type of certificate per type of station/year Maintenance programme:

---

[3] Appendix D. Presents list of indicators and information requested to the operators

Importance	Reactive	Proactive (Note: indicators require qualitative data for their interpretation)
		• Part of maintenance programme that is based on in-service experience Economy • Number of aircraft types • Number of implemented safety measures

Paper VII: Proposing safety indicators for helicopter offshore operation on the Norwegian Continental Shelf & Paper VIII Approaches to elaborate on the safety of helicopter operations

Table 5.2 lists indicators for a specific scenario. It shows the following elements:

- Indicators related to flight planning airworthiness of helicopter, helicopter deck-rig status, quality of weather information
- Communication, collaboration, competence helicopter-helideck personnel
- Penalties and influences of contracts

Indicators identified using FRAM are a combination of helicopter operators' indicators and interactions across organizations helicopter operators and helideck.

Table 5.2 Scenario specific indicators using FRAM
Helicopter Landing on Helideck (Paper VII)

Indicators	Operationalization
Airworthiness of helicopter (quality of technical condition)	• **Minimum Equipment List (MEL)**: The status of critical systems. MEL can be an indicator which says something about to what extent the organization maintains continuous airworthiness. • Data from continuous use of Health Usage Monitoring System (**HUMS**) data for early identification of errors • Indicators related to **performance of maintenance work**
Quality of rig/facility report	• Quality assurance of personnel that make reports. **Training of weather observers**
Quality of information, technical equipment on helidecks	• Indicators related to the status of technical systems, **reports from facility**
Quality of updated procedures, which describe practices according to the helicopter type	• Conducted through audits or observations of normal operation of the helicopter operators, ATM/ANS and helidecks. **Procedures, audits and compliance** • Active use of **Flight Data Monitoring (FDM)**
Quality of communication between helicopter and helideck personnel; procedures and practice	• Here there are individual differences as regards the use of language, and helicopter-related phraseology (especially for small floating facilities). Language knowledge varies on different facilities, particularly those with floating helidecks (ships). • **Observations** to provide a qualitative assessment of the development in regard to communication. • **OLF MANUAL/regulations update**. The number of changes in revisions. Harmonizing of data provided to pilots, use of radio frequency (in some areas there is increased use of VHF). Passenger manifest is provided with METAR and TAF
Quality of team work (*CRM - Crew Resource Management*)	• Quantitative goals are difficult. **Observations** more important than interviews, but both can be used, see what works well and what does not work well
Competence to operate on the Norwegian Shelf	• **Observations** of normal operations. Use line check proactively. **Observations** in line check are conducted once a year, with a simulator twice a year.
Contract conditions – "*penalties*"	• If penalties are experienced during delays.

Indicators	Operationalization
	Deviations in time from the operations centre; the number of bought days off per year, use of overtime
For floating helidecks, the occurrence of visual clues	• Status on helideck, according to CAP 437, OLF guidelines

Based on existing suggestions from the literature, RIF, FRAM, seminars and interviews with industry/experts representatives, Table 5.3 presents a subset of prioritized indicators proposed to the industry, helicopter operators and air navigation service providers.

Table 5.3 Example of prioritised indicators – helicopter operations (Papers VII & VIII)

Lagging	**Leading** **(Note: indicators require additional qualitative data for their interpretation)**
Incidents, number of repeating technical or pilot reports for safety critical systems e.g. communications, navigation, rotor	Helicopter operators Technical condition: • Continuous use of Health Usage Monitoring System • Number of deferred defect list items combined with minimum equipment list (indicate availability of spares and resources) Quality of procedures, which describe practices according to helicopter type, maintenance programme, revision of operational procedures • Compliance, audits and observations to reveal deviations between procedure and practice • Revision of procedures in the last period • Update in maintenance programme together with information send to technicians about updates • Number of "noticed to pilots" or "information to crew" (revision of procedure) Maintenance / pilots crew, workload • Planned crew versus real crew per station per shift • Use of overtime in relation to exemptions / number of purchased days off per year Cooperation and communication • Use interviews and observations to cover the condition of cooperation and communication Penalties • Follow-up penalties regime and how it influences the organization

There has been an evolution in the papers and proposed indicators:

- A combination of quantitative and qualitative data is advocated.
- In the first studies indicators were recommended to specific organizations while in the last studies indicators also are related to interaction across organizations.
- The differentiation between lagging and leading is assessed as important because it enables the focus to be placed on what it is relevant to continue operations.
- From indicators related to failures towards indicators that support production and everyday performance.
- Rather than a solution that can be applied at any time, a set of indicators is proposed requiring verification, interpretation and renewal in relation to specific context. The interpretation of these indicators together using a model in my case the FRAM model supports the analysis of the system as a "whole" on a specific context.

5.2 Scientific implications of the thesis

The material presented in the thesis contributes to new insights in the following areas: indicator concept and its criteria definition, utilization of different methods (FRAM, STEP, RIF and storytelling) to identify indicators, application of Resilience Engineering concepts and the FRAM in practice. The feedback from the method application contributes further to the development of the FRAM method.

Identification of indicators – Indicator concepts and positioning

Regarding indicator concepts, I support the definition of indicators as "proxy" measures about important aspects in the underlying safety model (Wreathall, 2009). Paper VIII demonstrates that there is a relation between the safety understanding and the identification of indicators (Papers VII and VIII). This understanding influences the selection and interpretation of indicators. Traditionally, the understanding and management of safety has been related to the monitoring of failures and accidents. In this case, most of indicators provide information after the fact. This type of indicators can be defined as lagging indicators. The Resilience Engineering perspective aims to ensure safety and that the functionality of the system stays intact under stress. The thesis argues for a balanced utilization of both lagging and leading indicators (Papers VII and VIII). The differentiation emphasizes not only on acting upon unwanted past events, but taking into account everyday operation to identify aspects that may become critical in the future. So leading indicators relate to possible future states and what may become critical. While there is a disagreement in the scientific community about the usefulness of the differentiation between "leading" and "lagging", I disagree with Hopkins (2009) in the aspect that the differentiation is might not be useful. I support Hopkins' (2009) argument about indicators about the clarification of the term indicator each time that it is used. In conclusion, indicator terms should be defined related to the specific context and agreement on what is their purpose. For example, past events can be defined as lagging while some indicators identified from the analysis of daily operation can be defined as leading. Papers VII to VII demonstrate that the utilization of specific indicator definitions is a good baseline for identification of indicators. Moreover, based on experience from the case studies and discussions with other researchers, the thesis proposes a more granular definition introducing lagging, current and leading indicators. This differentiation has not been tested in practical applications and further research is needed.

Identification of indicators – Diversity and modelling

One main lesson from the thesis is to advocate a diversity of methods and understanding. The approach proposed in the thesis agrees that a diversity of measures allow the selection of the ones that fit the context and provide multiple lens to see the same area of concern (Page, 2011). This thesis represents an invitation to polyphonic dialogue between different perspectives (Bakthin, 1986). This is not easy since we have different biases in our understanding (Reiman and Rollenhagen, 2010; Renn, 2008). With polyphony is here meant that applying more than one perspective not only enhance the number of issues to consider, but more important that the perspectives trigger and empower each other. These two objectives may be difficult to combine. When applying more than one perspective it is common to contrast the perspectives in order to see the strength of each perspective. In many cases this will also prevent their biases because issues considered to be the strength of one perspective are then often not considered from the other perspectives. In the work documented in this thesis it was possible to experience that different perspectives complement and trigger each other (at some limited extend). The RIF, the FRAM and inputs from industry experts triggered the identification of relevant topics to be studied by storytelling. The RIF modelling supported the identification of a relevant scenario for the FRAM modelling. Very few instances where identified where the methods triggered each other and gave source for resonance because that was not the focus of the work. On the other hand there are more instances that same findings from different methods supported the identification of recommendations.

Papers V to VIII use different approaches to the identification of indicators. Each approach directs attention to specific issues and obscures other issues. The traditional risk analysis studied represented by RIF modelling identified factors related to accidents and incidents. This type of modelling provides a static picture of the socio-technical system. FRAM analyses everyday operation and indicators based on this modelling are related to the factors that maintain the system functioning under specific context of operation. Papers VII and VIII show that it is possible to propose specific indicators based on FRAM and RIF. FRAM addresses indicators related to continuous operation while RIF is focused on indicators that represent threats or hazards to the operation. The effects of organizational changes were more explicit with storytelling.

The thesis represents a step forward towards the identification of a combination of leading and lagging indicators. Papers VII and VIII propose a combination of such indicators. Aviation improvements are often based on learning from rare unwanted events. Many improvements have been achieved. While lagging indicators enable to act after the fact. Leading indicators support foresight and acting before the threat materialize in an unwanted situation. For example Paper VIII identifies an indicator related to the utilization of penalties having impact as more pressure to operational control centre and pilots. Indicators should be analysed in relation to a specific context. Therefore, this example shows also that one indicator is not enough to have a picture of a situation. The effect of penalties pressure is different if resources (pilots, helicopter, time) available are available or not. Then, Paper VIII represents an argument for leading indicators as a set of indicators related to specific context. In addition, Paper VIII shows the utilization of modelling and criteria for identification of indicators. Sensitive is a criteria that is of particular relevance for leading indicators. In this way it is possible to assess the situation taking into account the socio-technical system, its changing environment of operation and significance of changes in a period of time.

The introduction of Resilience Engineering as a safety perspective generated different reactions in academia and industry. RE creates room for discussions, innovation and questions about established ways to address safety. It shifts focus from failures and unwanted events to unexpected events, successes and everyday operation. It aims to ensure that the functionality of the system stays intact under stresses at the same time that safety is not compromised. To understand what may go wrong it is necessary to understand what goes right (Hollnagel, 2011). One contribution to RE is related to the indicator concept. The leading indicators relate to different status of operation and look at tipping points from one state to a new state of operation. Another contribution is related to FRAM studies and feedback to its development. Paper IV analysing an incident contributes to test instantiations (sets of couplings among functions for specific time intervals). It concludes that FRAM provides new aspects, addresses operational context and dynamics of the system. Besides answering what happened and it answers the question why the event happened. Based on the results from Paper IV, FRAM was a potential candidate for the identification of indicators. Then, Paper VII uses FRAM to analyse everyday operations. Paper VII concludes with a proposal of concrete indicators. Cases have been developed in close cooperation with the industry. Results have been presented to the industry and academia contributing to the development of the method. More studies are necessary to generalize findings from these case studies.

Regarding different methods for the identification of indicators inspired by Resilience Engineering, the thesis addresses only some capabilities of resilient systems. Hollnagel (2009) proposes four essential capabilities. These capabilities are the capability of learning from experience, respond to regular and irregular threats, monitor in a flexible way and anticipate threats and opportunities. The methods proposed in the thesis are more related to the capacity to monitor and to anticipate. One method inspired by Resilience Engineering that applies in industry cases is FRAM. Papers VII and VIII proposed a set of indicators and there is still a need for organizations to use the proposed indicators on a regular basis in order to give feedback and improve the framework. Paper VII focused on the analysis of everyday operations and identification of indicators that support this kind of operations. This type of analysis contributes to a better understanding by considering the influence of the context in the actual performance. The method supports the analysis of the strengths and weaknesses of everyday operations. Other resilience work relates to situations where the system is challenged and is pushed towards its limits of operation (Woods and Wreathall, 2008; Branlat and Woods 2010; Lay, 2011). These studies are related to support the capacity to monitor, respond and learn from experience. The utilization of the concept of fundamental trade-offs can also be used to evaluate performance and adaptive capacity. The approaches used in the thesis aim to obtain a better understanding of functionalities that are critical for everyday successful operational performance. While some approaches address trade-offs, FRAM modelling captures the influence of the context, dynamics and areas that are required to continue operation. So in FRAM, some of the trade-offs (optimality-resilience, efficiency-thoroughness, specialist-generalist, distributed-concentrated) can be analysed as variability of normal performance having an impact on the performance of the functions. The perspicuity trade-off cannot be covered by FRAM since the method analyses the system from an RE perspective.

Assessment of indicators suitability and quality

The thesis work identifies key criteria for assessment of safety indicator quality and suitability. These criteria are meaningful, sensitive, reliable, measurable, verifiable, inter-subjective, operational and affordable. There are several studies about criteria for indicators and I agree with Tarrants (1980) and Kjellen (2000) that there is very little possibility that one indicator complies with all the criteria. One indicator that scores low in relation to one criterion might be complemented by other indicator scoring high. The quality of the indicators should have a minimum level of acceptance to ensure that the indicator can be used in a decision-making process. Therefore, a combination of indicators covering these criteria is recommended. Paper VII shows validation of the criteria for a specific industry with the organizations. In this study, representatives from the regulatory authorities, managers of the company, operational and technical personnel participated in the quality and suitability assessment of indicators based on the criteria presented above. A large number of candidates for indicators were identified in the helicopter study. Through a systematic evaluation process, each indicator candidate could be ranked against the criteria based on their suitability in a decision-making process for enhancing safety (Paper VII). The different approaches to identify indicators e.g. FRAM provide an exhaustive list of indicators. So, the criteria can be used as an instrument for the selection of indicators in order to assess their quality and suitability. For example the characteristics “inter-subjective” and “sensitive” are highly relevant for current and leading to have a consensus on aspects that are critical. It is not possible to generalize from one study. So the conclusion is that the criteria shown in Table 2.1 are a good start for the selection of indicators and need to be validated with the industry in each case and context they are applied. More case studies using this set of characteristics are needed to generalize this set of criteria for the aviation industry.

Use of indicators in a decision-making process

The focus of the thesis is on the identification of indicators. A main purpose of the indicator is support decision-making regarding when to decide, where and how to take action (Hale, 2009). The first step in the case studies is the definition and delimitations, the actors involved and who is going to use the indicators. As mentioned in Section 2, Rasmussen (1997) refers to decision-makers as “controllers” in relation to hazard sources.

Their interactions can be described from a hierarchical level structure from government, regulators, company planning, to physical processes and actors activities. Normative models do not emphasize what is done but rather what is to be performed. Indicators can relate to errors and deviations in relation to the work "as expected". In the RE literature, Woods and Branlat (2011) refer to polycentric control architectures, this structure is not hierarchical and it shows interactions and adaptation within and across organizations. The decisions are performed from different centres. Each centre adapts activities and coordinates activities in relation to other centres. RE modelling relates to models of "work as performed". Papers VII and VIII are related to aviation. Decision-makers identified can be related to "centres" as operational represented by pilots, maintenance, air traffic controller, helicopter deck operations. Other centres are represented by management engineering, operational, logistics, maintenance quality and safety, company CEOs and the regulator. Each centre has its own interest and needs (as already remarked by Rasmussen, 1997). Paper VII proposes a set of indicators in relation to "work as performed". The decision process uses many inputs for example the results from audits, and risk analyses. Indicators represent one input to this process. It is required to define at what point the indicators are related to a critical or non-critical situation. Since, indicators are often interdependent. The decision process for acting may require analysis and assessment of a set of indicators. Hence, it is necessary to analyse a set of indicators in a specific operational context. FRAM modelling is a proposed tool illustrating how the performance of the functions is affected by the context of operation and how the indicators are interdependent. The modelling supports an overall evaluation of the interrelation between indicators. The input for decision making based on indicators must therefore be based on an overall assessment of a set of indicators related to specific operation and not only individually. For operational areas a more comprehensive set of indicators related to specific context is proposed while for management and regulators a subset of these indicators is proposed. The thesis work presents the identification of indicators using different methods and perspectives. The papers have demonstrated that indicators can be identified from RE modelling. Further work is needed to implement the indicators in the safety management system of the organizations studied.

Theoretical choices, indicators, drift and anticipation

The term "drift" relates to the "work as planned" and the "work as performed in the real work" (ICAO, 2009). The literature shows how difficult is to define boundaries of operation that makes more difficult to identify drift (e.g. Rasmussen, 1997). Paper VII documents the FRAM approach selected to model and analyse the "work as performed". The thesis contributes to FRAM developments. Paper IV documents the use of instantiations showing the dynamics of the system. A scenario is illustrated as a "functional slide show"; where some particular functions are active at some specific time intervals. This approach is found to contribute to the gap in relation to enhance the identification of indicators in relation to specific operation and functionalities. Paper VII illustrates the relations between operational and economic influences. FRAM modelling using a systems view shows how variability in one function is damped or amplified by another function. Paper VII presents that FRAM enables the selection of indicators candidates related to wanted and unwanted variability. Further work is required in the operationalization and use of indicators. The thesis has also explore combination of established and RE methods (Papers IV, V, VI, VII and VIII). This combination is an example of diversity and use of different safety perspectives. It is a challenge to combine these different types of thinking. A question regarding interest for a unique model or harmonized way to present analysis and results was raised during the thesis work. The papers listed above represent argument for triangulation and utilization of different types of data and models rather than unique model. A unique model will prioritize and select only some aspects, while methods with different perspectives provide more rich data and can supplement each other.

5.3 Contributions to the industry and practical applications

The main contribution is related to the monitoring of everyday operations and new ways to identify indicators. Cases studies documented in Papers IV, V, VI and VII propose specific indicators to the industry.

Based on experience from the literature review, discussions with other researchers, industry experts and empirical studies carried out in this thesis, the conclusion is the utilization of different safety perspectives is central for the identification of indicators.
There are many developments in relation to the utilization of established approaches for identification of lagging indicators. For the utilization of FRAM to identify indicators, I recommend the following process:

1. Define the area of concern, the purpose of the specific study and the actors involved. The area of concern will enable the selection of scenarios where indicators need to be identified. Then, I propose the selection of scenarios related to safety critical operations. Paper VII presents offshore helicopter landing on a helicopter deck as a relevant scenario. For daily operations data can be captured by interviews and observations. A second alternative for scenarios is to analyse when system is pushed beyond the design envelope in this case near misses could be analysed (this last option is not studied in this thesis but has been addressed in the oil & gas industry, see www.buildingsafety.no). The actors involved should include decision makers and those whose performance is measured.
2. Define an industry expert group consisting of industry representatives and safety expertise. It is essential that the people who will be using the indicators are involved in the process of identification and selection of indicators
3. Present a definition of indicators, its purpose and criteria. Since there are different understandings of the indicator terms, a careful definition these terms is required in the context of the study. Validate the criteria with industry representatives. Table 2.1 is a useful starting point to characterize a good indicator.
4. Apply the Functional Resonance Analysis Method to identify and propose a candidate for indicators. The modelling is an iterative process that needs to be verified by the industry. The unwanted and wanted types of variability that cannot be damped within or across functions are candidates for indicators.
5. Use the criteria to define an initial set of indicators and use the indicators to monitor safety performance. Operationalize and validate these candidate indicators together with industry experts (personnel whose performance is evaluated and personnel who will use the indicators).

The studies presented in the thesis addressed points 1 to 5. There is still a need to use the proposed indicators for a period of time, to evaluate and if necessary renew the indicators.

5.4 Contribution to the objective and answers to the research questions

The objective of the thesis is to explore existing concepts and methods to address the question: *How do we identify that a system drifts or experiences sudden changes in the safety space?* sub-questions and specific objectives are developed to answer the main question. Eight papers address these sub-questions and represent the main results. The sub-questions and answers are recapitulated in this section.

1. What do we mean by safety performance indicators and related terms like leading and lagging indicators? The specific objective is to classify current knowledge in

relation to indicators that could be applied in the identification of drift or sudden changes.

As mentioned in Chapter 2, Papers II and III present a clarification of concepts and developments regarding safety performance indicators. Today, there is still confusion and disagreement about the terms leading and lagging indicators. This may be explained by the multidisciplinary nature of the safety community. Consequently, to avoid confusion the term should be clarified carefully each time it is used (Hopkins, 2009). The result will be that there is common understanding in a specific context rather than a generalization across different disciplines. The disagreement about the use of the terms lagging or leading indicators should not impede addressing the need for the reactive, proactive and predictive approaches that support the capability to continue safe operations. The difference is relevant to seek for a balanced composition of different types of indicators. Organizations where safety is critical use this composition as the basis for decision-making. Safety performance indicators cannot solely be based on hindsight and error tabulation. In general, leading and lagging indicators can be considered on a time scale where leading indicators precede and lagging indicators follow an unwanted or unexpected event. In everyday operations, when an unwanted outcome is not present there is still a need for safety performance indicators. The work in the thesis has been an iterative exercise and the understanding about indicators has been continuously updated. This process has resulted in the following distinction (Papers II, III and VII):

- Lagging indicators may be data that were registered or collected in the past, and which are now used to understand what has happened. They may also have been used or interpreted at the time, in which case they were current indicators. Lagging indicators may often include aggregated data that illustrate a historical development or document a trend. Examples of lagging indicators include event statistics and trends.
- Current indicators show the state of the system at the moment. Examples of current indicators include production rates, resource or inventory levels, or the number of aircraft in a sector.
- Leading indicators are the interpretations of measurements of current states with regard to what may happen in the future. In this way, measurements are used as predictors, rather than as status or performance indicators. Leading indicators are interpretations of a set of indicators that pertain to how the system as a whole copes with conflicting or demanding goals. The FRAM analysis supports this interpretation of non-linear and dynamic situations. Examples of leading indicators are the combined interpretation of available resources, technical status of safety critical components and time available (time available can be determined by contracts or regulations).

The distinction between lagging and leading is used in Paper VII and was presented to industry and the scientific community. Different responses came. First, there is acceptance about the need for leading indicators. Second, there is a wish to have simple indicators that can easily be implemented. Third, the industry is satisfied with the attempt to identify leading indicators. In particular the helicopter industry is working with the implementation of the indicators recommended in the study. The difference is important to create awareness of a balanced composition of leading and lagging indicators. The composition aims to recognize changes affecting performance. The leading indicators require the interpretation of their significance in a specific context of operation. Hence, leading indicators require competence for their interpretation and implementation. One way to distinguish lagging from leading indicators is to treat lagging indicators as naïve documentation of what has been going on, whereas leading indicators require understanding, knowledge and interpretation in order to point towards the future. The aim of this thesis has been to take a step forward to achieve a balanced use of a broad range of indicators. The concept of current indicators was not used in the studies but this definition is found useful since not all current indicators are leading. Still, these indicators provide information about current performance.

2. Do we need new approaches for the identification of safety performance indicators? The specific objective is to understand the contribution of different methods for the monitoring of safety performance.

Paper I discusses the need for new theories and approaches. Modern complex socio-technical systems require a combination of known approaches and new approaches for the identification of safety performance indicators. As mentioned in Chapter 2, most of the work on indicators is still focused on the collection and utilization of accident, incident and failure data. In aviation, measurements of safety performance are mainly based on accidents and incidents that are further decomposed in categories to identify particular safety issues. This categorization has enabled several improvements on specific issues and accidents are fortunately rare. The aviation industry has experienced several improvements by use of lagging indicators (reactive methods). More use of such indicators may be seen in this sector in the future. Still, there is common understanding that the absence of accidents is not a sufficient measure of safety performance (Van Steen, 1996). This trend has changed with the introduction of surveys, audits, use of flight data monitoring and direct observation methods like Line Operations Safety Audit (flight operations) and Normal Operations Safety Survey (air traffic management). These management tools are based on managing errors and threats. Hopkins, (2000) argues that organizations tend to attend to what is being measured rather than what is important. Moreover, accidents are rare and there are very few approaches looking at everyday safe operations. It is necessary to understand the way that operations are carried out and the adjustments that are required to continue operations. Therefore, new approaches are still needed to provide organizations with adequate understanding of the current state of the system, and predict possible problems and opportunities.

3. Which methods can be applied to identify lagging and leading indicators in the perspective of Resilience Engineering? The specific objective is to explore and test different methods to identify safety performance indicators.

Resilience Engineering sees safety as “the ability to succeed under varying conditions” (Hollnagel, 2011). A main objective is to increase the number of successful performances. Four essential abilities are identified to achieve this. These abilities are related to coping with the factual, the actual, the critical and the potential as described in Section 2.3. The ability to monitor what might become a threat (dealing with the critical) and the ability to anticipate threats and opportunities (dealing with the potential) are addressed in this thesis. Therefore, it explores and tests established and relatively new methods to understand how the system operates under many constraints. One method based on Resilience Engineering principles is the Functional Resonance Analysis Method (FRAM; Hollnagel, 2004). FRAM studies have been performed to explore the identification of indicators. By applying FRAM, the thesis has contributed to the development and dissemination of the method (Papers IV and VII). In addition, a combination of methods is explored. Papers VII and VIII present a combination of storytelling, risk influence modelling and FRAM. These papers show how the methods provide complementary information.

Each method represents different approaches to the understanding of safety. Each of these methods is described in Chapter 3. The studies demonstrate how a combination of methods enhances current knowledge and understanding as well as supports the identification of indicators. Most of the established methods look into failures and deviations. However, the indicators identified using FRAM aim to monitor changes in everyday performance that normally are not picked up or reflected using established methods. These indicators do not necessarily represent failures or deviations. For example, Paper VII documents a relation between time allowed, penalties and flight planning revealed by the FRAM analysis. The study identified the need to improve the management of contracts reducing unwanted variability such as unnecessary pressures on the front line (pilots, maintenance personnel and the operation control centre). Hence, one advantage of FRAM modelling is that it considers the influence of the context on actual performance.

Table 5.4 presents the strengths and weaknesses of the methods explored in the thesis.

Table 5.4 Comparison of STEP, FRAM, RIF, storytelling and Resilience Engineering.[4]

	Functional Resonance Analysis Methods (FRAM)	**Sequentially Timed Events Plotting (STEP) or Risk Influence Modelling (RIF)**	**Triangulation, list of activity and outcome indicators**	**Storytelling**
Model type	Systemic	Sequential	Mixed data sources	Narratives
Coping with the critical, monitoring performance & coping with the potential, anticipate threats and opportunities	Accident analysis & Monitor normal operations	Incidents and accidents are mapped	Based on previous studies, use quantitative company data	Depends on the analysis. Stories are sources of rich information that support sense making
Type of systems	Address socio-technical systems	Main focus is on systems where the logical structure is easy to understand	Main focus on organizational aspects	Organizations
Strengths and weaknesses	Multidisciplinary study	Multidisciplinary study	Multidisciplinary study	Multidisciplinary study
	Non-rigorous description of every step	Rigorous description of every step	Non- rigorous description of every step	Open design of the study
	Focus on function and not the individual	Focus on failures, errors and what may go wrong	Focus on numbers quantitative data and qualitative data supplement each other	-
	Address dependencies and operational context	Weak on system dependencies and context		Address context and dependencies
	Analyst with training and practical experience	Require analyst with some experience	Require analyst with some experience	Require analyst with experience
	Difficult to communicate	Easy to communicate	Easy to communicate	Easy to communicate

4. What are the relevant safety performance indicators that could be applied to aviation and are they accountable to safety? The specific objective is to explore which indicators may be identified for safety monitoring

The way that safety is understood drives the identification and utilization of the indicators. Resilience Engineering sees safety as an integrated and essential part for production.

---

[4] Based on experience from the thesis and methods applied, several cases are needed to validate and generalize these findings.

Corresponding to this understanding monitoring the variability in everyday performance is required to maintain operations. The application of the FRAM analysis can support the identification of patterns, relations and local adaptations required for a successful operation. Paper VII presents a combination of lagging and leading indicators using established and relatively new methods.

The indicators which are identified are related to specific cases (Paper VII). These indicators are a combination of quantitative and qualitative data. Paper VII demonstrates that the modelling supports their identification. It is also possible to illustrate the relation between indicators and the actual performance. An interesting aspect is that some indicators relate to economy and other to operations i.e. penalties and regularity.

The application of storytelling illustrates other aspects such as feelings and frustrations that are not addressed by the other methods explored. For example the combination of changes in the framework conditions, conflicts and frustrations are not reflected in the other methods. The stories are rich and together represent an early warning for the organizations. It shows aspects that require improvements such as the integration of two management cultures. One management model is a Norwegian one in which management and employees represent different interests but cooperate. Another one is a more authoritarian model. An example revealed by the storytelling is the need to improve spare parts availability to solve technical problems in time.

A set of technical, human and organizational indicators related to a systemic perspective may be accountable for safety. Thus, indicators should provide information about how the system as a whole operates. It is necessary to revise understanding to enable early identification of new opportunities or threats. This thesis represents a step forward in the identification of indicators and further work is required. Examples of such indicators are provided in Table 5.5

Table 5.5 FRAM, RIF, storytelling and example of indicators.

	Functional Resonance Analysis Methods (FRAM)	**Risk Influence Modelling (RIF)**	**Triangulation, list of activity and outcome indicators**	**Storytelling**
Model support interpretation	Candidate indicators are identified related to a specific context. Monitoring performance variability or adaptive capacity	Candidate indicators are identified related to incidents or unwanted events	Indicators based on previous research	Stories are indicators and cannot be decomposed in factors without losing meaning. Factors can be identified to support the RIF work.
Examples	Indicators related to the airworthiness of the helicopter Indicators related to the use of penalties	Incorrect weather information Oil leakage Fuelling events	Near misses Backlog Recurrent training	Unintended consequences of new organizational structure Lack of spare parts Quality of training

5. How is a specific operation carried out in the real world and what implications for indicators and models are available to improve the monitoring of safety? The specific objective is to explore which indicators may be identified for the monitoring of normal operations.

The monitoring of everyday operations enhances the understanding of how the system works and the influence of its context of operation. This approach complements normative descriptions which look how the system should operate or operated. Hence, this approach enables one to understand how system as a whole operates and adapts under varying conditions to continue operations. The aim is to identify indicators related to variability of performance (Hollnagel, 2004) or to the adaptive capacity (Woods, 2011). Indicators derived from monitoring specific operations using the FRAM method make it possible to monitor and control variability that is critical to ensure the functioning of the system. This implies an invitation to use models and identify leading indicators that allow the organization to be able to respond to changes and disturbances.

The indicators are related to a specific context and the utilization of the model enables its interpretations. The application of the FRAM method shows that it is possible to identify functions that are essential for operations. Candidates for indicators could be derived from these functions. Indicators are related to the quantity and quality that are required to successfully perform an operation. These indicators depend on inter-subjectivity and consensus between the analyst and those involved in the operation that is analysed. This is required to differentiate and identify unintended system interactions (Paper VII). Rather than using single indicators, both model and a set of indicators are used for interpretation. Since, the purpose is to address dynamic phenomena, it is necessary to revise the model periodically to identify important changes.

5.5 Evaluation of the research

Safety researchers and the aviation industry are target groups identified as beneficiaries of this work. These two groups represent different criteria for the evaluation of the benefits achieved. One is related to the contribution to new knowledge and the other is related to a practical approach that will enhance safety and operations.

The Research Council of Norway (2000) proposes three criteria for the evaluation of the quality of applied research. First, originality related to the extent to which the research presents innovative use of theories and methods. Second, solidity related to the extent to which the statements and conclusions are properly substantiated. Third, relevance related to the extent to which the work contributes to fill gaps in previous research and contributes to bring the research a step forward. This section discusses the thesis against these criteria.

The thesis presents a degree of originality by applying a combination of established and relatively new methods. For example the combination of risk influence modelling, the functional resonance analysis method and storytelling may be considered innovative. The monitoring of a specific operation when nothing goes wrong using some of the premises from the Resilience Engineering perspective is relatively new. Paper VI on leading indicators borrows ideas from an ecology perspective (Gunderson, 2002) to illustrate how changes in the socio-technical system are influenced by internal and external processes.

Regarding solidity, the methods for data gathering and analysis have been documented to support the findings. The studies have been subjected to internal and external quality reviews. The results have been discussed with industry. Inputs from other researchers who are active in the area of indicators have been implemented in the papers. The papers have been subjected to peer review in international journals and conferences.

Regarding relevance, this work addressed and identified the need to provide clarification of concepts and tests of relatively new and established methods. The thesis thus contributes to advancing research. For example the comparison between STEP and FRAM provides some of the new aspects addressed by systemic methods. In the scientific community and industry there is increasing interest in systemic methods. I participated in seminars to present FRAM to the industry and discuss the development in the Functional Resonance Analysis Models. The FRAM method is relatively new and some steps are not defined in detail so it was necessary to improvise during the seminars. More developments are still needed in the FRAM method. The thesis work has also contributed to gain more knowledge about Resilience Engineering. This knowledge has been presented to students through lectures, to industry through conferences and courses and to the research community through conferences and academic workshops.

The following presents a quality evaluation of the responses to the sub-questions, progress in relation to the theoretical framework and practical solutions to the industry.

What do we mean by safety performance indicators and related terms like leading and lagging indicators?
The aim of this question is to classify current knowledge in relation to indicators that could be applied in the identification of drift or sudden changes. Papers have been published providing an overview and clarification of definitions. While most of the approaches focused on analysis of failures, this thesis promotes the analysis of daily operations. This approach will enable the understanding of how the system actually performs and sustains operations. Specific indicator definitions, criteria for good indicators and methods have been proposed to the academia and the industry. These concepts have been applied to specific case studies. Further work is needed to assess the utilization of the proposed indicators.

Do we need new methods for the identification of safety performance indicators?
The specific objective of this question is to understand the contribution of different methods to the monitoring of safety performance. Most of the methods related to the identification of indicators use accident, incident and failure data. Other methods based on the perspective of RE need to be tested. New methods focusing on the ability to continue safe operations in the presence of stress are needed. The proposed indicators in the papers match specific operational situations and RE. The thesis contributes to the use of different approaches that provide complementary information to established approaches. Concrete methods and examples are proposed to the industry in Papers V to VII.

Which methods can be applied to identify leading and lagging indicators in the perspective of Resilience Engineering?
The question aims is to explore and test different methods to identify safety performance indicators. New methods are recently proposed using the RE perspective. The thesis contributes to the development of the Functional Resonance Analysis Method, FRAM, which is based on RE principles. Other approaches like the proposed by Wreathall and Woods (2006), Øien et al. (2010), Woods and Branlat (2011) or questionnaires using RE perspective have not been tested and compared. FRAM is still under development and more clarifications are needed for the analyst. Paper VII shows that FRAM supports the identification of candidates for leading indicators.

What are the relevant safety performance indicators that could be applied to aviation and are they accountable for safety?
The specific objective of the question is to explore which indicators may be identified for safety monitoring. Understanding safety guides the identification of indicators. Ensuring safety through RE relates to the ability to adjust its operation so the system can sustain operation. The thesis work succeeded in modelling everyday operations and in identifying indicators that are critical for specific cases. It contributes to the utilization of different methods for the analysis of daily operations. Indicators proposed are context specific and

there is no one solution that fits all. The indicators are interdependent and should be analysed as a set. FRAM is one framework for interpretation and recognition of its combined significance. More work is needed regarding the use of these indicators in the decision-making context.

How is a specific operation carried out in the real world and what implications for indicators and models are available to improve the monitoring of safety?
The specific objective is to explore which indicators may be identified for the monitoring of “normal” daily operations. Most of the analyses focus on how the system should operate or has operated. The thesis contributes to two developments: 1) Indicators are related to factors essential to daily operations and 2) Models and analysis related to the analysis of situations where failures have not occurred.

6 Conclusions and recommendations for further research

The way safety is understood guides the identification of indicators. Resilience Engineering sees safety as something that the system does and not something that the system has. Complex socio-technical systems operate under many operational and financial constraints. Indicators are one source of information to act on and continue successful operations under varying conditions. For modern systems where accidents are rare, it is not possible to have a simple list of indicators. It is necessary to understand the operation of the system to be able to identify a set of indicators related to a specific context of operation. In relation to monitoring what is critical or to anticipate threats and opportunities, the indicators represent patterns of operation related to the variability of performance or its adaptive capacity. These indicators are closely related to production and regularity. They show how the system responds to cope with demands. The papers give examples of such indicators. Further work is needed to refine the indicators and provide guidance in their use in a systematic manner.

The objective of the thesis is to explore existing concepts and methods to address the question: *How do we identify that a system drifts or experiences sudden changes in the safety space?* When I started the thesis, I followed safety management approach as proposed by ICAO. It mainly associates drift and deviations with errors and failures, which is negative. Then this approach studies methods and measures that try to correct and minimize drift and deviations. Extensive research, industry practice both propose and use indicators related to factors associated with errors and failures. Due to my industrial experience and lessons learned during the thesis work, I adopted Resilience Engineering. This perspective proposes alternative ways to understand socio-technical systems. Resilience Engineering sees drift and deviations as manifestation of adaptations to cope with today´s complexity and conflicting goals. They are required to continue operations. The aim is to understand how the system operates and the impact of these adaptations. It is necessary to identify when the system becomes brittle and enhances resilience. So the system is prepared to deploy responses when exposed to disturbances. Then, methods and measures aim to manage processes in context. Indicators are related to the analysis of real cases and identification of which indicators are valid for the specific activity. Unfortunately, indicators have a "tunnel effect" focusing on some particular issues and obscuring others. So, indicators need to be revised and changed.

Rather than following one safety perspective, it is recommended to use a variety of methods based on different perspectives. Established methods explored during the thesis look at system failures that may be explained by the state of their components. Systemic methods like the Functional Resonance Analysis Method (FRAM) claim that the functioning of the system cannot solely be explained by the separate components. These methods look for dynamic and contextual descriptions of the system performance and encourage the analyst to look beyond a specific failure into the conditions of normal work. The storytelling method captures feelings and rich contextual data. These methods are considered to be complementary and can trigger a broader understanding of the system.

To identify indicators it is necessary to understand how operational constraints affect performance. Therefore, the main argument in this thesis is the identification of indicators related to failures and successful operations. For that purpose, it is reasonable to distinguish between three types of performance indicators: 1) lagging indicators, which refer to what has occurred in the past; 2) current indicators, which refer to what is occurring now; and 3) leading indicators, which refer to what may occur in the future. While, there is considerable data available for lagging indicators and disagreements about the indicator concept, a balanced composition of lagging, current and leading indicators is needed. Lagging indicators support action after the fact, whereas leading indicators support foresight and enable correction prior to an unwanted situation. When it comes to drift and sudden changes that affect operations and safety, the differentiation between leading and lagging is considered important. Since lagging indicators basically capture the outcome of a statistical process there is large uncertainty with respect to monitor drift and sudden changes. The aim of leading indicators is to measure factors and conditions that in the future will affect operations and safety. If leading indicators are based on context specific understanding and knowledge it is therefore believed that leading type of indicators will have higher strength in order to reveal trends and sudden changes. The studies demonstrate that it is possible to identify candidates for leading indicators related to successes and failures which then is believed to increase the capability of an organization to identify trends and more actively cope with changes and adaptations close to the safety boarders. These indicators enable the identification of issues that require improvements. The thesis contributes to more focus on everyday operations, leading indicators and analyses using systemic methods.

Further research could address the following:

- The studies presented in the thesis focused on the identification of indicators. The next step is the operationalization and utilization of indicators and their impact on performance in close cooperation with industry.

- A possible continuation is the implementation and use of safety indicators fitting decision-making processes within an organization and across organizations.

- Perform more studies in relation to everyday operations to identify trade-off and their impacts on performance, management practices and design of socio-technical systems.

- The Functional Resonance Analysis Method is under development and various needs are identified. One possible approach is to refine rules and provide better guidelines to analysts for the identification and evaluation of variability. Another is to refine a more systematic approach to the generation of recommendations in terms of barriers and indicators.

- Study a combination of RE methods as a basis for monitoring and managing complex socio-technical systems. Explore the meaning of Margin of Manoeuvre and Polycentric organizations in real industry cases

- Explore the impact of Resilience Engineering in relation to current safety management practices and establish cooperation with ICAO in further SMS developments.

- Explore the contribution of Resilience Engineering in systems design, regulations and monitoring performance in relation to future aviation developments towards more automation and implementation of advanced technologies.

While at the beginning of the helicopter study FRAM is used as a representative of Resilience Engineering. Today, I think that the study contributes to show how different perspectives represented using a variety of approaches (FRAM, Risk Influence Modelling and Storytelling) represent diversity providing a broader view that proposes different indicators.

References

Accident Investigation Board/Norway, (AIBN). (2005). *Safety in Norwegian Aviation during the Process of Change, Lillestrøm*, Norway.

ACARE (Advisory Council for Aeronautics Research in Europe), (2010). Aeronautics and Air Transport: Beyond Vision 2020 (Towards 2050).

Alderson, D. L., Doyle, J. C. (2010). Contrasting views of complexity and their implication for network-centric infrastructures. IEEE Systems, Man and Cybernetics, Part A, 40(4), 839-852.

Ale, B.J.M., Bellamy, L.J., Cooke, R.M., Goossens, L.H.J., Hale, A.R., Roelen, A.L.C., Smith, E. (2006). Towards a causal model for air transport safety – an ongoing research project. Safety Science (44), 657–673.

Ale, B.J.M., Bellamy, L.J., Van der Boom, R., Cooper, J., Cooke, R.M., Lin, P.H., Morales, O., Roelen, A.L.C., Spouge, J. (2008). Using a Causal Model for Air Transport Safety (CATS) for the Evaluation of Alternatives, ESREL2008, Valencia, Spain.

Allford, L., (2009). Process safety indicators: Response to Andrew Hopkins. *Safety Science* 47, 466.

Amalberti, R. (2001a). The Paradoxes of Almost Safe Transportation Systems, *Safety Science* 37, 109-126.

Amalberti, R. (2001b). Revisiting safety and human factors paradigms to meet the safety challenges of ultra complex and safe systems, In B. Willpert, & B. Falhbruch, Challenges and pitfalls of safety interventions, Elseiver : North Holland.

Andersen, H., Casa, J., Dandrieux, A., Debray, B., De Dianous, V., Duijim, N.J., Delvosalle, C., Fieves, C., Goossens, L., Gowland, R.T., Hale, A.J., Hourtolou, D., Mazzarotta, B., Pipart, A., Planas, E., Prats, F., Salvi, O., Tixitier, J., (2004) ARAMIS – User Guide, EC Contract number EVG1-CT-2001-00036

Ashby, W.R. (1981). Self-regulation and requisite variety. In F.E. Emery (ed.). *Systems Thinking*. Volume One. Harmondsworth: Penguin Education, 100-120. Earlier published as Chapter 11 in W.R.

Ashby, W.R. (1956). *Introduction to Cybernetics*, Wiley.

Australian Transport Safety Bureau (ATSB), (2005). *Aviation Safety Indicators. A report on safety indicators relating to Australian Aviation*. Aviation Research Investigation Report B2005/0046.
http://www.atsb.gov.au/publications/2005/safety_indicators.aspx

Aven, T., Sklet, S., Vinnem, J.E. (2005). Barrier and operational risk analysis of hydrocarbon releases (BORA-Release); Part I Method description. Journal of hazardous materials

Baker, III J., A, Leveson N., Bowman, F., L., Priest, S., Erwin, G., Rosenthal, I., Gorton, S., Tebo, P.,V., Hendershot, D, Wiegmann, D.,A, Wilson, L., D., (2007). *The Report of the BP U.S. Refineries Independent Safety Review Panel*.

Bakthin, M. (1986). *Speech genres and other late essays*. Austin, TX: University of Texas Press.

Beer, S. (1985). *Diagnosing the System for Organizations*. New York: John Wiley.

Behm, M., Veltri, A., Kleinsorge, I. (2004). Analyzing the Cost of Safety. Professional Safety – *Journal of the American Society of Safety Engineers*. 49:4, 22-29.

Bellamy, L.J., Papazoglou, I.A., Hale, A.R., Aneziris, O.N., Ale, B.J.M., Morris, M.I., Oh, J.I.H. (1999). I-Risk: Development of an integrated technical and management risk control and monitoring methodology for managing and quantifying on-site and off-site risks. Contract ENVA-CT96-0243. Report to European Union. Ministry of Social Affairs and Employment. Den Haag.

Bourrier, M. (1996). Organizing maintenance work at two American nuclear plants. *Journal of Contingences and Crisis Management* 4, 104-106.

Bourrier, M. (2002). Bridging Research and Practice: The Challenge of `Normal Operations Studies. . *Journal of Contingences and Crisis Management* 10, 173-180.

Borys, D., Else, D., Legget, S. (2009). The fifth age of safety: the adaptive age. *Journal of health & safety research & practice* 1-1, 19-27.

Civil Aviation Authority (CAA-UK), (2003). *Aviation Maintenance Human Factors* CAP 716 (EASA / JAR 145). http://www.caa.co.uk/docs/33/CAP716.pdf

Davoudian, K., Wu, J.-S., Apostolakis, G. (1994a). Incorporating organizational factors into risk assessment through the analysis of work processes. Reliability Engineering & System Safety 45, 85–105.

Davoudian, K., Wu, J.-S., Apostolakis, G. (1994b). The work process analysis model (WPAM). Reliability Engineering & System Safety 45, 107–125.

Dekker, S.W.A. (2004). *Ten questions about human error: A new view of human factors and system safety*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc

Dekker, S.W.A. (2005). *Why we need new accident models*. Technical Report 2005-02, Lund University School of Aviation. Sweden. http://www.lu.se/upload/Trafikflyghogskolan/TR2005-02_NewAccidentModels.pdf

Dekker, S.W.A. (2006). Resilience engineering: Chronicling the emergence of confused consensus. In E. Hollnagel, D. D. Woods & N. Leveson (Eds.), *Resilience Engineering: Concepts and precepts*. Aldershot, UK: Ashgate

Dekker, S.W.A. (2011). *Drift into failure: From Broken Components to Understanding Complex Systems*. Ashgate Publishing, UK.

Department of Energy (DOE), (2005). *Implementation guide, aviation program, performance indicators (metrics)*. https://www.directives.doe.gov/directives/current-directives/440.2b-EGuide-1a/view
DHJIT (Deepwater Horizon Incident Joint Investigation), 2010. Marine Board of Investigation into the Marine Casualty, Explosion, Fire, Pollution, and Sinking of Mobile Offshore Drilling unit Deepwater Horizon, with the Loss of Life in the Gulf of Mexico April 21–27, 2010. The US Coast Guard (USCG/ Bureau of Ocean Energy Management, Regulation and Enforcement (BOEMRE) Joint Investigation Team (JIT).

Doyle, J.C. (2000). Multiscale networking, robustness and rigor. In T. Samad & J- Weyrauch (Eds.) *Automation, control and complexity: an integrated approach*. NY: John Wiley & Sons, Inc. New York, 287-301

European Aviation Safety Agency (EASA), (2009). *Annual safety review 2010*.
http://easa.europa.eu/communications/docs/annual-safety-review/2010/EASA-Annual-Safety-Review-2010.pdf

Eisenhardt, K.E.(1989). *Building Theories from Case Study Research*. The academy of Management Review, Vol.14, No.4, pp.532-550. Stanford University. USA.

Electrical Power Research Institute (EPRI), (2000). "Guidelines for Trial Use of Leading Indicators of Human Performance: The Human Performance Assistance Package". EPRI (U.S. Electric Power Research Institute), Palo Alto, CA, 1000647.

Electrical Power Research Institute (EPRI), (2001). "Final report on Leading Indicators of Human Performance". EPRI, Palo Alto, CA, and the U.S. Department of Energy, Washington, DC, 1003033. e

EUROCONTROL, (2009). *Improving European ATM safety through SMART safety indicators*. 3rd SAFRED TF Report to Provisional Council. European ATM safety performance indicators.

Fitts, P.M (1951). *Human Engineering for an Effective Air-Navigation and Air traffic-Control System*. National Research Council. Washington, D.C.

Foster, H.D. (1993). Resilience theory and system evaluation. In J.A. Wise, V. D. Hopkin and P. Stager (eds).*Verification and Validation of Complex Systems: Human Factors Issues*. Berlin: Springer

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry* 12 - 2. 219-245. Sage Publications.

Gabriel, Y. (2000). *Storytelling in organizations*. Oxford: Oxford University Press.

Gibson, J. J. (1961). The contribution of experimental psychology to the formulation of the problem of safety – a brief for basic research. In Behavioral Approaches to Accident Research, New York: Association for the Aid of Crippled Children, pp. 77-89. Reprinted in W. Haddon

Gosling, G.D. (1998). *Development of system safety performance measures in support of the global analysis and information network.* Working paper UCB-ITS-WP-98-3. Institute of Transportation Studies. University of California at Berkeley. USA. Available from: http://www.escholarship.org/uc/item/9n39p3fm

Guba, E.G., Lincoln, Y.S. (1981). *Effective evaluation: Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic* Approaches. Jossey-Bass. San Francisco.

Gunderson, L.H., Holling, C.S. (2002). *Panarchy Understanding Transformations in Human and Natural systems*. Island Press. USA.

Haddon-Cave, C. (2009). *The Loss of RAF NIMROD XV230. A failure of leadership, culture and priorities.* ISBN: 9780102962659. Published by The Stationary Office.
http://www.official-documents.gov.uk/document/hc0809/hc10/1025/1025.asp

Haddon, W. (1970). On the escape tigers: An ecologic note. *Technology Review*. Vol 72 no. 7. Massachusetts Institute of Technology.

Haddon, W. (1973). *Energy damage and the ten countermeasure strategies*. The journal of trauma. Vol 13. No. 4. Pp 321-331.

Hale, A.R., Heming, B.H:J., Carthey, J., Kriwan, B. (1997). Modelling of safety management systems. *Safety Science* 26. No. 1/2, 121-140

Hale, A., Baram, M., and Hovden, J. (1998). Perspectives on Safety Management and Change. In: Hale, A., and Baram, M., (Eds), *Safety Management: The Challenge of Change*. Pergamon

Hale, A. R. and Hovden, J. (1998). "Management and culture: The third age of safety". In A.-M. Feyer & A. Williamson (Eds.), *Occupational injury. Risk, prevention and intervention*. London: Taylor & Francis.

Hale, A., Heijer, T. (2006). Defining Resilience. In: E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts*. Aldershot: Ashgate Publishing

Hale, A. (2009a). Editorial. Special Issue on Process Safety Indicators. *Safety Science* 47, 459.

Hale, A. (2009b). Why safety performance indicators? *Safety Science* 47, 479-480.

Health and Safety Executive (HSE), (2006). *Developing process safety indicators: a step-by-step guide for chemical and major hazard industries*.

Heinrich, H. W. (1931). *Industrial accident prevention*. McGraw-Hill: New York.

Helmereich, R.L., Klinect, J.R., Wilhem, J.A. (2003). Managing Threat and Error: Data from Line Operations. In: Edkins, G., Pfister, P, editors " *Innovation and consolidation in aviation*". Ashgate, Aldershot

Hendrick, K., Benner, L. (1987). *Investigating accidents with STEP*. Marcel Dekker Inc. New York.

Herrera, I.A., Håbrekke, S., Kråkenes, T., Hokstad, P., Forseth, U. (2010). *Helicopter Safety Study 3*. Main report. SINTEF report no. A14973, Trondheim, Norway.

Hitchins, D.K. (2004). *Putting Systems to work*.
http://www.hitchins.net/Books&samplers/e-Putting%20SystemsToWork.pdf

Hokstad,P., Jersin, E., Sten, T. (2001). A risk influence model applied to North Sea helicopter transport. *Reliability Engineering and System Safety 74*: 311–322.

Hollnagel, E. (2004). *Barriers and accident prevention*. Aldershot, Ashgate.

Hollnagel, E., Woods, D.D., Leveson, N., (2006). Resilience Engineering: Concepts and Precepts. Ashgate, Aldershot.

Hollnagel, E. (2008a) "Preface" and "Safety management, looking back or looking forward" In: Hollnagel E, Nemeth CP, Dekker S, editors. In *Resilient Engineering Perspectives, vol. 1, Remaining Sensitive to the Possibility of Failure*. Ashgate, Aldershot, USA.

Hollnagel, E. (2008b). The changing nature of risks. *Ergonomics Australia Journal* 22, 1-2, 33-46

Hollnagel, E. (2009). *The ETTO Principle: Efficiency-Thoroughness Trade-Off. Why Things That Go Right Sometimes Go Wrong*. Ashgate, UK.

Hollnagel, E. (2009). The four cornerstones of resilience engineering. In: Nemeth C., Hollnagel E. and Dekker S. (Eds.), *Resilience Engineering Perspectives, vol. 2, Preparation and Restoration*. Ashgate, Aldershot, UK.

Hollnagel, E. (2011). Prologue: The scope of Resilience Engineering and Epilogue: RAG – The Resilience Analysis Grid. In: Hollnagel E., Pariès, J., Woods, D.D., and Wreathall, J. (Eds.), *Resilience Engineering in Practice: A Guidebook*. Ashgate, Aldershot, UK

Holmberg, J., Laakso, K., Lehtinen, E., Johanson, G., (1994). *Safety evaluation by living probabilistic safety assessment and safety indicators*. TemaNord 1994:614. Copenhagen, Denmark: The Nordic Council of Ministers.

Hopkins, A. (2000). *Lessons from Longford. The ESSO Gas Plant Explosion*. CHC Australia Limited

Hopkins, A. (2008). *Failure to learn: the BP Texas City refinery disaster*. CCC Australia Limited.

Hopkins, A. (2009a). Thinking about process safety indicators. *Safety Science* 47, 460-465.

Hopkins, A. (2009b). Reply to comments. *Safety Science* 47, 508-510

International Atomic Energy Agency (IAEA), (1986). *Summary report on the post-accident review meeting on the Chernobyl accident* (IAEA safety series No. 75-INSAG-1). Vienna: International Atomic Energy Agency.

International Civil Aviation Organization (ICAO), (2006). *Worldwide and regional trends in aviation safety. Working Paper* DGCA/06-WP/2. Montreal, Canada.

International Civil Aviation Administration (ICAO) (2009). *Safety Management Manual* (SMM). Doc. No. 9859, AN/474. http://www2.icao.int/en/ism/ Guidance%20Materials/DOC_9859_FULL_EN.pdf

Kinnison, H.A. (2004). Maintenance Management. McGraw-Hill. New York.

Kjellén, U. (2000). *Prevention of accidents through experience feedback*. Taylor & Francis, London.

Kjellén, U. (2009). The safety measurement problem revisited. *Safety Science* 47, 486-489.

LaPorte, T. R. and Consolini, P.M. (1991). Working in practice but not in theory: Theoretical challenges of "High-Reliability Organisations". *Journal of Public Administration Research and Theory*, 1, 19-47.

Leveson, N.G., (2001). *Evaluating Accident Models using Recent Aerospace Accidents*, Technical report, MIT Dept. of Aeronautics and Astronautics. Available at: http://sunnyday.mit.edu/accidents/nasa-report.pdf

Leveson. N. (2002) A Systems Model of Accidents, International Conference of the System Safety Society, Denver.

Leveson, N., (2002) Model-Based Analysis of Socio-Technical Risk. Technical Report, Engineering Systems Division, Massachusetts Institute of Technology

Leveson, N., (2004). Model-based analysis of socio-technical risk. Massachusetts Institute of Technology ESD-WP-2004-08

Leveson, N. (2004). A New Accident Model for Engineering Safer Systems. *Safety Science 42*, 237-270

Leveson, N. (2011). Managing Process Safety: Lessons from Deepwater Horizon. IO conference. Trondheim. Norway

Leveson, N. (2011). Engineering a safer world. Draft book available at: http://sunnyday.mit.edu/safer-world/

Leveson, N., Dulac, N., Zipkin, D., Cutcher-Gershenfeld, Carrol, J., Barret, B., (2006) Engineering Resilience into Safety-Critical Systems. In: E. Hollnagel, D. D. Woods, & N. Leveson (Eds.), *Resilience Engineering: Concepts and Precepts*. Aldershot: Ashgate Publishing

Lincoln Y.S., Guba, E.G. (1985). *Naturalistic inquiry*. Sage Publications, Inc. Newbury Park.

Linköping University (LiU), (2007) *Fram Visualizer*. http://code.google.com/p/framvisualizer/

Mohaghegh, Z. (2010): Development of an Aviation Safety Causal Model Using Socio-Technical Risk Analysis (SoTeRiA). 10th International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSAM10).

Mohaghegh, Z., Mosleh, A. (2009). Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: Principles and theoretical foundations. Safety Science 47 (2009) 1139–1158.

Macchi, L. (2010). *A Resilience Engineering approach to the evaluation of performance variability: development and application of the Functional Resonance Analysis Method for Air Traffic Management safety assessment*. Ph.D. Thesis. École Nationale Supérieure des Mines de Paris, France.

Maurino, D.E., Reason, J.T., Johnston, A.N., Lee, R.B. (1995). *Beyond Aviation: Human Factors*. Aldershot: Avebury Aviation.

McDonald, N. (2006). Organizational resilience and industrial risk. In: Hollnagel, E., Woods, D., Leveson N. (Eds.) *Resilience Engineering: Concepts and Precepts*. Ashgate, England.

McDonald, N., Ward, M., Morrison, R. (2012) Achieving Impact in Ergonomic Research. Paper presented during18th World Congress on Ergonomics. Recife Brasil.

Miller, R. (1953). A method for man-machine task analysis.

National Transportation Safety Board (2002). *Loss of control and impact with Pacific Ocean, Alaska Airlines Flight 261 McDonnell Douglas MD-83, N963AS, about 2.7 miles north of Anacapa Island, California, January 31, 2000 (AAR-02/01)*. Washington, D.C.: NTSB

Organization for Economic Co-operation and Development (OECD) (2002) *Frascati Manual*, Sixth edition, 2002, http://www.oecd.org/document/6/0,3343,en_2649_34451_33828550_1_1_1_1,00.html

Organization for Economic Co-operation and Development (OECD), (2005). *Guidance on safety performance indicators. Guidance for industry, public authorities and communities for developing SPI programmes related to chemical accident prevention, preparedness and response*. OECD Environment, Health and Safety Publications. Series on Chemical Accidents 11.

Organization for Economic Co-operation and Development (OECD), (2005). *Glossary of statistical terms*. http://stats.oecd.org/glossary/download.asp

Page, S.E. (2010). *Diversity and Complexity.* Princenton, NJ: Princenton University Press.

Pantakar, M., Taylor, J. (2004). *Risk Management and Error Reduction in Aviation Maintenance*. Aldershoot, Ashgate Publishing Limited.

Papazoglou, I.A., Bellamy, L.J., Hale, A.R., Aneziris, O.N., Ale, B.J.M., Post, J.G., Oh, J.I.H. (2003). I-Risk: development of an integrated technical and management risk methodology for chemical installations. Safety Science 16, 757-591

Patton, Q.M., (2002). *Qualitative Research & Evaluation Methods*. 3rd ed. Sage Publications

Perin, C. (2005). *Shouldering risk. The culture of control in the nuclear power industry*. Princenton University Press. New Jersey

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies*. New York: Basic Books.

Pettersen, K.A. (2008). The social production of safety. Theorising the human role in aircraft line maintenance. *PhD thesis. University of Stavanger*, Norway

Pidgeon, N.;O'Leary, M. (2000). Man-made disasters: why technology and organizations (sometimes) fail. *Safety Science* 34. Pp 15-30

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science*, 27, 183-213.

Rasmussen, J. and I. Svedung (2000). *Proactive Risk Management in a Dynamic Society* (Swedish Rescue Services Agency, Karlstad, Sweden).

Rasmussen, J. (2000). Human factors in a dynamic information society: where are we heading? *Ergonomics*, 43, 869-879.

Reason, J. (1990). *Human error*. Cambridge: Cambridge University Press.

Reason, J. (1997). *Managing the Risk of Organizational Accidents*, Ashgate, Alsershot, USA

Reason, J.T., Hobbs, A. (2003). *Managing Maintenance Error*. Aldershot , UK: Ashgate.

Reason, J., Hollnagel, E., Paries, J., (2006). *Revisiting the "Swiss Cheese" model of accidents*. Eurocontrol EEC Note No. 13/06

Reason, J. (2008). *The Human Contribution. Unsafe Acts, Accidents and Heroic Recoveries*. Ashgate Publishing, Aldershot.

Reiman, T. & Oedewald, P. (2009). *Evaluating safety critical organizations. Focus on the nuclear industry*. Swedish Radiation Safety Authority, Research Report 2009:12.

Reiman, T., Pietikäinen, E. (2010). *Indicators of safety culture. Selection and utilization of leading safety performance indicators*. Swedish Radiation Safety Authority, Research Report 2010:07.

Reiman, T., Rollenhagen, C. (2010). Identifying the typical biases and their significance in the current safety management approaches. *Proceedings of the 10th International probabilistic safety assessment and management conference*. Seatle. USA.

Reiman, T. (2010). Understanding maintenance work in safety-critical organizations – managing the performance variability. *Theoretical Issues in Ergonomics Science*, In press.

Renn, O. (2008). Essay 1 A guide to interdisciplinary risk research. *Risk Governance: coping with uncertainty in a complex world*. Earthscan. UK.

Research Council of Norway (Norges forskningsråd) (2000). *Kvalitet I norsk forskning. En oversikt over begreper, metoder og virkemidler*. ISBN 82-12-01361-8. Oslo, Norway.

Rigaud, E., Guarnieri, F. (2006). Proposition of a conceptual and a methodological modelling framework for resilience engineering. *Proceedings on the second Resilience Engineering Symposium*. Juan Les Pins, France. http://www.resilience-engineering.org/REPapers/Rigaud_Guarnieri_R.pdf

Ringdal, K. (2001). *Enhet og mangfold: samfunnsvitenskapelig forskning og kvantitativ metode*. In Norwegian [Unity and diversity: social science and quantitative methods] Bergen, Fagbokforlaget

Rollenhagen C., Andersson O. (1999). Experience from the MTO Programme at the Forsmark NPP", A paper presented at ENS TopOperation conference, Berlin.

Rosness, R., Guttormsen, G., Steiro, T., Tinmannsvik, R. K. and Herrera, I. A. (2004). *Organisational accidents and resilient organisations: five perspectves*. Trondheim, SINTEF, Industrial Management, Safety and Reliability. SINTEF report no. STF38 A04403

Shappell, S.A., Wiegmann, D.A., (2000). *The Human Factors Analysis and Classification System - HFACS.* FAA. US Department of Transportation

Silverman, D. (2006). *Interpreting qualitative data: methods for analyzing talk, text and interaction*. London, Sage.

SINTEF, Ingstad, O., Rosness, R., Sten, T., Ulleberg, T., Rausand, M. & Lydersen, S. 1990. Helicopter Safety Study. Main Report. SINTEF Report no. STF75 A90008.

SINTEF, P. Hokstad, E. Jersin, G. K. Hansen, J. Sneltvedt and T. Sten (1999). Helicopter Safety Study 2. SINTEF Report no. STF38 A99423

SINTEF, I. A. Herrera, S. Håbrekke, T. Kråkenes and P. Hokstad (2010). Helicopter Safety Study 3. SINTEF Report no. STFA14973.

Sklet, S., Vinnem, J.E., Aven, T. (2006). Barrier and operational risk analysis of hydrocarbon releases (BORA-Release). Part II: Results from a case study. Journal of Hazardous materials submitted in 2005.

Snook, S.A. (2000*). Friendly fire: The accidental shootdown of U.S. Black Hawks over northern Iraq*. Princenton, NJ:Princenton University Press.

Step-Change in Safety (2001). Leading performance indicators: a guide for effective use . [http://www.stepchangeinsafety.net/stepchange/News/StreamContentPart.aspx?ID=1517]

Stroeve, S.H., Sharpanskykh, A., Kirwan, B. (2011) Agent-based organizational modelling for analysis of safety culture at an air navigation service provider. Safety Science 96. Pp 515-533

Størseth, F, Tinmannsvik, R., K., Øien, K. (2009). "Building safety by resilient organization – a case specific approach". European Safety and Reliability Association Annual Conference (ESREL), 7 – 10 September, Prague, Czech Republic.

Swain, A. D., Guttman, H. E. (1983). *Handbook of human reliability analysis with emphasis on nuclear power plant applications (NUREG CR 1278)*. Washington, DC NRC.

Tarrants, W. E. (1980). *The Measurement of Safety Performance.* New York, USA.

Task Group on Regulatory Effectiveness (TGRE), (2004). *Direct indicators of nuclear regulatory efficiency and effectiveness*.

Thagaard, T. (2003). *Systematikk og innlevelse: en innføring i kvalitativ metode*. In Norwegian [Systematic and insight: introduction to qualitative methods] Bergen, Fagbokforl.

Tinmannsvik, R.K. (2005). Ytelsesindikatorer for flysikkerhet - noen resultater fra svensk luftfart, (Performance based indicators for flight safety - some results from Swedish aviation). SINTEF, Trondheim, Norway

Turner, B. A. (1978). *Man-made disasters*. London: Wykeham Science Press.
Turner, B. A., Pidgeon, N. F. (1997). *Man-made disasters. 2nd Editi*on. London: Butterworth-Heinemann.

Yin, R.K. (2009). *Case study research: design and methods*. (4th ed.) Thousand Oaks, CA: Sage

Van Steen, J. (1996). *Safety Performance Measurement.* Institution of Chemical Engineers U.K

Vaughan, D. (1996). *The Challenger launch decision*. Chicago: University of Chicago Press

Vinnem, J.E., et al. 2007. Operational safety condition – concept development. ESREL 2007, Stavanger, 25–27 June, 2007.

Vinnem, J.E., 2008. On Causes and Dependencies of Errors in Human and Organizational Barriers Against Major Accidents. Paper Presented at ESREL 2008, Valencia, 22–25 September, 2008

Weingberg, G.M., Weinberg, D. (1979). *On the design of stable systems*. New York: Wiley.

Wilpert, B., Fahlbruch, B. (1998). Safety related interventions in inter-organisational fields. In A. Hale & M. Baram (Eds), *Safety Management: The Challenge of Change*. Pergamon

Webb, P., 2009. Process safety performance indicators: A contribution to the debate. *Safety Science* 47, 502-507.

Weick, K.E. (2001). *Making sense of the organization*. Oxford: Blackwell.

Weick, K.E. and Sutcliffe, K.M. (2001). *Managing the Unexpected*. San Francisco: Jossey-Bass.

Weick, K.E. & Sutcliffe, K.M. (2007). *Managing the unexpected: Resilient Performance in an Age of Uncertainty*. San Francisco: Jossey-Bass.

Weick, K.E. (2009). *Making sense of the organization: the impermanent organization*. John Wiley & Sons.

Westrum, R. (1993). Cultures with Requisite Imagination. In J.A. Wise, V. D. Hopkin and P. Stager (eds). *Verification and Validation of Complex Systems: Human Factors Issues*. Berlin: Springer, 401-416.

Widalsky, A. (1988). Searching for Safety. New Brunswick. CT: Transaction Books.

Woltjer, R. (2009). Functional modeling of constraint management in aviation safety and command and control. Linköping Studies in Science and Technology, Dissertation No. 1249, Linköping University, Sweden.

Woods, D. D. (2003). *Creating foresight: How resilience engineering can transform NASA's approach to risky decision making.* Testimony on the future of NASA for the Committee on Commerce, Science and Transportation, John McCain, Chair. Washington, D.C.
http://ergonomics.osu.edu/pdfs/Press%20Releases/Press%20Release%20Oct03-Creating%20Foresight.pdf

Woods, D.D. (2010). *Creating Safety by Engineering Resilience or On Being Resilient in the Age of Complexity*. Key note during the 41st Anural Conference of the Association of Canadian Ergonomist. Available from:
http://www.ace-ergocanada.ca/lib/file.php?file=/ACE_2010/keynotes/keynote-woods.pdf

Woods, D.D. (2011). Resilience and the ability to anticipate. In: Hollnagel E., Pariès, J., Woods, D.D., and Wreathall, J. (Eds.), *Resilience Engineering in Practice: A Guidebook*. Ashgate, Aldershot, UK

Woods, D.D., Wreathall, J. (2003). Managing Risk Proactively: The mergence of Resilience Engineering.
http://csel.eng.ohio-state.edu/woods/error/working%20descript%20res%20eng.pdf

Woods, D.D., Dekker, S., Cook, R., Johannesen, L., Sarter, N. (2010). *Behind Human Error*. Second Edition. Ashgate. UK.

Woods, D.D., Branlat, M. (2011). In: Hollnagel E., Pariès, J., Woods, D.D., and Wreathall, J. (Eds.), *Resilience Engineering in Practice: A Guidebook*. Ashgate, Aldershot, UK

Wreathall, J. (2006). Property of Resilient Organization: An Initial View, Resilience Engineering In: Hollnagel, E., Woods, D., Leveson N. (Eds.) *Resilience Engineering: Concepts and Precepts*. Ashgate, England.

Wreathall, J. (2007). Some rumblings on a framework for identifying metrics and their use in resilience. Handed out at Resilient Risk Management Course, Juan les Pins, France.

Wreathall, J. (2009). Leading? Lagging? Whatever! Safety Science 47, 493-494.

Yin, R.K. (2009). *Case Study Research: Design and Methods*. Sage Publications, Inc. California

Øien, K, Massaiu, S, Tinmannsvik, R.K., Størseth, F. (2010). Proceedings of the 10th International Probabilistic Safety Assessment and Management Conference (PSAM), Seattle, USA

Øien, K., Utne, I.B., Herrera, I.A., (2009). Building Safety Indicators. Part 1 - Theoretical background. Safety Science.

Øien, K.(2001). *A Focused Literature Review of Organizational Factors' Effect on Risk, in Risk Control of Offshore Installations: A framework for the establishment of organizational risk indicators*. Doctoral Thesis. The Norwegian University of Science and Technology.

Appendix A. Acronyms and abbreviations

AIBN	Accident Investigation Board, Norway
ATS	Air Transport System
ATSB	Australian Transport Safety Bureau
BBN	Bayesian Belief Network
CAA	Civil Aviation Authority
CATS	Causal Model for Air Transport Safety
CFIT	Controlled Flight into Terrain
CMA	Continuous Monitoring Approach
CRM	Crew Resource Management
CSF	Contributing Success Factors
DOE	Department of Energy
DRM	Dynamic Risk Modelling
EC	European Commission
EASA	European Aviation Safety Agency
EPRI	Electrical Power Research Institute
ESD	Event Sequence Diagram
FDM	Flight Data Monitoring
FRAM	Functional Resonance Analysis Method
FTA	Fault Tree Analysis
HAZOP	Hazard and Operability Analysis
HFACS	Human Factors and Classification Systems
HCI	Hybrid Causal Logic
HSE	Health and Safety Executive
HSS	Helicopter Safety Study
HRO	High Reliability Organizations
HUMS	Health Usage Monitoring System
IAEA	International Atomic Energy Agency
ICAO	International Civil Aviation Organization
I-RISK	Integrated Risk
LIOH	Leading Indicators of Organizational Health
LOFT	Line Oriented Flight Training
LOSA	Line Operations Safety Audit
MESH	Management Engineering and Safety Health
METAR	Meteorological observation message for routine aviation
MTO	Man, Technology and Organozation
NAT	Normal Accident Theory
NOSS	Normal Operations Safety Survey
NTSB	National Transport Safety Board
OECD	Organization for Economic Co-operation and Development
OLF	Norwegian Oil Industry Association
RAG	Resilience Analysis Grid
RE	Resilience Engineering
REWI	Resilience Early Warning Indicator
RIF	Risk Influence Factors
SADT	Structured Analysis and Design Technique
SD	System Dynamics
SMS	Safety Management System
SoTeRiA	Socio Technical Risk Analysis
STAMP	Systems-Theoretic Accident Modelling and Process
STEP	Sequentially Timed Events Plotting Methods

STPA	STAMP HazardAnalysis
TAF	Terminal Aerodrome Forecast
THERP	Technique for Human Error Rate Prediction
TOPAZ	Traffic Organization and Perturbation AnalyZer
WPAM	Work Process Analysis Model

Appendix B. Terms and definitions

Air Transport System
The system of systems within the broad aviation sector including: civil aviation authorities (CAA), air traffic management (ATM), flight operations (FO), aircraft maintenance (MX), original equipment manufacturers (OEM), and airport services (e.g. ground services, baggage handlers, fuel suppliers, catering etc.)

Brittleness
A brittle system collapses suddenly when confronted with a disturbance or event that challenges the system's boundaries or assumptions (key example is the Ariane 501 failure). A resilient system exhibits graceful degradation in the face of potential challenge events by actively monitoring for and anticipating new challenge events, by deploying or mobilizing new responses to preserve key system functions in the face of challenge events, and by learning improved approaches for adapting to potential challenges (Alderson and Doyle 2010; Hollnagel et al., 2006; 2011)

Drift
ICAO´s SMS (2009) uses the term practical drift. This term is related to a gradual drift from a baseline design performance towards an operational performance. The baseline performance is determined by the initial design of the system. This design is based on assumptions regarding the technology needed, the way the technology is operated, the regulations and procedures. The operational performance is the collective performance taking into account that the technology does not always operate as predicted, the procedures cannot always be executed as planned due to dynamic operational conditions and regulations do not always take into account contextual limitations. ICAO states that it is necessary to capture the practical drift to learn about successful adaptation or to develop strategies to prevent that this practical drift develops to far that an unwanted event materializes in form of incident or accident.

Emergence
How a system's properties and behaviour arise from the relationships and interactions across parts, and not from the individual parts in isolation or properties of components.

Functional resonance
The variability of individual functions may combine in an unwanted and unexpected way. This is the result of functional couplings in the system. Any part of the system variability can be a "signal" and the "noise" is determined by the variability of the functions in the system. Thus the variability of a number of functions may resonate, i.e., reinforce each other and thereby cause the variability of one function to exceed normal limits.

Indicator
Its origin corresponds to the 'one who points out.' It is related to Latin indicare or 'to point out.' The term indicator has been used in several ways by the safety community, which means that there are many definitions (Øien et al., 2010). See safety indicator for the definition adopted in the thesis.

Instantiation
In the FRAM modelling this term is used to describe a set of couplings among functions for specific time intervals (Herrera et al., 2010)

Intractable
A system which cannot be described in every detail and where the functioning and therefore is not completely understood. Intractable systems are only partly predictable.

Model
It is a representation of something else, of phenomenon or event such an accident or of a system such as an organization (Reason et al., 2006).

- Retrospective model is the basis for explaining or understanding something
- Prospective model is the basis for predicting something, including measurements of present states as an indicator of possible future states.

Performance variability
It relates to the ways in which individual and collective performances are adjusted to match current demands and resources, in order to ensure that things go right.

Predictive methods
ICAO (2009) relates it to routine operational data captured in real time. It is based on the notion to trying to find trouble, not just waiting for it to show up. Therefore, indicators are related to potential problems, emerging safety risks from a variety of sources.

Proactive methods
ICAO (2009) considers less serious events related to proactive monitoring with little or no damaging consequences. This monitoring is based on the notion that system failures can be minimized before the system fails. Examples of indicators are related to mandatory and voluntary reporting systems, safety audits and safety surveys.

In this thesis predictive and proactive monitoring relates to the present state. It is proposed that predictive and proactive monitoring support a feed forward strategy for safety management. The proactive focuses on indication related to the current state while the predictive focuses on the potential for future problems or opportunities.

Reactive methods
It relates to the identification, reporting and investigation of incidents and accidents (HSE, 2006; Baker, 2007). It provides a feedback in relation to safety performance and allows the identification of deficiencies related to specific incidents or trends. It relates to an event with considerable damaging consequences. Monitoring based on reactive methods uses the notion of waiting until “something breaks to fix it”. Indicators related to accidents and serious incidents are examples of reactive navigation aids (ICAO, 2009).

Resilience
The definition applied in the thesis is “*the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions*” (Hollnagel et al., 2011).

Resonance
It is proposed as a principle that explains how disproportionate large consequences can arise from seemingly small variations in performance and conditions.

Safety
The state in which the possibility of harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and safety risk management (ICAO, 2009)

Using an analogy between safety and fitness, it is possible to understand safety as the ability of the organization to function efficiently so it can meet its objectives.

Resilience Engineering sees safety as "the ability to succeed under varying conditions (Hollnagel, 2011)

Safety indicator
It is an observable characteristic of an operational unit, presumed to bear a positive correlation with the safety of the system (based on Holmberg, 1994)

Wreathall (2009) defines indicator as "*proxy measures for items identified as important in the underlying model(s) of safety. As such they are uncertain and often only distantly connected to the idealized measures that rarely available in practice*"

Safety Management System
A Safety Management System (SMS) is a systematic approach to manage safety, including the necessary organizational structures, accountabilities, policies and procedures (ICAO). ICAO through various Annexes to the Chicago Convention has incorporated requirements for service providers in various domains of aviation to have an SMS.

Safety Risk
It is an the assessment, expressed in terms of predicted probability and severity, of the consequences of a hazard, taking as reference the worst foreseeable situation ICAO (2009).

Serious Incident
An incident involving circumstances indicating that an accident nearly occurred (ICAO Annex 13 Chapter1)

Socio-technical system
The term refers to a technical system where people who operate and maintain the system to a great extent influence the effectiveness of the system. The efficiency of the technology is therefore largely dependent on the people who operate and maintain it, and there is a complex interaction between people and technology (HSE, 2002).

Appendix C. Additional work

The following conference papers and publications were produced during the PhD work period. These publications are not included in this thesis:

- Hansson, L., Herrera I.A., Konsvik, T. (2008). *Applying the resilience concept in practice: A case study from the oil and gas industry*. Proceedings of the European Safety and Reliability Association Annual Conference (ESREL), Valencia, Spain
- Herrera, I.A., Håbrekke, S., Kråkenes, T., Hokstad, P., Forseth, U. (2010). *Helicopter Safety. Study 3. Main report*. SINTEF report no. A14973, Trondheim, Norway
- Herrera, I.A., Hollnagel, E., Macchi, L., Woltjer, R. (2010). *Exploring Resilience Engineering Contribution to Risk Analysis in Air Traffic Management*. EUROCONTROL Safety and Human Factors Seminar. Brétigny-sur-Orge, France
- Rosness, R., Grøtan, T.O., Guttormesen, G., Herrera, I.A., Steiro, T., Størseth, F., Tinmannsvik, R., Wærø, I. (2010). *Organisational Accidents and Resilient Organisations: Six Perspectives. Revision 2.* SINTEF. Trondheim. Norway.

International workshop presentations:
- Herrera, I. A., Woltjer, R. (2008). *A comparison of the FRAM and STEP models in the aviation domain.* FRAM Workshop. Sophia Antipolis, France
- Herrera, I.A., Tveiten, C. (2008). *Modelling normal performance variability with FRAM and risk analysis with FRAM*. FRAM Workshop. Sophia Antipolis, France
- Herrera, I.A., Macchi, L. (2009). *What to look for when everything is normal*. FRAM Workshop. Sophia Antipolis, France
- Herrera, I.A., Macchi, L. (2009). *What to look for when everything is normal*. Joint VTT-NTNU/SINTEF-MINES ParisTech, Sophia Antipolis, France
- Herrera, I.A. (2010). *Presentation of recent research: Indicators and Aviation*. Workshop on indicators and evaluation of safety; Joint VTT-NTNU/SINTEF-MINES ParisTech, Helsinki, Finland

Other international conferences and courses
- Participation in Resilience Engineering Workshop. June 2007. Vadstena, Sweden
- Participation in the Resilient Risk Management Course. November 2007. Juan les Pins, France
- Preparation and presentation. A practical introduction to Functional Resonance Analysis Method (2008), Trondheim, Norway.
- Lecture on Safety Performance Indicators (2009). PhD course NTNU-Mines Paris Tech, Sophia Antipolis, France.
- Preparation and presentation. A practical introduction to Functional Resonance Analysis Method as event analysis (2010). Trondheim, Norway.
- Lectures on Breaking the myth of human error in maintenance or Human factors in maintenance 2008, 2009, 2010
- Lecture on “The Functional Analysis Method, An alternative Method” for Eni Norway as part of a course on accident investigation.

National workshops:
- Herrera, I. A., Woltjer, R. (2008). STEP & FRAM workshop. Accident Investigation Board Norway. Lillestrøm, Norway
- Organization of seminar with Andrew Hopkins (2009). Thinking about safety performance indicators, acceptable risk and lessons from Longford accident.

- Workshop on safety performance indicators with participation from Snorre Sklet (Statoil), Teemu Reiman (VTT), Urban Kjellén and Marius Aardal (DetNorske)

National Conferences

- Herrera, I (2008). Safety Days, Hva proaktive indikatorer kan brukes til?, Trondheim
- Herrera, I & Hollnagel, E. (2009). Safety Days, Workshop - Måling av sikkerhet, Trondheim

Appendix D. List of indicators identified in the studies

Investigating non-existing accidents, Norwegian aviation under major organizational changes
The maintenance review is considered as airlines´ representative to consider if the implemented organizational changes have significance to aviation safety. Indicators identified in the AIBN study, presented in Papers V and VI and trended as far as possible, period of five years 2000-2005, for three airlines and two helicopter operators included:

- Flight time production, number of aircraft, type of aircraft
- Reported incidents
- Technical and pilot reports
- Minimum Equipment List (MEL) reports
- Open items in the back-log and Hold Item List items (HIL)
- Cancellations and unscheduled downtime
- Maintenance costs
- Changes: staff per station, training, maintenance programme

The information gathered from the airlines and helicopter operators included:

- Airline activity, economy and personnel situation from 1999 until 2004
  - Number of aircraft per aircraft type and model
  - Number of flight hours per aircraft type and model
  - Number of cycles per aircraft type and model
  - Maintenance cost direct and indirect (specification items are included)
  - Maintenance cost for in house maintenance per aircraft type
  - Maintenance cost for subcontract maintenance per aircraft type
  - Total maintenance cost per aircraft type and model
  - Number of maintenance employees per category and base line and heavy maintenance (engineers, technicians, inspectors, logistics, etc.)
- Management documentation and competence
  - Safety Management Manuals
  - Maintenance Management Organisation Exposition Manuals
  - Description of personnel responsibilities and functions
  - “Release to service” procedures
  - Inspection of maintenance activities procedures
  - Dispensations procedures
  - Training and recurrent training programme and requirements
- Reports, statistics and follow-up
  - Incidents reports, internal investigation reports, air safety reports, flight occurrence reports, technical failures reports, ground occurrence reports
  - MEL reports
  - HIL reports
  - Carry forward / dispensations
  - Technical failures per type of aircraft per system (ATA) if possible which have been reported by pilots and by technical personnel (Techreps - Pireps)
  - Cancellations due to technical problems
  - Number of internal audits per type per year

Over last 10 years, there has been one helicopter accident in the Norwegian sector…Safety monitoring cannot based on the absence of accidents

Indicators identified from the scenario helicopter deck landing are presented in Table D 0.1 and Table D 0.2. Helicopter landing on helideck is consider as critical operation and considered relevant for identification of case specific indicators.

Table D 0.1 Helicopter landing indicators identified using FRAM

Indicators	Operationalization
Airworthiness of helicopter (quality of technical condition)	• **Minimum Equipment List (MEL)**: The status of critical systems. MEL can be an indicator which says something about to what extent the organization maintains continuous airworthiness. • Data from continuous use of Health Usage Monitoring System (**HUMS**) data for early identification of errors • Indicators related to **performance of maintenance work**
Quality of rig/facility report	• Quality assurance of personnel that make reports. **Training of weather observers**
Quality of information, technical equipment on helidecks	• Indicators tied to status of technical systems, **reports from facility**
Quality of updated procedures, which describe practices according to the helicopter type	• Conducted through audits or observations of normal operation of the Helicopter operators, ATM/ANS and helidecks. **Procedures, audits and compliance** • Active use of **Flight Data Monitoring (FDM)**
Quality of communication between helicopter and helideck personnel; procedures and practice	• Here there are individual differences as regards the use of language, and helicopter-related phraseology (especially for small floating facilities). Language knowledge varies on different facilities, particularly those with floating helidecks (ships). • **Observations** to provide a qualitative assessment of the development in regard to communication. • **OLF MANUAL/regulations update**. The number of changes in revisions. Harmonising of data provided to pilots, use of radio frequency (in some areas there is increased use of VHF). Passenger manifest is provided with METAR and TAF
Quality of team work (*CRM - Crew Resource Management*)	• Quantitative goals are difficult. **Observations** more important than interviews, but both can be used, see what works well and what does not work well
Competence to operate on the Norwegian Shelf	• **Observations** of normal operations. Use line check proactively. **Observations** in Line check are conducted once a year, with a simulator twice a year.
Quality of safety management in connection with organization changes	• Implemented **analysis of safety related consequences** of change
Contract conditions – ”*penalties*”	• If penalties are experienced during delays. **Deviations in time** from the operations centre; the number of bought days off per year, use of overtime
For floating helidecks, the occurrence of visual clues	• Status on helideck, according to CAP 437, OLF guidelines

Table D 0.2 Prioritized indicators for helicopter operations offshore by relevant organization

Indicator	
Topic	**Name/definition**
Helicopter operators – technical condition	
Health Usage Monitoring System (HUMS) data	**Continuous use of HUMS data** for early identification of errors

Indicator	
Topic	**Name/definition**
Helicopter technical condition: *Deferred Defect List* (DDL) & *Minimum Equipment List* (MEL)	**Number of DDLs** is a better indicator than MEL in relation to sufficient resources to uphold maintenance **Number of MEL** reports per year. Number of deviations/errors on systems that can influence safety. Flying with an MEL mark is possible and the helicopter is considered airworthy. If MEL increases or decreases, this can indicate something about parts and accessories, shortage of components, ability to correct faults
Predeparture check	**Quality of predeparture check**, competence, experience
Procedure compliance	**Number of deviations from procedure**. **Audits and observations** reveal whether there are deviations between procedure and practice.
Revision of procedures	**Number of updated revisions of procedure in the last period**
Maintenance programme, Updating	**The number of updates per year**, is seen together with information sent to the technician, e.g. technical information with updating. The number does not say anything about quality; therefore it is necessary to view this together with other indicators for procedure revision and compliance.
Change in maintenance programme	**Number of changes in important programmes and for tasks with short intervals**
Back-log	**Average back-log in maintenance tasks** per company per year. An alternative indicator is repair times in relation to MEL.
Maintenance, crew, scope	**Planned crew versus real crew per station per shift. The number of maintenance hours per flight hour**
Cooperation	**Quality of cooperation.** Quantitative goals are difficult. **Interviews** better cover the status of the cooperation.
Communication	**Quality of communication.** Quantitative goals are difficult. **Interview** better covers the condition of the cooperation.
Workload Sufficient resources and slack	**Average work time (hours per day) for employees per year.** Use of overtime is seen in relation to exemptions and in relation to crew
Helicopter operators – flight operative condition	
Revision of procedures	**Number of ”*notice to pilots*”** or **”*information to crew*”** (revision of procedure)
Procedure compliance	**Active use of FDM analyses.** Audits and observations reveal if there are deviations between procedure and practice.
Training, cooperation and communication	Proactive use of ***Line Check*** in relation to observations of “normal operations”. Simulator training, number of hours and training in beyond regulatory requirements.
Crew, sufficient and shortfall in resources	Number of purchased days off per year.
Penalties	Follow up of the penalty regime, and how this influences the organization. If there are penalties associated with delays and stress in the organization to maintain regularity.
Exemptions	**Average number of applications for exemption related to aviation safety** per company per year (i.e. in relation to maintenance interval and DDL)
Helicopter lagging indicators	
Aviation incidents	**Number of serious aviation incidents per 100 000 flight hours.** If there are serious incidents the organization must act.

Indicator	
Topic	**Name/definition**
ATA-code (*Air Transport Association, code to classify systems within aviation) reports.*	**The number of repeating deviations per ATA per period.** This is analysed in the Maintenance Review Board meetings. Analysis includes when the same deviations (deviation within the same ATA chapter) occur on two or more flights within a time period. Systems that can be regarded as critical are related to communication (ATA 23), navigation (ATA 34) and rotor (ATA 65-70) The number of technical errors per system. i.e. • *Windshield cracking* • *Chip warning* • *Door open warning* • *Oil leakage detected by walk-around* • Error involving main rotor, gearbox (UK, based on accident, 2009)
Pilot reports	**Number of pilot reports per year.** This is analysed in the Maintenance Review Board meetings
Air Traffic/Navigation Service (ATS/ANS)	
Radar/ADS-B coverage; surveillance coverage, controlled airspace	**Percentage of the area with radar coverage** (Ekofisk, Ula, Sleipner, Heimdal, Statfjord CTA, Haltenbanken, Norne and the Barents Sea)
Radio communication	**Number of/per cent redundant communication systems per area** (Ekofisk, Ula, Sleipner, Heimdal, Statfjord CTA, Haltenbanken, Norne and the Barents Sea)
Exemptions	**Per cent exemptions.** Many flights with exemptions result in a negative impact on the work situation (Avinor has the best overview of exemptions granted directly on an ad hoc basis)
Procedure inquiry	Audits reveal potential deviations between procedure and practice. Can be acquired from feedback (interview) from pilot – user survey – focus on standardization. This can be seen in connection with the use of FDM (Flight Data Monitoring).
Cooperation, phraseology, communication	Observations are more important than interviews, but both can be used, must see what works well and what does not work well.
Crew	Average work time per position per unit per year, seen in relation to exemptions. The number of persons per control unit in relation to sector. Can analyse if the development in the on call duty lists is expedient and if there is a possibility for improvement.
Lagging indicators	
Aviation incidents in MESYS (Confidential fault reporting, Avinor)	**Number of serious aviation incidents per 100 000 flight hours.**
Trend in reporting to MESYS	**Number of reported incidents in MESYS.** (People are not as careful with reporting errors in technical systems, as they are with incidents)
Helideck	
Revision procedure	**The number of facilities with a logbook/system which tracks that people sign off that they have read the** revisions. It is important to have a system and an active use of the distribution list. The challenge lies in shift work and ensuring that everyone has the same information. Experience transfer and handover are IMPORTANT in relation to changes of procedure on shifts. **OLF MANUAL; regulations update**. The number of changes per revision. Harmonising of data which is provided for pilots, better use of frequency (in some areas there is an increased use of VHF and facilities for planning fuel). Passenger manifest is given with METAR and TAF.

Indicator	
Topic	**Name/definition**
Procedure compliance	**Number of deviations** (in relation to following procedures, procedure is available). Audits reveal potential deviation between procedures and practice.
Training, weather observers	**Per cent persons with radio responsibility that are trained per year per facility.** BSL-G-MET is finished and will be implemented. Important to have a continuous process that keeps the personnel professionally up-to-date.
Technical condition and lighting on helideck	Here there is a need to look at the status of helidecks in relation to OLF guidelines and CAP 437
Lagging indicators	
Incidents	Number of reports of undesirable incidents per year in relation to helideck/HFIS

PART II - Papers

- Paper I
  Is there a need for new theories, models and approaches to occupational accident prevention?

- Paper II
  Building Safety Indicators, Part 1 - Theoretical foundation

- Paper III
  Leading indicators applied to maintenance in the framework of resilience engineering: A conceptual approach

- Paper IV
  Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis

- Paper V
  Aviation safety and maintenance under major organizational changes, investigating non-existing accidents

- Paper VI
  Key elements to avoid drifting out of the safety space

- Paper VII
  Proposing safety performance indicators for helicopter offshore on the Norwegian Continental Shelf

- Paper VIII
  Approaches to elaborate on the safety of offshore helicopter operations

Paper I

Is there a need for new theories, models and approaches to occupational accident prevention?

Safety Science (2010)
vol 48, pp. 950-956

Invited paper for prevention of occupational accidents in a changing work environment, WOS 2008, held in Crete, Greece, on 30th September -3rd October 2008

Safety Science 48 (2010) 950–956

Contents lists available at ScienceDirect

Safety Science

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

Is there a need for new theories, models and approaches to occupational accident prevention?

Jan Hovden [a,*], Eirik Albrechtsen [a,b], Ivonne A. Herrera [c]

[a] Dep. of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[b] Dep. of Safety Research, SINTEF Technology and Society, Trondheim, Norway
[c] Dep. of Production and Quality Engineering, Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Article history:
Received 19 January 2009
Received in revised form 7 May 2009
Accepted 26 June 2009

Keywords:
Accident model
Occupational accidents
Systemic accidents

ABSTRACT

This paper discusses occupational accident modelling challenges associated with a changing working life, and asks whether ideas from models developed for high-risk, complex socio-technical systems can be transformed and adapted for use in occupational accident prevention. Are occupational accidents mainly simple component failures or is a systemic approach to the phenomenon of some interest and value?

1. Introduction

The aim is to invite to a discussion about theories and models in the field of occupational accident prevention. Is the current knowledge base satisfactory, or is there a need for inspiration and approaches taken from other fields of risk research? Is there a need for radical changes, for modification of traditional approaches and knowledge bases; or do the problems of occupational accident prevention mainly amount to a question of priorities, resources and implementation of known remedial actions?

The intention is not to give clear and finite answers to these questions. Rather, the paper invites to reflections on the needs for and uses of accident models in occupational accident prevention through: (1) a brief review of established, mainstream accident models applied in this field; (2) a description of changes in working life with possible impacts on the need to rethink the paradigms for accident modelling and safety management approaches; (3) a brief review of new theoretical approaches to high-risk complex socio-technical systems; and (4) a discussion of the purposes and functions of occupational accident models in a new context which takes into account the impact of economic, political, organisational, and technological stressors on safety performance.

A delimitation: Approaches to technical risk analysis are not dealt with, and systemic models are presented only briefly and discussed solely in relation to their relevance for occupational accidents. The discussions are primarily based on today's situation in the Nordic countries both with regard to safety management practices and the associated challenges, thus defining the scope of the paper. A common basis for a Nordic framework for occupational accident prevention is described by the Finnish Institute of Occupational Health (1987).

2. Background – the established models in occupational accident prevention

Accident definitions converge in certain assumptions which describe an accident as a hazard materializing in a sudden, probabilistic event (or chains of events) with adverse consequences (injuries). Classification is used as a tool to standardise the collection and analyses of data on accidents. There are four main standard categories (Kjellén, 2000):

- *Damage/loss:* includes injuries and fatalities, material and economic losses, reputation, etc.
- *Incident:* subdivided into Type (fall, slip, explosion, etc.) and Agency (machine, vehicle, tool, etc.).
- *Hazardous condition:* covers defective tools, unsafe design, housekeeping, etc.
- *Unsafe act:* covers errors and omissions.

In addition, accidents can be categorised according to arena, i.e. where the accident happens, the type of activity involved, system characteristics, etc. The ESAW methodology for statistics on

---

[*] Corresponding author. Address: NTNU, IOT, NO 7491 Trondheim, Norway. Tel.: +47 73 59 35 07; fax: +47 73 59 31 07.
E-mail address: hovden@iot.ntnu.no (J. Hovden).

0925-7535/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ssci.2009.06.002

accidents at work (European Commission, 2001) present a characterisation of variables which is representative for an epidemiological approach to occupational accident data and statistics; e.g. properties of the enterprise, exposure and employee, organisation and workplace, working conditions, sequence of events, and the victim. *Occupational accidents* are distinguished from other accidents by the facts that they happen in a working life context and that the main consequences are limited to injuries on the involved workers. Furthermore, the worker is often the agent as well as the victim of the injury. Injuries are classified according to the nature of the injury (cut, fracture, burns, etc.), the part of the body affected by it (head, neck, etc.), and its severity.

Most accident models and theories applied in the field of occupational accidents are still based on the ideas in Heinrich's (1931) domino model, Gibson's (1962) and Haddon's (1968) epidemiological models of energy-barriers, and are using a closed system safety mindset with mechanistic metaphors to describe the conditions, barriers and linear chains of an accident process. In the 1960s and 1970s there was typically a focus on technical faults and human errors (Kjellén and Hovden, 1993).

Competing modelling approaches evolved: (1) causal sequences similar to the domino model, e.g. ILCI (Bird and Germain, 1985), (2) descriptive models of accident processes in terms of sequentially timed events and/or phases, e.g. STEP (Benner, 1975; Hendrick and Benner, 1987), and OARU (Kjellén and Larsson, 1981), (3) system models based on a mixture of causal sequences and epidemiological models, e.g. TRIPOD, and the ''Swiss cheese''model (Reason et al., 1988; Reason, 1997; Reason, 2008), (4) logical, risk analysis inspired models, e.g. the fault three based MORT method (Johnson, 1980) and the similar SMORT method (Kjellén, 2000) tailored for occupational accident investigation. In recent years technical risk analysis modelling has also been used for quantifying occupational accident risks (Ale et al., 2008). Systemic accident models are mainly within the domain of major accident risk research and will be presented in part 4. There are many ways of classifying accident models, and the grouping above is a mixture of Kjellén (2000) and Hollnagel (2004). Epidemiological models are mainly used in statistical accident reporting systems for monitoring safety (Andersson, 1991), whereas more complex models are used in in-depth accident investigations.

The 1980s was the era of creative occupational accident modelling activities, and a number of different occupational accident models were developed in the Nordic countries in addition to OARU, such as a Finnish model (Touminen and Saari, 1982), and a Danish one (Jørgensen, 1985). For a review of accident models, see Kjellén (2000), Leveson (2001), Sklet (2004), and Lundberg et al. (in press). As a reaction to all these efforts in accident modelling, Hovden (1984) asked provocatively, ''Do we need accident models?'' at the yearly Nordic conference in accident research,[1] questioning the utility of these analogue models of boxes and arrows in relation to the progress of safety science and improved accident prevention in industry. The pessimistic conclusion was that the models were not scientific enough, or practical enough, and neither were they specific enough, nor holistic enough to serve this purpose.

Andersson's (1991) work on the role of accidentology in occupational injury research discusses classifications of accident theories and models and revealed a split between traumatology and epidemiological approaches on one side, and technological and cross-disciplinary approaches on the other. The history of accident modelling is very much about a positioning on model power between different disciplines, with technologists, psychologists, other social scientists and so on claiming to be holistic and cross-disciplinary in their combining of human factors, technology and organisational aspects, while combining these according to their own biased mental models.

From the mid-eighties the focus changed from accident modelling to an interest in management tools for safety monitoring and safety auditing (Kjellén and Hovden, 1993). Hale and Hovden (1998) described management and culture as the third age of safety. The first age was preoccupied with technical measures, whereas the second focussed on human factors and individual behaviour (Hale and Glendon, 1987). The latter was influenced by ergonomics and later merged with the technological approaches. In the 1980s the socio-technical approaches based on the Tavistock School, which had a long tradition in working environment studies, influenced accident modelling. During the same period large international companies such as DuPont became role models for many companies through their focus on management responsibility, workers' behaviour, and safety performance indicators based on incident reporting.

''The three ages of safety'' (Hale and Hovden, 1998) are about an expansion of perspectives on accident phenomena. The perspectives are not substituting each other, but supplement each other. Technical safety is still important, and human factor research methodological development on accident risks has – also flourished the last decades, e.g. HERMES, ATHEANA, CREAM – for a review see Hollnagel (1998). This review of approaches to accident modelling and prevention must necessarily be brief. Nevertheless, it reveals a great variety of perspectives on accident phenomena and preventive strategies: we find what we look for, and fix what we find (Lundberg et al., in press). However, do the dominant occupation accident models and approaches provide good enough understandings of current and future challenges of occupational accident prevention in a changing working life environment?

3. Changes in working life

Wilpert (2009) presents a comprehensive discussion of the impact of globalization on human work. He describes the impact of new information technologies, changing work structures in industrialized countries and changing industrial relations systems. In a report from the European Agency for Safety and Health at Work (Op de Beek and Van Heuverswyn, 2002) the changes are listed as: (1) changing industrial organizations, (2) the free market, privatization, downsizing, subcontracting, (3) technology, (4) the growing use of remote operations, homework, changes in working hours, work pace and workload, (5) changing labour market with an increase in part-time jobs, temporary work, self-employment, women in employment, the ageing of the workforce, etc. According to that report the emerging risks of the changes should be met by a dynamic safety management system emphasizing participation, leading performance measurements, communication and life-long learning.

In their process of coping with accident risks, the levels and layers within organisations are subjected to stress from a number of external forces and counter forces. The main contextual stressors influencing working life risks are changing political climate and public awareness, market conditions and financial pressures, competence and education concerns, and the fast pace of technological change (Rasmussen, 1997). Adaptation to these stressors has changed the everyday reality of work, the contents of work processes and the socio-technical systems at traditional workplaces.

At the microlevel, or ''the sharp end'', very little has changed in manual work tasks: climbing a ladder at a construction site or performing maintenance work in a chemical plant has not become complex, dynamic or intractable. However, does the increased

---

[1] These Nordic conferences (NOFS) were a precursor to the WOS conferences.

use of information and communication technologies as integral parts of manual work and the construction of new, distributed industrial organisations change the characteristics of the work in such a way that alternative approaches to accident prevention are needed?

The use of information systems creates new types of communication, improved ability to store and retrieve data and more effective information processing, and all of these factors influence the modern organisation of work (Groth, 1999). New developments in these areas create advantages such as the automation of work processes, more effective planning and communication, and improved employee availability. However, the "information revolution" also creates pitfalls such as information overload, high demand for information, and communication problems. Example No. 1 below shows an example of how information technology can influence occupational safety both positively and negatively in relation to sharp-end activities.

> *Example No. 1: Monitoring work performance at offshore installations*
>
> Within the Norwegian oil and gas industry, there is an ongoing transition to the concept of integrated operations, i.e. the use of information technology and real-time data to improve decision-making processes and cooperation across disciplines and organisations. One of the implications of this development is increased monitoring of offshore workers. This implies that operators onshore can watch offshore workers' performance by use of camera equipment and monitors. On the one hand, this creates a secure and safe environment, as offshore workers' performance is monitored by a 'watchful eye', making it possible to prevent and stop unwanted actions. It also provides offshore workers with decision support from onshore experts. On the other hand, monitoring may lead to workers feeling uncomfortable at being evaluated all the time and even result in a sense of mistrust.

Globalization has also reached working life today. A study performed for the Norwegian Labour and Welfare Administration shows that one out of three Norwegian companies used manpower from the EU countries in 2007. Sectors which have traditionally had a high rate of occupational accidents top the statistics for use of foreign manpower: primary industries (58% have used foreign manpower), hotels and restaurants (49%), manufacturing, and the building and construction industry (both 43%). Half of the foreign workers were hired on temporary contracts. Swedish and Polish workers are the most widely used foreign nationality workers in Norway (Perduco, 2007). On the one hand, this work immigration provides much needed labour and expertise; but on the other hand, it also creates challenges in relation to occupational accidents (See example 2), as well as working environment challenges regarding social rights. Mearns and Yule (2009) have studied occupational safety and how globalization process affect attitudes, beliefs and behaviour of "national" workforces for multi-national companies. The study remarks that management commitment is a more important determinant of behaviour than national culture. Thus organisational issues like management and leadership emerge as important for safety performance also in a new context of working life.

As a result of the impact of information technology, globalization and a dynamic post-industrial society, work is organized in new ways. For industrial workplaces, automation is an evident change to the organizing of the work (see example 3). However, Zuboff (1988) argues that the impact of information technology is dual: it both automates and informates the organisation. It automates manual activities and it generates information about underlying productive and administrative processes, which can be used to understand, improve and plan activities. For example, an organisation is informed by access to data and information that produce new ways of achieving more effective planning, as well as faster input to safety methods and tools. The informated organisation, i.e. an organisation that utilizes the benefits of information technology, occurs in many configurations (Groth, 1999), such as car manufacturing with its tight cooperation with subcontractors based on just-in-time principles.

> *Example No. 2: Language barriers creating an occupational accident*
>
> At a Norwegian chemical plant, a Finnish welder was hired from a contractor to stop a leakage from a pipe containing lye. Due to language problems, the Finnish welder misconceived the mission and thought is was an air pipe that was to be repaired. As a consequence, the welder failed to wear the necessary protective equipment for the job, putting on only a mask and gloves. While preparing to fix the leak, the welder discovered that a green liquid was coming out of the pipe, and only then understood that he was dealing with a chemical liquid. During his search for the leakage, drops of lye landed on his neck. He reacted to the drops by touching his neck with his gloves, which had already been in contact with the lye.

> *Example No. 3: Automation of manual work*
>
> A study of occupational accidents and costs in the Norwegian furniture industry revealed that automation of the production line reduced the number of injuries, especially the cutting of fingers, but that maintenance and handling of disruptions resulted in more severe injuries, e.g. the amputation of arms.

As part of the wave of globalization, we see a trend of deregulation and new concepts in business administration related to profit, time and cost cutting. These concepts include capital cost reduction, outsourcing, downsizing, management, contracting, leasing, strategic alliances, joint venture/partnership, enterprises in network, lean production, just-in-time (Kanban), business process re-engineering, flexible specialization, and virtual organizations, plus learning organizations, knowledge management, and change management. Example No. 4 shows an example of complex organisation of work resulting from new contexts and types of organisation. The question here is how these new realities fit into occupational accident models, and how they are considered and dealt with by the safety management.

Four elements are identified in relation to subcontracting and occupational safety (Mayhew et al., 1997). The first is related to economic pressures, where occupational safety interventions are not usually perceived as good investments. The second is disorganization, leading to a fragmentation of work place where major firms produce specific safety manuals having little or no effect on subcontractors. The third is dealing with inadequate regulatory controls. Regulations primarily deal with a traditional, stable employer/employee relationship in mind and not addressing the issue of small organizations pressured by time and costs, and changing frequently work site. The increased use of outsourcing is a challenge for regulation and control. Finally, the fourth element addresses the ability of workers to organize. Self-employed workers do not normally unionize and address compensation claims. Their injuries and illness are not properly recorded. This has an implication on the visibility of subcontracting/outsourcing on occupational safety performance.

Example No. 4: Complex organisation of maintenance work

Maintenance activities in aviation are typically spread over multiple locations, the task complexity varies, the working environment is non-optimal (in terms of light, access, noise), working times differ, downsizing, increase of subcontracting and seasonal recruitment are common. Several programmes have been introduced to improve worker and operational safety. In aviation, there is still a need to take a more realistic approach; one that considers human behaviours and decision-making processes in operational contexts. A proposed solution is to monitor normal operations. This monitoring introduces the challenge of negotiating between approaches that decompose data into quantitative factors and those that use interpretation of qualitative data to increase the understanding of normal performance.

Working life has changed along with the transformation of the stable industrial society of former times into the dynamic knowledge society of today. Post-industrial working life is characterized by provision of services, handling of information, and knowledge-intensive work. While the industrial society was perceived as stable, the post-industrial society is dynamic, with its technological development, international competitions, efficiency demands and changes. This implies that the post-industrial organisation of work is different from that of industrial bureaucratic organisations. Some examples of changes in different types of work include, by industry:

- *Craft industries*: in general, more mobile phones, more foreign workers, automated tools replace hammers and saws, etc.
- *Manufacturing*: increased automation (see Example No. 3).
- *Farming and fishing*: less manual work, increased production volumes and energy consumption.
- *Process and petroleum industry*: increased automation, integrated operations.

Types of industrial systems and context of which they operate have changed. Leveson (2004) enumerate changes on systems that are stretching the limits of current models and safety analysis techniques. Changes affecting the working life are the fast pace of technological changes, the changing nature of accidents, new types of hazards decreasing tolerance to single accidents and changing regulatory and public view on safety. Beyond these changes, work place accident prevention has also been affected by organisational changes. Even if prevention of occupational accidents is adequate, the introduction of innovations and organisational changes may have an impact in the way safety practice is carried out. Traditional firms implement changes without systematic consultation of workers. Workers in successful modern firms where major technological and organisational changes are introduced seem to be more involved in the improvement of their working condition (Harrisson and Legendre, 2003).

The world economy has been through a long period of revival which now has turned into a recession. ILO asks: “Will the financial crisis push us back in the struggle for safer and healthier workplaces?” (Al-Tuwaijri, 2009). The research literature on the effects of economic cycles on occupational safety seems scarce. But the research field of road traffic accidents tells that safety increase in recessions more than the effects of the decrease in exposure (Wilde, 1998).

To sum up, technologies, knowledge, organisations, people, values, and so on are all subject to change in a changing society. Nonetheless, when it comes to occupational accident prevention most experts and practitioners still believe in the domino model and the iceberg metaphor (Heinrich, 1931; Bird and German, 1985; Hale, 2000).

4. New approaches to safety in complex and dynamic socio-technical systems

In view of the changes briefly described above and the challenges of vulnerability in complex, dynamic socio-technical systems, theories and models have evolved in relation to the high-risk industries and in transportation. There is some overlap between these theories, and our intention here is restricted to highlighting some key elements and possible lessons to be learned and applied in the field of occupational accidents. Two prominent schools which have addressed the organisational aspects of safety are Normal Accident Theory (NAT) and High Reliability Organisation (HRO) theory.

NAT (Perrow, 1984) introduced the idea that in some systems accidents are inevitable and normal. Such system accidents involve *the unanticipated interaction of multiple failures.* NAT presents a two-dimensional typology of socio-technical systems based on degree of interaction and couplings. Perrow uses these two dimensions in a two-by-two table to indicate that different systems may need different ways of organizing. If the system is both interactively complex and tight coupled, there is no possibility for identifying unexpected events, and the system should be abandoned. In such systems, simple, trivial incidents can develop in unpredictable ways with potentially disastrous consequences. The changes in working life, as described above in part 3, have resulted in increased safety, but also increased vulnerabilities at most workplaces over the last twenty years. Therefore, it may be important to reduce interactive complexity and tight coupling in the design of workplaces. Bellamy and Geyer (1992) argue that in the new work environment there are too many tightly coupled and complex systems to be abandoned as Perrow recommend. Perrow (2007) discusses safety potential of an alternative model for organisation, the “Network of small firms”. In this model the dependencies are low, with multiple sources, and single, unexpected failures will not disrupt interdependencies since other firms can change or absorb the business. The decentralized nature of small firms has positive and negative effect on safety. The small firms require good safety practice towards client requirements. On the other hand these firms have limited resources for occupational safety investments.

HRO researchers claim to counter Perrow’s Normal Accident Theory (LaPorte and Consolini, 1991). HRO theory is based on studies on organisations that successfully handle complex technologies. The cost of failures in such organisations is socially unacceptable. The main characteristics of HROs include managing of complexity through: (1) continuous training, (2) use of redundancy, and (3) numerous sources of direct information. Furthermore, HROs rarely fail even if they experience unexpected events (Weick and Sutcliffe, 2007). They redefine HROs as ‘mindful’ organisations. Organizing for high reliability requires continuous anticipation and containment in the running of the organisation. A mindful organisation operates according to the following three principles for anticipation: (1) an ability to become aware of unexpected events through preoccupation with failures, reluctance to simplify, and sensitivity to operations; (2) commitment to resilience (involving abilities to absorb and preserve, to recover and to learn); and (3) deference to expertise (migration of decision-making to the levels were people come together to solve a problem).

Some researchers argue that the HRO approach does not contradict or falsify NAT at all because the conclusions of the HRO theory are based solely on a few case studies which do not fulfil Perrow’s definition of complex interactivity or of tight coupling (Marais et al., 2004).

In an information processing perspective, the accident is viewed as a breakdown in the flow and interpretation of information

(Turner, 1978). This perspective highlights how the individuals and the organisation perceive and make use of information. A key point is to establish how information and knowledge are related to the accident and how misinformation may arise. The model includes factors such as wrong interpretation of signals, information ambiguities, disregard for rules and instructions, and overconfidence and organisational arrogance. A response to this perspective was the description of how organisations treat information in a (1) pathological, (2) bureaucratic, and (3) generative way (Westrum, 1993). This also became a basis for classifying and ranking safety cultures (Reason, 1997).

Rasmussen (1997) directs the attention to the migration of activities towards the boundary of acceptable performance – a migration which is influenced by the pressure towards cost-effectiveness in an aggressive and competitive market. He argues that it is feasible to provide the necessary decision support to operators, and proposes a distributed decision-making system in order to cope with the dynamics of modern organisations. He also recommends studying normal work processes rather then focusing on deviations, errors and incidents. Rosness et al. (2004) point out that in the migration model, regulations and procedures keep the actors within the boundary of safe operation and prevent conflicts between activities when decision-making is distributed.

From the classical definition of safety as freedom from unacceptable risk, through safety seen as a dynamic non-event (HRO), to the ability to predict, plan and act to sustain continuous safe operation, the Resilience Engineering school presents an alternative or supplementing perspective, claiming that instead of focusing on failures, error counting and decomposition, we should address the capabilities to cope with the unforeseen. The ambition is to "engineer" tools or processes that help organisations increase their ability to operate in a robust and flexible way.

Hollnagel et al. (2006) define resilience engineering as the "intrinsic ability of an organisation (or system) to adjust its functioning prior to or following changes and disturbances to continue working in the face of continuous stresses or major mishaps". The premises of this definition are the following: (1) the increase of complexity has made the systems intractable, and therefore under-specified; (2) people are seen as an asset because they are flexible and can learn to overcome design flaws, they can adapt to meet demands, interpret procedures, detect and correct when things go wrong, and use "requisite imagination" (Westrum, 1993) to cope with the unexpected; and (3) systems balance efficiency and thoroughness to meet demands. Hence resilience engineering encompasses research on successes and failures in socio-technical systems, organisational contributions and human performance. A systemic view is encouraged in order to understand how the system as a whole dynamically adjusts and varies for the sake of continuing safe operations. The focus is on the proactive side of safety management and the need to make proper adjustments in terms of anticipation, updating of risk models and effective use of resources.

People, organizations and technology are under continuous change. Do these changes represent a growing complexity in the working life? As a parallel to Perrow's description of complexity and tight coupling, Hollnagel (2008) proposes the concepts "tractable" and "intractable": a system, or a process, is tractable if the principles of functioning are known, if descriptions are simple and with few details, and most importantly if the system does not change while it is being described. Contrary, a system or a process is intractable if the principles of functioning are only partly known or even unknown, if descriptions are elaborated with many details, and if the system may change before the description is completed". Accident models and theories provide different "glasses" that will influence the way we look for, understand, analyse and provide recommendations.

5. Discussion

Can theories from the domain of high-risk complex and advanced socio-technical systems such as the ones advocated by Perrow, Rasmussen, Weick, Hollnagel and others contribute to better understandings and practices in relation to preventing traditional and often seemingly simple and trivial occupational accidents? Do they have something substantial to add to this area, or do they represent a different world of risk problems?

These questions cannot be answered with a simple yes or no. The traditional approaches may be good enough; suited to some workplaces but not to others, and suited to understanding some accidents but not others. In occupational accident prevention most problems may be solved by looking at simple, direct causes and triggering events. In most industrial domains, there is a high potential for achieving low injury rates through continuous work to improve performance through deviation control. Saari (2001) states that humans tend to underestimate known risks and overvalue new risks. Still falls cause a large proportion of fatalities at workplaces, but are old and well known.

The need for new models can therefore be considered as low in the daily work of accident reporting and surveillance. Merely identifying a proximate cause as the "root cause" may, however, lead to the elimination of symptoms without much impact on the prospect of reducing future accidents (Marais et al., 2004; Leveson, 2004). In order to identify systemic causes, one may need to supplement with models representing alternative mindsets in order to spark the imagination and creativity required to solve the accident risk problem.

The use of accident models can be discussed in a framework of learning loops at different levels (Freitag and Hale, 1997). At the sharp end, i.e. the "execution" or work processes level (Hale et al., 1997), very simple and rather iconic models for reporting and communication may be needed in order to achieve valid information and immediate actions based on first order learning (van Court Hare, 1967). At the meso level, i.e. "planning" by safety professionals, more advanced analogue models such as TRIPOD, ILCI, etc., may be appropriate for second order learning by monitoring and auditing. For emerging events related to new technologies and changes in a context which are difficult to understand and specify, it may be helpful to look at modelling approaches based on system dynamics, or at more rare approaches and paradigms from anthropology, e.g. ones that are based on story telling and text mining, studying of normal work processes, etc. Developments in information technology make such approaches to accident prevention more applicable.

At the level of "structure" or strategic management, it is important to distinguish events that suggest that fundamental changes are needed in the safety management system or the regulatory regime from those that suggest that greater efforts are needed with respect to implementing the systems and preventive measures already in place (Hale, 1997). Important tasks at this level are to conduct a *change analysis* related to impacts on safety caused by changes in technology, organisation and work processes, and to consider remedial actions within a framework of cost-benefit for the company and regulatory constrains imposed by the government. For these tasks the basic ideas of resilience engineering seem appropriate. "Resilience" has become a popular buzzword in many research areas. It seems to inspire a feeling that it represents an answer to the threats and uncertainties associated with the fast-paced changes of modern society. The ongoing developments in the field of "resilience engineering" are promising in relation to needs in strategic occupational accident risk management, but the field is still immature with regard to practical and applicable tools for the industry.

Accident models affect the way people think about safety, how they identify and analyse risk factors, and how they measure performance. Accident models can be used in both reactive and proactive safety management. Many models are based on an idea of causality. Accidents are thus the result of technical failures, human errors or organisational problems. Most applied performance indicators do not take into account whether the consequences of failures are major or minor, e.g. Lost Time Injury (LTI), and are build on pre-assumptions based on an iceberg metaphor for the relationship between unsafe acts, injuries and fatalities (Heinrich, 1931; Hale, 2000). Many models, e.g. the Swiss cheese model (Reason et al., 1988), have an underlying idea that actions at the "sharp end" are influenced by conditions set at the "blunt end". The measurement of performance is based on the status or effectiveness of the risk control systems, such as barriers, maintenance error, failure to control hot work, etc. (Hopkins, 2007).

Recently, two systemic models have been introduced, namely the Functional Resonance Accident Model (FRAM) (Hollnagel, 2004), where failures and successes are the result of adaptations to cope with complexity; and the Systems-Theoretic Accident Model and Processes (STAMP) (Leveson, 2004). These models may inspire a more creative search for alternative and proactive (leading) safety performance indicators.

The accident model applied guide the choice of performance indicators and gives a reference point from which they can be interpreted (Hollnagel, 2008). Herrera and Hovden (2008) define leading indicators as precursors that when observed, imply the occurrence of a subsequent event that has an impact on safety and performance. Leading indicators are indicators that change *before* a change has occurred in the calculated risk. In FRAM the idea of causality is replaced by emergence, whereby a combination of factors in a given context can produce an unexpected outcome. At sharp-end level, leading indicators are factors such as overtime, seasonal recruitment, and the quality of training, adequate feedback from reporting, sick leave levels, how risk management processes are systematically integrated into normal activities (use of safe job analysis), and interpretation and update of procedures. At the organisational level, Wreathall (2001) suggests leading indicators related to management commitment, awareness, preparedness, and flexibility.

Is there a need for models that are more flexible in the sense that they can be adapted and tailored to specific work contexts and local needs? If yes, it reveals a need to develop taxonomies of types of workplaces, relevant features of the socio-technical systems, the phenomenology of incidents and energy involved and so on, merged with a categorisation of main accident theories, models and approaches to accident prevention. This task may be approached by developing a representative list of accident scenarios as a basis for defining the contents of the taxonomies. This is huge research challenge – a challenge which we leave for further research to address.

There are many reasons for discussing the need for accident models, namely to:

- Create a common understanding of accident phenomena through a shared simplified representation of real-life accidents.
- Help structure and communicate risk problems.
- Give a basis for inter-subjectivity, thus preventing personal biases regarding accident causation and providing an opening for a wider range of preventive measures.
- Guide investigations regarding data collection and accident analyses.
- Help analyse interrelations between factors and conditions.
- Different accident models highlight different aspects of processes, conditions and causes.

Therefore, *many* different and competing models are welcome as they highlight different aspects of the risk problem (Kjellén, 2000). They are simplified representations of real-life accidents, not right or wrong, and should be evaluated on their applicability in different risk arenas and on the guidance they can offer in terms of proper and effective remedial actions.

6. Implications and conclusions

Organisations today are under stress from a number of dynamic factors in their environment, such as technological changes, globalization, and market conditions. Modern socio-technical systems are characterized by increased complexity and coupling, and are as a consequence increasingly intractable (Hollnagel, 2008). However, it can be argued that working life at the sharp end has remained largely unaltered, although some changes have occurred at this level as well. Examples of such changes include automation of manual work, the increased use of migrant workers and multicultural challenges at workplaces, and new use of information technology to coordinate work and to communicate effectively. The question addressed is whether new theories from other fields of risk research can play a constructive role in occupational accident prevention. There is no straightforward answer to this question. There seems to be little need for new models and approaches for the sake of understanding the direct causes of occupational accidents in daily work at the sharp end. For this purpose, Gibson's (1961) basic energy-barrier model and Haddon's (1968) 10 strategies for loss prevention will never be outdated.

However, as a result of the changes at higher levels than the sharp end in post-industrial society, theories, models and approaches to high-risk complex socio-technical systems have the potential of enriching occupational safety management activities such as learning from accident models (understanding root causes), planning (expecting and responding to the unexpected) and change analysis.

Normal accident theory, the theory of high reliability organisations, and resilience engineering have all been developed and used within the context of complex high-risk socio-technical systems. Theories from such risk research domains are nevertheless important contributors to discourses on occupational safety management approaches, as they represent an invitation to consider whether new models and approaches can supplement and improve current approaches to this subject area.

Based on these arguments presented, there is a need for further discussions and research on the development of new tools to be added to the occupational safety management toolkit. Examples of areas to be explored are leading indicators, mapping and understanding normal operations (work as actually performed), improvements of accident models and approaches to accident investigation.

References

Ale, B.J.M., Baksteen, L.J., Bellamy, L.J., Bloemhof, L., Gossens, L., Hale, A., Mud, M.L., Oh, J.I.H., Papazoglou, I.A., Post, J., Whiston, J.Y., 2008. Quantifying occupational risk: the development of an occupational risk model. Safety Science 46 (2), 176–185.

Al-Tuwaijri, S., 2009. Will the financial crisis push us back in the struggle for safer and healthier workplaces? ILO Global Job Crisis Observatory Note. <http://www.ilo.org/public/english/support/lib/financialcrisis/featurestories/story12.htm>.

Andersson, R., 1991. The role of accidentology. Arguments for a scientific approach. Doctoral Dissertation, Arbete och Hälsa 17

Bellamy, L. J., Geyer, T. A. W. 1992. In: (ed Williams, J. C.,) Organisational, Management and Human Factors in Quantified Risk Assessment. HSE Contract Research Report No. 33/1992. <http://www.hse.gov.uk/research/crr_pdf/1992/crr92033.pdf>.

Benner, L., 1975. Accident investigations. Multilinear events sequencing methods. Journal of Safety Research 7 (2), 67–73.

Bird, F.E., Germain, G.L., 1985. Practical Loss Control Leadership. Inst. Publ. Division of International Loss Control Institute, Atlanta, Georgia.

European Commission, 2001. European Statistics on Accidents at Work (ESAW). Methodology. DG Employment and Social Affairs, UNIT D-5. Health, Safety and Hygiene at Work. Brussels. <http://osha.europa.eu/en/publications>.

Finnish Institute of Occupational Health, 1987. Successful Accident Prevention. Recommendations and Ideas Field Tested in the Nordic Countries. Review 12. Finish Institute of Occupational Health, Helsinki.

Freitag, M., Hale, A.R., 1997. Structure of event analysis From Accidents to Organisational Learning. In: Hale, A.R. et al. (Eds.), After the Event. Pergamon, Oxford.

Gibson, J.J., 1961. The contribution of experimental psychology to the formulation of the problem of safety – a brief for basic research. In: Behavioral Approaches to Accident Research. Association for the Aid of Crippled Children, New York, pp. 77–89.

Groth, L., 1999. Future Organizational Design. The Scope for the IT-based Enterprise. John Wiley, Chichester.

Haddon, W., 1968. The changing approach to epidemiology, prevention and amelioration of trauma. American Journal of Public Health 58 (8), 1431–1438.

Hale, A.R., 1997. Introduction: The goals of event analysis. In: Hale, A.R. (Ed.), After the Event. From Accidents to Organisational Learning. Pergamon, Oxford.

Hale, A., 2000. Conditions of occurrence of major and minor accidents. 2me séance du séminaire Le risque de défaillance et son contrõle par les individus et les organisations, Gif sur Yvette.

Hale, A.R., Glendon, I., 1987. Individual Behaviour in the Control of Danger. Elsevier, Amsterdam.

Hale, A.R., Hovden, J., 1998. Management and culture: the third age of safety. A review of approaches to organizational aspects of safety, health and environment. In: Feyer, A.M., Williamson, A. (Eds.), Occupational Injury. Risk Prevention and Intervention. Taylor & Francis, London.

Hale, A.R., Bellamy, L.J., Gildenmund, F., Heming, B.H., Kriwan, B., 1997. Dynamic modelling of safety management. In: Presented at ESREL'97.

Hare, V.C., 1967. System Analysis: A Diagnostic Approach. Harcourt Brace & World, New York.

Harrisson, D., Legendre, C., 2003. Technological innovations, organizational change and workplace accident prevention. Safety Science 41, 319–338.

Heinrich, H.W., 1931. Industrial Accident Prevention. McGraw-Hill, New York.

Hendrick, K., Benner, L., 1987. Investigating Accidents with STEP. M. Dekker, New York.

Herrera, I.A., Hovden, J., 2008. Leading indicators applied to maintenance in the framework of resilience engineering: a conceptual approach. In: Proceedings of the third Symposium on Resilience Engineering.

Hollnagel, E., 1998. Cognitive Reliability and Error Analysis Method (CREAM). Elsevier Science Inc., New York.

Hollnagel, E., 2004. Barrier and Accident Prevention. Ashgate, Hampshire, England.

Hollnagel, E., 2008. The Changing Nature of Risks. Ecole des Mines de Paris, Sophia Antipolis, France.

Hollnagel, E., Woods, D.D., Leveson, N., 2006. Resilience Engineering. Concepts and Precepts. Ashgate, Aldershot.

Hopkins, A., 2007. Thinking about Process Safety Indicators. In: Presented at the Oil and Gas Industry Conference, Manchester, England, November 2007.

Hovden, J., 1984. Behøver vi ulykkesmodeller? Fordeler og ulemper ved ulike modeller. In Norwegian [Do we need accident models?] In Fjärde Nordiska Olycksfallsforskningsseminariet, Technical Research Centre, Symposium, 55, Espoo.

Johnson, W.G., 1980. MORT Safety Assurance System. Marcel Dekker, New York.

Jørgensen, K., 1985. Forskning om arbejdsulykker. In Danish. [Reserach on occupational accidents]. In Hovden, J. et al. (Eds.), nordiske konferanse for ulykkesforskere. 5. SINTEF, Trondheim.

Kjellén, U., 2000. Prevention of Accidents Through Experience Feedback. Taylor & Francis, London.

Kjellén, U., Hovden, J., 1993. Reducing risks by deviation control – a retrospection into a research strategy. Safety Science 16, 417–438.

Kjellén, U., Larsson, T.J., 1981. Investigation accidents and reducing risks – a dynamic approach. Journal of Occupational Accidents 3, 129–140.

LaPorte, T.R., Consolini, P.M., 1991. Working in practice but not in theory: theoretical challenges of high-reliability organizations. Journal of Public Administration Research and Theory 1, 19–47.

Leveson, N., 2001. Evaluating Accident Models using Recent Aerospace Accidents, Technical Report, MIT Dept. of Aeronautics and Astronautics. <http://sunnyday.mit.edu/accidents>.

Leveson, N., 2004. A new accident model for engineering safer systems. Safety Science 42, 237–270.

Lundberg, J., Rollenhagen, C., Hollnagel, E., in press. What-you-look-for-is-what-you-find – the consequences of underlying accident models in eight accident investigation manuals. Safety Science. doi:10.1016/j.ssci.2009.01.004.

Marais, K., Dulac, N., Leveson, N., 2004. Beyond normal accidents and high reliability organizations: the need for an alternative approach to safety in complex systems. In: Engineering Systems Division Symposium. MIT, Cambridge, MA.

Mayhew, C., Quinlan, M., Ferris, R., 1997. The effects of subcontracting/outsourcing on occupational health and safety: survey evidence from four Australian industries. Safety Science 25 (1–3), 163–178.

Mearns, K., Yule, S., 2009. The role of national culture in determine safety performance. Challenges for the global oil and gas industry. Safety Science 47, 777–785.

Op de Beeck, R., van Heuverswyn, K., 2002. New Trends in Accident Prevention Due to the Changing World of Work. Report, European Agency for Safety and Health at Work, Bilbao. <http://osha.europa.eu/en/publications/reports/208>.

Perduco, 2007. En bedriftsundersøkelse om arbeidskraft. Utarbeidet for NAV-EURES. [A company survey on manpower, the Norwegian Labour and Welfare Administration]. <www.nav.no/binary?id=805394484&download=true> in Norwegian.

Perrow, C., 1984. Normal accidents. In: Living with High-Risk Technologies. Princeton University Press, Princeton, NJ.

Perrow, C., 2007. The Next Catastrophe: Reducing our Vulnerabilities to Natural, Industrial, and Terrorist Disasters. Princeton University Press, Princeton, NJ.

Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem. Safety Science 27 (2–3), 183–213.

Reason, J., 1997. Managing the Risks of Organizational Accidents. Ashgate, Aldershot.

Reason, J., 2008. The Human Contribution. Unsafe Acts, Accidents and Heroic Recoveries. Ashgate Publishing, Aldershot.

Reason, J.T., Shotton, R., Wagenaar, W.A., Hudson, P.T.W., Groeneweg, J., 1988. TRIPOD: A Principle Basis for Safer Operations. Report for Shell International, Univ. of Manchester and Univ. of Leiden.

Rosness, R. Guttomersen, G., Steiro, T., Tinmannsvik, R, Herrera, I.A., 2004. Organisational Accidents and Resilient Organisations: Five Perspectives. SINTEF Report STF38 A 04403, Trondheim, Norway.

Saari, J., 2001. Accident Prevention Today. Magazine of the European Agency for Safety and Health at Work, No. 4, 2001.

Sklet, S., 2004. Comparison of some selected methods for accident investigation. Journal of Hazardous Materials 11 (1–3), 29–37.

Touminen, R., Saari, J., 1982. A model for analysis of accidents and its application. Journal of Occupational Accidents 4, 263–273.

Turner, B.A., 1978. Man-made Disasters. Wykeham, London.

Weick, K.E., Sutcliffe, K.M., 2007. Managing the Unexpected: Resilient Performance in an Age of Uncertainty. Jossey-Bass, San Francisco, CA.

Westrum, R., 1993. Cultures with requisite imagination. In: Wise, J.A. et al. (Eds.), Verification and Validation in Complex Man–Machine Systems. Springer, New York.

Wilde, G.J.S., 1998. Risk homeostasis theory: an overview. Injury Prevention 4, 89–91.

Wilpert, B., 2009. Impact of globalization on human work. Safety Science 47, 727–732.

Wreathall, J., 2001. Final Report on Leading Indicators of Human Performance. US Department of Energy Washington, DC, EPRI, Palo Alto.

Zuboff, S., 1988. In the Age of the Smart Machine. The Future of Work and Power. Basic Books, New York.

Paper II

Building Safety Indicators, Part 1 - Theoretical foundation

Safety Science (2011)
vol 49, pp. 148-161

Safety Science 49 (2011) 148–161

Contents lists available at ScienceDirect

Safety Science

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

Building Safety indicators: Part 1 – Theoretical foundation

K. Øien [a], I.B. Utne [b,*], I.A. Herrera [b]

[a] SINTEF Technology and Society, Safety Research, Trondheim, Norway
[b] Department of Production and Quality Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

ARTICLE INFO

Article history:
Received 26 June 2009
Received in revised form 21 April 2010
Accepted 16 May 2010

Keywords:
Indicators
Safety
Risk
Risk management

ABSTRACT

Development of early warning indicators to prevent major accidents – to 'build safety' – should rest on a sound theoretical foundation, including basic concepts, main perspectives and past developments, as well as an overview of the present status and ongoing research. In this paper we have established the theoretical basis for development of indicators used as early warnings of major accidents. Extensive work on indicators have been carried out in the past, and this could and should have been better utilized by industry, e.g., by focusing more on major hazard indicators, and less on personal safety indicators. Recent discussions about safety indicators have focused on the distinction between leading and lagging indicators; however, a discussion on terms should not impede the development of useful indicators that can provide (early) warnings about potential major accidents.

1. Introduction

One strategy to avoid accidents is to be continuously vigilant through the use of indicators. Often, hindsight has shown that if signals or early warnings had been detected and managed in advance, the unwanted event could have been prevented. This includes, e.g., the accident at the Esso natural gas plant in Longford, Australia in 1998, killing two workers (Hopkins, 2000), and the accident at the BP Texas City refinery in 2005, killing 15 workers (Baker et al., 2007). Recognizing signals/early warnings through the use of proactive safety indicators will reduce the risk of such major accidents.

Building Safety[1] is a research project which addresses safety opportunities and challenges in petroleum exploration and production in the northern regions, with emphasis on the Goliat field outside the northern coast of Norway. Oil and gas exploration in the Barents Sea is controversial, and avoiding major accidents is critical in order to gain political acceptance. One of the main research issues in Building Safety is to develop new models and methods for the establishment of indicators, which can unveil early warnings of major accidents (SINTEF, 2010). Goliat will be the first oil development in the Barents Sea with planned production start in 2013–2014, and it is subject to strict environmental requirements. Early warning indicators will improve the ability to produce oil and gas without any harmful spills.

The purpose of this paper is to establish the theoretical basis for development of indicators used as early warnings of major accidents. This includes basic concepts, main perspectives and the past developments of major hazard indicators, and constitutes Part 1 of the research.

More recent developments are included in Part 2; a separate follow-up paper (Øien et al., 2010). Here, we have presented examples of applications and practices from selected major hazard industries.

Presently, there are many discussions about the use and development of major hazard indicators (e.g., Hopkins, 2009a; HSE and CIA, 2006[2]; Duijm et al., 2008; Grabowski et al., 2007; Saqib and Siddiqi, 2008; Körvers and Sonnemans, 2008; Osmundsen et al., 2008), but there are few attempts, if any, to structure and summarize past work in this research field. We believe that a theoretical basis, including a thorough review of developments in the past (i.e., 'the history of safety indicators') as well as an overview of the current status, applications and practices, constitute a necessary foundation for future development of safety indicators.

Our own research on safety/risk indicators covers both developments in the past (e.g., Øien et al., 1996, 1997, 1998; Øien and Sklet, 1999a,b, 2000; Øien, 2001a,b,c) and recent and ongoing developments (e.g., Øien, 2008; Øien et al., 2010). This constitutes a useful basis for the structuring of perspectives and developments in this field of research, which started in the early 1980s.

A recent discussion about safety indicators focuses on the distinction between leading and lagging indicators (Hopkins, 2009a; Hale, 2009a). Apart from the fact that leading indicators are of particular interest for the development of early warning indicators,

---

* Corresponding author.
E-mail address: ingrid.b.utne@ntnu.no (I.B. Utne).

[1] http://www.sintef.no/buildingsafety.

[2] CIA – Chemical Industries Association.

0925-7535/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ssci.2010.05.012

this debate demonstrates the need for a theoretical foundation and good knowledge of past developments.

In this paper, we focus on early warnings in the form of safety or risk indicators with, e.g., the following properties:

- They provide numerical values (such as a number or a ratio).
- The indicators are updated at regular intervals.
- They only cover some selected determinants of overall safety or risk, in order to have a manageable set of indicators.

We do not cover other types of early warnings, such as safety bulletins providing information about events experienced in other companies, or continuous (on-line) control systems, such as risk monitoring systems or process control systems. However, information from, e.g., process control systems can in principle also be used as input for safety indicators.

Furthermore, we do not include personal safety (occupational accidents), or research related solely to classification and evaluation of organizational factors. The emphasis is on industries exposed to major hazards, such as the nuclear power industry, the chemical process industry and the petroleum industry.

This paper is divided into five main sections: Section 2 discusses important concepts and perspectives. Section 3 covers the research related to safety indicators in general and their historical development. Section 4 covers the latest discussions on safety indicators, with main focus on a special issue of Safety Science on process safety indicators (Hale, 2009a). Conclusions are stated in Section 5.

2. Concepts and perspectives

Research on indicators is carried out in the borderline between social and natural science. There are diverse opinions about measurements of safety or risk within and across these disciplines. In this section, we discuss the concepts safety indicator and risk indicator, and we focus on two main perspectives of the research on indicators.

2.1. Concepts

The term indicator may be used in several ways, which means that there exist many definitions. Some definitions of indicators are:

- ‘A safety performance indicator is a means for measuring the changes over time in the level of safety (related to chemical accident prevention, preparedness and response), as the result of actions taken’ (OECD, 2003).
- ‘A safety indicator is an observable characteristic of an operational nuclear power plant unit, presumed to bear a positive correlation with the safety of the reactor. The safety indicators have been selected, among other means, for the purpose of supervision of safety. The safety indicators can be related to defense lines according to defense-in-depth such as physical barriers and safety functions’ (Holmberg et al., 1994).
- ‘An indicator is a measurable/operational variable that can be used to describe the condition of a broader phenomenon or aspect of reality’ (Øien, 2001b).

The last definition is a broad definition that also covers the first two definitions. We will pursue and explain this last definition in some detail, not only to have a thorough understanding of what an indicator is, but also to know what an indicator is not, because there has been an increasing tendency to put ‘everything’ under the umbrella of ‘indicators’.

The last definition is based on the combination of the following two definitions; ‘an indicator is a measurable/operational definition of a theoretical variable, i.e., it is an operational variable’ (Hellevik, 1999); ‘indicators are measures used to describe the condition of a broader phenomenon or aspect of reality’ (Gray and Wiedemann, 1999).

A risk influencing factor (RIF) is defined as ‘an aspect (event/condition) of a system or an activity that affects the risk level of this system or activity’ (Øien, 2001c). A given RIF (e.g., an organizational factor) might not be directly measurable. This is denoted ‘the measuring problem’ within social science research methodology (Hellevik, 1999). Instead we need an operational definition of the RIF (denoted an ‘operational variable’) that represents the theoretical variable. This is illustrated in Fig. 1 (Øien, 2001c).

This operational variable is what we denote an indicator (Hellevik, 1999). The indicator is not the RIF itself, just a measurable representation of the RIF. Important to notice is that the measuring of one RIF may be performed by a set of indicators. Making a (theoretical) variable operational means giving an instruction on how to measure the theoretically defined variable (Grenness, 1997). This transformation is both controversial and a possible source of errors. Gray and Wiedemann (1999) discuss this problem and state that ‘the basic, inherent difficulty with indicators is that they are selective. They each represent one measure of one aspect of any situation’. This means that there is always room for discussion and even disagreement about whether they really represent what one wants to measure; whether people want to measure the same thing; and whether the measure is understandable to ‘non-experts’.

The terms safety indicator and risk indicator are sometimes used interchangeable, but it may be appropriate to distinguish between these two terms: If the RIFs are included in a risk model (a logic system structure), such as a probabilistic risk assessment (PRA), then it is possible to determine (within the limitations of the model) the effect on risk (measured by some risk metric) of a change in the indicator value of a given RIF. We then talk about risk indicators or risk-based indicators. If we do not have such a risk model, we can still identify some of the same factors and also establish some of the same indicators. However, the effect on safety has to be related to some other measures (than risk metrics), such as number of accidents or incidents, or purely qualitatively without quantifying safety. The indicators and the corresponding factors are then often selected, based on either an assumed effect on safety, or through correlation. These indicators should be denoted safety indicators to avoid confusion with risk-based indicators, and the corresponding factors are most appropriately denoted safety influencing factors. Our definition of risk indicator is then; ‘a risk indicator is a measurable/operational definition of

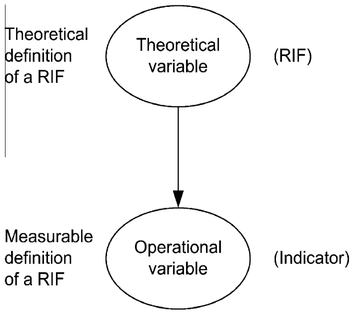


Fig. 1. General measurement model (Øien, 2001c).

a RIF' (Øien, 2001c) in which it is implicitly implied that the RIF is linked to a risk metric through a risk model.

Risk indicators are developed from a risk based approach (Øien, 2001b), whereas safety indicators may be developed from various approaches, such as a safety performance based approach (HSE and CIA, 2006), an incident based approach (Øien, 2008) or a resilience based approach (Øien et al., 2010). Thus, a probabilistic risk assessment is only one basis for the development of indicators.

From the early days of research and development of safety and risk indicators, an indicator has usually been restricted to a numerical value such as a number, a ratio, etc., and the updating of the indicator values are rather frequent. Infrequent assessments of safety, using questionnaires, are covered by safety audit methods and tools, and although these questions sometimes are referred to as 'indicators', they belong to the domain of safety audit systems. These systems are not part of the research on safety and risk indicators, and are therefore not included in this paper.

2.2. Perspectives of the research on indicators

This paper is structured according to a combination of two perspectives. The first perspective is related to the development in the search for causes of accidents, moving from technical, to human, and further to organizational causes, i.e., further back in the causal chain (Leveson, 2004). However, this perspective is viewed in the light of a second perspective, which is the question of a predictive versus a retrospective view. It makes a big difference whether we try to predict the possibility of having a major accident 'tomorrow', including all possible causes, or if we 'only' try to establish the causes after-the-event (in retrospect).

Based on these two presented perspectives; the technical-human-organizational, and the predictive-versus-retrospective, we establish a conceptual model in order to structure and illustrate the previous research. This simplified model is shown in Fig. 2. Only some possible topics related to quantitative risk assessment (QRA) are illustrated here, e.g., fault tree analysis (FTA), event tree analysis (ETA) and human reliability analysis (HRA).

The 'technical–human–organizational' perspective is illustrated horizontally and the 'retrospective-versus-predictive' perspective is illustrated vertically. For retrospective purposes, such as accident investigation, organizational factors have been included at least since the Three Mile Island accident in 1979. For predictive purposes, organizational factors have only more recently been included or attempted to be included.

If we limit the understanding of organizational factors to accident investigation, that is, hindsight, then we can talk about different 'ages' in the development moving from technical, to human, and further to organizational causes. We can even look for more remote causes as external pressure and regulation. Wilpert (2000) suggests that we have entered the period of 'inter-organizational relationships'. However, Reason (1997) raises the question whether 'the pendulum has swung too far' in our search for the origins of major accidents. This search should add explanatory, predictive and/or remedial value, but particularly the added remedial value is questionable when we move far back in the causal chain, and we should concentrate on the changeable and controllable.

Another potential pitfall 'rushing' further to even more remote causes is the impression that we now can cope with organizational causes. This is both a false impression and a potentially dangerous one. The organizational factors' effect on safety/risk is by no means well understood. One token of this can be found in Wilpert (2000). There is a general lack of consensus regarding the classification of organizational factors, and none are identical. About 160 different factors have been suggested in those 12 classifications assessed by Wilpert (each of the classifications usually consists of 10–20 factors).

For the prediction of risk, as for accident investigation, we can talk about a development from technical, to human, and even to organizational causes. This does not imply that all features of risk assessment can be classified according to a technical–human–organizational 'scheme'. There are features that cut across these aspects, such as dependent failure analysis and uncertainty analysis. However, some aspects can be attached to primarily one of the causal categories, such as human reliability analysis (HRA) attached to the human causes of accidents.

The proactive approaches for the assessment of underlying factors' effect on safety or risk can be illustrated by reversing the arrows in Fig. 2, illustrated and simplified in Fig. 3.

There has been a long tradition, especially within social sciences, to assess the effect of organizational factors on safety, and this is shown in the upper part of Fig. 3. In addition, we have modified the 'technical–human' aspects and concentrated on the organizational aspect. A major obstacle to the assessment of organizational factors' effect on safety with respect to industrial accidents is that these accidents are so rare that a direct measure of safety is not possible. Instead indirect safety measures are sought, termed performance indicators, safety indicators, safety performance indicators, direct performance indicators, indirect programmatic performance indicators, etc. These safety performance indicators are either assumed to have an effect on safety, or efforts are put into establishing correlation between the indicators and 'safety'.

Within the probabilistic approach (lower part of Fig. 3) the emphasis is not on measuring the effect of organizational factors on risk as an isolated effort, since it is dealing with potential accidents and then it does not matter how rare the events are (except with respect to uncertainty and credibility). Risk is estimated based on the existing risk model, and the focus is on how this risk estimate changes and perhaps becomes more correct when the organizational factors are explicitly accounted for. Davoudian et al. (1994a) claim that risk is underestimated if effects from organizational factors are not accounted for.

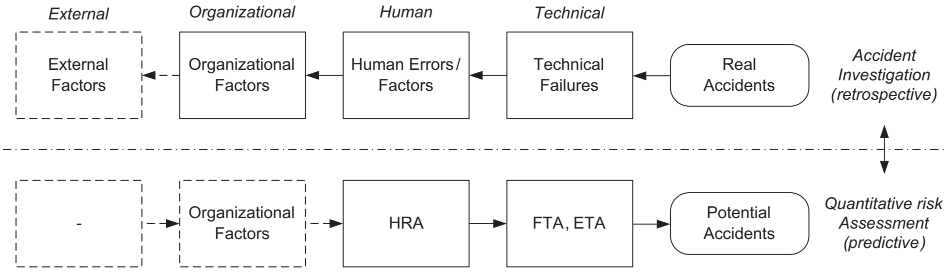


Fig. 2. Accident investigation versus predictive assessment (Øien, 2001b).

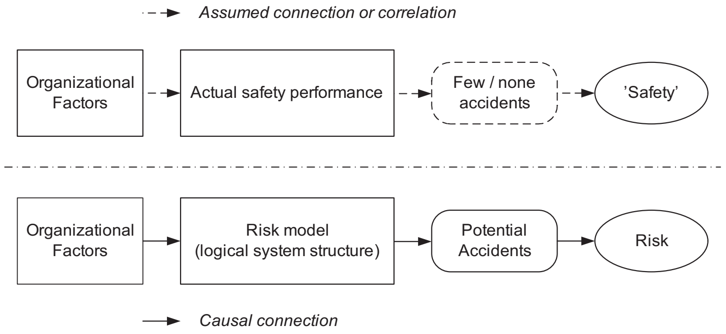


Fig. 3. Proactive approaches to assess organizational factors' effect on safety/risk (Øien, 2001b).

More recently work based on a resilience engineering perspective has also been carried out (Herrera and Woltjer, 2008; Leveson, 2004) in which non-causal models of accidents and systemic causes are applied.

3. Indicator development

The research on indicators started with the need to measure safety or risk. The main function of a measure of safety performance is to describe the safety level within an organization, establishment, or work unit. The term indicator in the safety field is rather new, but safety measurements were performed in the 1980s and before, but then with terms like index, rate, and measurements (Tarrants, 1980). Nowadays, the terms indicator and key performance indicators are commonly applied (e.g., Kjellén, 2000; Mearns, 2009).

The description of the indicator developments is structured as follows:

1. Work initiated by the United States' Nuclear Regulatory Commission.
2. WANO performance indicators.
3. Operational safety indicators.
4. Safety performance indicators.
5. Operator specific safety indicators.
6. Probabilistic indicators.
7. PSA based risk indicators.
8. Accident sequence precursors.
9. The resilience engineering perspective on indicators.

The nuclear power industry has been a key driver in the development of major hazard indicators, which is also reflected here. In addition, there have been some contributions from the chemical process industry and the offshore petroleum industry (including our own research).

The last subsection constitutes a transition from the developments in the past to ongoing developments. It is noteworthy that one of the key players in the past developments, Wreathall et al. (1990), is also a key player in the ongoing developments (Wreathall, 2009).

3.1. Work initiated by the United States' Nuclear Regulatory Commission

From the early 1980s, the US Nuclear Regulatory Commission (US NRC) initiated a lot of work on the effect of organizational factors on safety. Most of this work belongs to the upper part of Fig. 3, and uses indicators for measurement. Even though there are exceptions to this, we treat the US NRC initiated work collectively in this section.

Osborn et al. (1983a,b) carried out a major literature review in a project with the objective of carrying out initial research on the feasibility of applying organizational factors in nuclear power plant safety assessment. They state that 'instead of working from technically identified problems or functions back toward causes, the [organizational] perspective shows patterns which predict and/or precede success and failure'. In a way this depicts the transition from a retrospective perspective (upper part of Fig. 2) to a predictive perspective (upper part of Fig. 3). Another interesting statement is that 'NRC, utility, and industry officials themselves agree that utility management is important even as they disagree over which factors are important and whether NRC should develop new regulations in this area'. This is probably still a valid statement, and reflects the 'political tension' in this field of research.

Osborn et al. (1983a,b) found that the organizational perspective on safety was a new one, and that the existing studies within other industries focused on the individual worker in terms of the causes and consequences of unsafe behavior. They further found that while both management and organization had been identified as root causes for many problems within the industry, analysts could rarely point to measurable factors that could be logically linked to safety. Finally, they recommended that the lack of empirical organizational analyses of the nuclear industry should be rectified immediately, which led to the next NRC project.

In the project 'Initial Empirical Analysis of Nuclear Power Plant Organization and Its Effect on Safety Performance' (Olson et al., 1984), the objective was to assist the NRC in developing a technical basis for evaluating management capabilities of utilities seeking nuclear power plant (NPP) operating licenses. In this study they found that organizational structure (the way the work of the organization is divided, administered and coordinated) appeared to be an important predictor of plant safety performance. One of the conclusions was that 'the overall results show that plants with better developed coordination mechanisms, shorter vertical hierarchy, and a greater number of departments tend to perform more safely'. They recommended that the feasibility of developing indicators more directly relevant to sub-areas of the plant should be explored.

In 1985, Olson et al. (1985) conducted an initial attempt to derive and validate, on an empirical basis and by using existing performance data, objective indicators of safety-related performance for use in the assessment of organizational factors. They state that 'few topics in the nuclear industry are as constantly and hotly debated as that of how to assess the safety performance of plants. One reason for the lack of agreement about measuring and evaluating safety performance is that direct measures of safety are not usually available'. This has, as mentioned before, led to the search for indirect safety measures, not only in the nuclear industry, but in other high-hazard industries as well.

In lack of direct measures of safety, Olson et al. (1985) defined 'penultimate' measures of safety (e.g., number of potentially significant events), and further used these in an attempt to validate other direct performance measures. Thus, they did not try to develop organizational performance measures, but rather 'outcomes' of organizational performance. The direct performance indicators have actually much in common with the later developed INPO[3] and WANO[4] indicators. The conclusion of this study was that 'the analysis lends considerable support to the position that measures derived from LERs,[5] operating and outage data, and violations data are useful in assessing Licensee performance and in predicting subsequent performance on the penultimate safety measures'.

Olson et al. (1988) also carried out a project that was both developmental and empirical. This time efforts were put into modeling the connection between organizational factors and safety, and validating the different elements of the model via correlation. The objective of the study was to develop and validate a set of programmatic performance indicators (PPIs)[6] of Licensee performance for monitoring safety performance of operating Nuclear Power Plants (NPP) and to assist the regulatory decision-making process. The main areas of concern were 'maintenance', 'management and administration', and 'training and experience'. One specific feature of this study was that they attempted to not only correlate the programmatic indicators with direct performance indicators, but also to correlate them with a kind of 'safety audit' of the plant, termed Systematic Assessment of Licensee Performance (SALP) and carried out by NRC staff. The results showed that for most of the evaluated programmatic performance indicators, there were either data problems or lack of data. Thus, the overall conclusion was that 'the development and implementation of new data collection techniques and requirements will be necessary in order to develop preferred indicators of programmatic performance'. Unfortunately, this is probably valid across all industries, even today.

Marcus et al. (1990) performed a project for the NRC, where the objective was to develop organizational effectiveness indicators for use by NRC. The purpose of such indicators can be drawn from the following statement: 'If good indicators of management and organization were available, NRC, as well as plant and utility management, would be in a better position to anticipate potential problems and to do something about them'. In this study Marcus et al. (1990) developed a logical framework linking the management and organizational factor classes (environment, context, organizational governance, organizational design, and emergent process) to intermediate outcomes (efficiency, compliance, quality, and innovation) and further to direct performance indicators. Empirical (correlation) analysis was carried out with respect to organizational indicators' relation to the direct performance indicators. This resulted in two additional candidate management indicators: 'utility resources', and 'lagged recognition and correction of problems'.

Particular attention was paid to the goal conflict between safety and efficiency. The results showed that 'profitability, in particular earlier profitability, tended to be significantly positively related to the safety indicators'. 'This confirms the single proposition that utilities have to be able to afford safety, that safety costs money, and without adequate resources, it cannot be achieved'. A final statement about this conflict was 'for regulators, managing and operating NPPs safety may be the 'true goal' with efficiency falling to the status of a critical constraint, but for utility managers efficiency is more likely to be the main goal with safety occupying the status of a critical constraint'.

Prior to this last project, NRC initiated two other more probabilistically oriented projects. The first of these was carried out by Boccio et al. (1989). The objective of this project was to develop more responsive indicators of system performance using available data basically associated with safety performance. It was probably one of the first attempts to develop risk-based indicators. However, they focused on just one specific indicator being the Safety System Function Trend (SSFT). This risk-based indicator was suggested to replace the previous safety performance indicator; Safety System Failures (SSF). Focusing on probability of component failures instead of just actual performance of systems, provided indications of a potential declining system performance before loss of system function was observed. The conclusion was that 'the SSFT indicators correlated with the SSF indicators and provided much faster response. Because the SSFT indicator is also risk-based, it can provide more direct measures of impacts to risk'. However, like the study of Olson et al. (1984), this study did not treat organizational aspects.

The second probabilistically oriented 'indicator project' was carried out by Wreathall et al. (1990). Here, the objective was to identify specific programmatic performance indicators related to nuclear plant maintenance. The purpose of using indicators in general is given in the following statement: 'Management and control of any operational enterprise, from landing a spacecraft to developing investment portfolios, require the development and tracking by indicators of the enterprise's performance throughout time. Properly developed indicators can serve as signals to management to allow them to take appropriate control action'. Based on a literature survey, Wreathall et al. (1990) started out with 78 candidate indicators. This list was screened down to a shortlist of nine indicators based on qualitative criteria, and after a quantitative evaluation (validation), they were left with two indicators. These were 'number of inadvertent emergency safety feature actuations, due to test and maintenance', and 'gross heat rate/daily power loss'. The probabilistic element of this study is that, instead of correlating the programmatic performance indicators with direct performance indicators, they used a risk model for validation and calculated the conditional core damage probabilities in a similar way as in accident sequence precursor (ASP) analyses.

Haber et al. (1991) developed the so-called Nuclear Organization and Management Analysis Concept (NOMAC). In the project ('Influence of Organizational Factors on Performance Reliability'), the initial objective was to identify methods, which had undergone substantive prior scrutiny that could be used in observing and assessing the impact of organizational factors on NPP safety, and that would provide products useful to US NRC staff, NPP personnel, and probabilistic risk assessment (PRA) practitioners. This work can be divided into three parts:

1. The first part is the 'organizational concept' that draws heavily on Mintzberg's model of 'a machine bureaucracy' (Mintzberg, 1979, 1983, 1988). This part covers the identification of organization and management factors.
2. The second part is the data collection methods. These include functional analysis, behavioral observation technique, and organizational culture assessment. This leads to qualitative rating of the organization and management factors, and not to quantitative measurement using performance indicators. In fact, there is no mention of indicators at all in this work. Thus, this part of NOMAC may be categorized as being a type of safety audit method.

---

[3] INPO – Institute for Nuclear Power Operations.

[4] WANO – World Association of Nuclear Operators.

[5] LER – Licensee Event Report.

[6] PPIs are indicators that assist in assessing the quality and performance of various programs, functions, and activities relating to the safety of the plant.

3. The third part is the data analysis leading to the three different applications: inspection activities, regulatory insight and PRA support. For regulatory insight, NOMAC was judged as being more objective than for example SALP.

The application of NOMAC to PRA leads to a probabilistic approach, and is in fact the starting point of what is later to become the Work Process Analysis Model (WPAM method) (Wu et al., 1991; Davoudian et al., 1994b). As part of the 'NOMAC project', work was subcontracted to the University of California at Los Angeles (UCLA), where a preliminary scheme for quantifying the impact of organizational factors in PRA was developed. It was suggested that the Success Likelihood Index Methodology – Multi Attribute Utility Decomposition Method (SLIM–MAUD) (Embrey et al., 1984) might be a viable technique to utilize in the integration of organizational and management factors into PRA.

Wreathall et al. (1992) developed the Integrated Safety Model (ISM) framework as a tool to integrate and discuss the relevance of performance indicators used by the US NRC. A key model in the ISM framework is the so-called diamond tree. This model is reapplied in a later project (Youngblood et al., 1999), which discusses an approach to performance-based regulatory oversight.

US NRC also initiated a project called 'Management and Organizational Factors in PRA' that led to the preliminary development of a framework named The Socio-Organizational Contribution to Risk Assessment and the Technical Evaluation of Systems (SOCRATES) (Gertman et al., 1998; Blackman et al., 1998). Surprisingly, the US NRC terminated the project and no final report exists. Like the NOMAC work, this is neither a 'safety performance indicator' type of project. The US NRC has continued its work on performance indicators and has emphasized the importance of developing risk-based performance indicators (RBPIs) (Baranowsky et al., 1999), but the main emphasis is not on the organizational factors.

There is also a close relationship between the RBPIs and the ASP program (Belles et al., 1998); the importance of the RBPIs is determined based on ASP analyses using the simplified plant analysis risk (SPAR) models (Long et al., 1998). ASP analyses are in general focusing on operational events, not on organizational inadequacies. One exception to this was the previous mentioned work by Wreathall et al. (1990).

3.2. WANO performance indicators

The international nuclear power community formed the World Association of Nuclear Operators (WANO) after the Chernobyl accident in 1986. The worldwide standardization of the nuclear power plant performance indicators was seen by WANO as one of the important aims. In 1990, after an international development effort, WANO established for international use a set of 10 performance indicators in the areas of nuclear power plant safety, reliability, efficiency, and personnel safety (Holmberg et al., 1994):

1. Unit capability factor.
2. Unplanned capability loss factor.
3. Unplanned automatic scrams per 7000 h critical.
4. Safety system performance.
5. Thermal performance.
6. Fuel reliability.
7. Collective radiation exposure.
8. Volume of lower level solid radioactive waste.
9. Chemistry index.
10. Lost-time accident rate.

Every plant operator has been given a detailed use description for these indicators. Still, concerns have been raised regarding the extent of safety emphasis in the WANO indicator set. A practical problem for the operators is to sort out the most important information from the large flow that comes in every day. Improved learning from experience can be achieved by using more power plant specific safety indicators. Thus, further development and implementation of more detailed and plant specific indicators (for surveying safety critical activities and uncovering deviations in the power plant) were considered useful among nuclear plant operators as well as regulators (IAEA, 1991). However, for most of the suggested indicators, the degree of correlation between the indicators and safety/risk is unknown (Øien and Sklet, 1999b).

The WANO indicators may also be classified as direct indicators, that is, outcome indicators that utilize different types of experience data. Whereas work has continued in developing direct indicators, emphasis has also been put into the development of indicators that can give early warnings. These early warning type of indicators are classified as indirect indicators which can measure the performance of the functional units within an organization, such as operation, maintenance, training, and engineering support (Holmberg et al., 1994). Often, the indirect indicators are called predictive indicators.

3.3. Operational safety indicators

A work group in IAEA has developed a set of indicators for surveillance of the operational safety within a power plant (IAEA, 1999a). They present a framework for identifying performance indicators for circumstances related to safety, as well as indicators for economic issues, as economy has a huge impact on the safety level. The safety indicators are supposed to show the trends and developments over time in order to give the operators a chance to analyse the causes to changes. The work group also emphasizes the importance of adapting the indicators to plant specific conditions, and thus the proposed indicators are meant to function as a framework for this work.

The framework has a hierarchical structure, in which the topmost level is the power plant's operational safety ('NPP operational safety performance'). Level 2 consists of operational safety attributes, from which 'operational safety performance indicators' can be identified. The attributes are:

- The power plant operates smoothly.
- The power plant operates with low risk.
- Effective plant management processes.

At the level below the attributes, seven paramount indicators were worked out to evaluate the relevant safety aspects. Further, 14 strategic indicators were developed, resulting in 38 specific indicators. Ideally, the specific indicators should be able to reveal potential problem areas so that detailed analyses can be initiated, and risk reducing efforts implemented, before safety is further reduced.

Morenõ et al. (1998) describe a framework for establishing operational safety indicators within the chemical process industry. They emphasize that the purpose of using such indicators is to identify any negative development in safety in order to implement accident preventive efforts at a facility.

The framework is prepared for indicators at two levels:

- Paramount indicators (High Level Indicators – HLI).
- Specific indicators (Low Level Indicators – LLI).

The paramount indicators are to be used by management as a basis for making decisions. The specific indicators constitute a foundation for a pyramid structure, and are developed to get information about specific conditions regarding technical systems, operational and organizational conditions, and may be used by 'systems responsible persons', or management at different levels in the organization. The complete set of operational safety indicators (OSIS) consists of a large amount of LLI and a small amount of HLI.

Indicators should be established for six different functional areas:

1. Management, organization, and administration.
2. Design of facility and processes.
3. Training and qualification.
4. Operation.
5. Maintenance.
6. Emergency preparedness planning.

When establishing a preliminary set of indicators, it is either possible to use existing parameters, or to establish a new set of indicators. All indicators are classified according to functional area, as previously described. Further, the indicators should be evaluated to determine whether they are feasible as paramount indicators directly, or which specific indicators that may be combined into paramount indicators. Identification of new indicators should be based on assessments of the existing operational experience and search for underlying causes to unwanted events that have occurred.

The operational safety indicators cover technical, operational, and organizational conditions, but the relationship between the indicators and the risk level is unclear. There have not been any attempts to link these indicators to a risk model in order to quantify the effect on the risk level.

3.4. Safety performance indicators

In Scandinavia, a safety indicator project has been carried out, described by Holmberg et al. (1994). Some of the reasons for carrying out this project were the uncertainty about the correlation between the WANO indicators and safety, and the usefulness of having plant specific indicators. Two types of indicators were established; direct indicators and indirect indicators. Preferably, these indicators can be used to evaluate safety by assessing the performance level, and by evaluating the performance trend. The former implies comparison of indicator values to a pre-determined reference value. The latter necessitates a trend analysis to reveal significant increasing or decreasing changes in the indicator values.

In the Nordic project, the barriers in the defense-in-depth strategy along with the risk analysis were identified as a reasonable framework for identification and structuring of the safety performance areas. The defense-in-depth strategy is illustrated in Fig. 4.

Fig. 4 illustrates several barriers preventing the energy process from getting out of control at a nuclear power plant and becoming a threat to human beings. Threats to the system, such as component failure, human failure, and external influences are defeated by this strategy. The physical barriers shall be able to uphold their integrity through a set of safety barriers.

The performance areas defined, based on the defense-in-depth strategy, were:

- Safety management (Level 1 safety barrier).
- Control of operation (Level 2 safety barrier).
- Safety functions (Level 3 safety barrier).
- Physical barriers (Physical barriers 1–4).

Levels 4 and 5 (crisis management and emergency preparedness) were not used as basis for development of indicators. About one hundred safety indicators have been collected and further developed, and thereafter specified by name, function, purpose, definition, need for data, use, and results. Examples of such indicators can be found in Table 1.

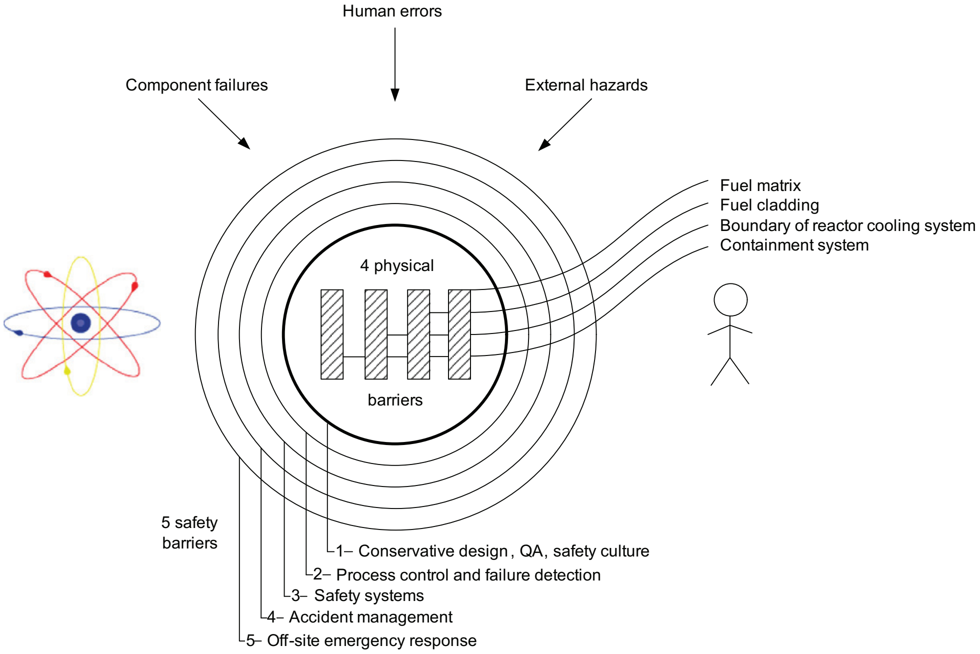


Fig. 4. Illustration of the defense-in-depth strategy (based on IAEA, 1988).

Table 1
Examples of safety indicators (selected from Holmberg et al., 1994).

Safety management	Control of operation	Safety functions	Physical barriers
Recurrent fault modes	Transient index	Safety systems performance*	Tightness index
Maintenance ambition index	Mean time between repairs of components	Common cause failures	Crack index
Safety issues backlog	Unplanned capability loss factor*	Length of component unplanned outage	Fuel reliability index*

* Also WANO indicator.

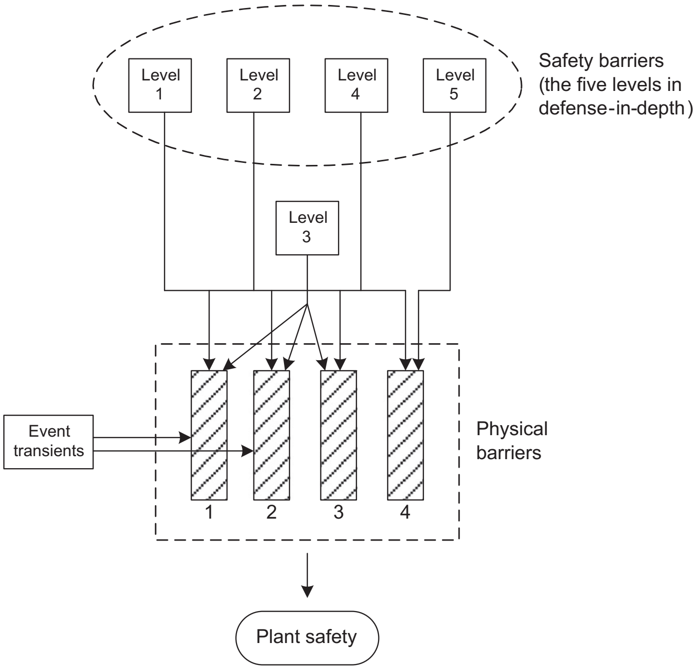


Fig. 5. Vattenfall's framework for defining indicators (based on Holmberg et al., 1994).

3.5. Operator specific safety indicators

Vattenfall developed operator specific safety indicators in cooperation with the Nordic project described by Holmberg et al. (1994). The framework for defining Vattenfall's indicators is illustrated in Fig. 5.

Fig. 5 may be viewed as a further development of Fig. 4, of which the four physical barriers should resist initiating events threatening the safety of the nuclear power plant. The figure shows what kind of safety barriers that influence the physical barriers. All safety barriers, except the safety systems (barrier, level 3), may be considered organizational or administrative systems, i.e., non-physical systems.

The physical barriers, together with the safety systems, give possibilities for establishment of direct indicators, while the relationship between the organizational/administrative barriers and safety is indirect, and mainly possible to monitor through indirect indicators.

Note that the 'safety systems' (level 3 in defense-in-depth) consist of equipment (hardware) monitored by use of direct indicators. The identified indicators are shown in Table 2.

Nine of the indicators were accepted for common use following the development project, while six of these have been used as a tool for communication between the operator's central management and each of the power plant units. An analysis and evaluation of these indicators at the different plant units are reported quarterly to the power plants and central management.

A positive correlation between the indicators and safety is assumed, even though the real effect on safety has not been evaluated. However, the use of these indicators has revealed safety conditions which would have remained hidden in the large amount of reports and data available. The safety indicators have a potential to improve the experiential learning.

3.6. Probabilistic indicators

The Swedish Nuclear Power Inspectorate (SKI) manages an experience database, STAGBAS II. Statistical diagrams in the so-called 'event catalogue' visualize the annual number of events originating from different safety functions, safety systems and safety components. From this catalogue, 'indicator candidates' are identified by evaluating trends.

The safety relevance for each candidate (e.g., isolation valves), as unit specific indicators, is also assessed by using the plant specific risk analysis to express their significance related to the total risk. The underlying failure cause of the candidates is investigated in detail to identify the real problem area and to confirm trends. Afterwards, the unit specific indicators can be established. The following 'observation areas' are common for several plant units:

Table 2
Operator specific indicators (based on Holmberg et al., 1994).

Safety indicators			Type	Routine use	Reported each 4 month
Events		Unplanned automatic scrams	D	x	x
		Transient index	D/I	x	x
Physical barriers	Fuel cladding (1)	Fuel reliability	D	x	
	Primary circuit	Chemistry index	D/I	x	
	Pressure boundary (2)	Crack index	D		
	Containment	Tightness index	D	x	x
Safety barriers	Safety culture, QA (1)	QA index	I		
		Exemption index	I		
		LER significance index	D	x	x
		Recurrent failure index	D/I	x	
(‘Defense-in-depth’)	Control of operation (2)	Maintenance quality index	I		
		Maintenance ambition index	I		
		Work order management index	I		
		Unplanned capability index	D	x	
	Safety systems (3)	Safety system performance	D	x	x
		Valve failure index	D		

D – direct indicators. I-indirect indicators.

- Hydraulic scram system or control rod drives (reactivity control).
- Fire protection system.
- Electric power supply.
- Isolation valves.
- Core spray system and containment vessel spray system.

Thus, the probabilistic safety indicators are developed based on a detailed study of each of the ‘problem areas’. Note that these indicators measure a change, e.g., in isolation valves, where the effect on risk is evaluated through use of the plant specific risk analysis. This means that the connection between the indicators and safety (risk) is known, as well as the effect.

3.7. PSA based risk indicators

A Probabilistic Safety Assessment (PSA) for a nuclear power plant includes a lot of information about safety, at the same time as it quantifies the risk. The PSA can be used to identify conditions and establish indicators for those conditions most significant to safety. Fig. 6 shows a framework for establishing PSA based indicators suggested by IAEA (1999b). The structure is consistent with the framework for identifying operational safety indicators.

PSA based risk indicators are established at different levels. Level 1 indicators deal with the total risk of the facility. The total risk target depends on the scope of the analysis. For level 1 PSA,[7] the frequency of core damage (CDF) is used. Level 2 indicators should cover the possibility for undesirable events (i.e., the frequency of initiating events), the power plant’s inability to tackle incidents in a way that core damage does not occur (i.e., the probability for safety system failure), and the plant’s inability to deal with accidents (i.e., the probability of failure in systems that should prevent the release of radioactive materials following core damage).

Level 2 indicators can be divided into several sub levels. The primary indicator for the plant’s ability to tackle events is the probability of core damage for each initiating event. Further, the safety function unavailability[8] and system unavailability indicators are found. The unavailability can be further decomposed to compressor train and component level. A corresponding decomposition can be carried out for the plant’s ability to tackle accidents.

PSA based indicators are a safety information tool that may be used in several ways. Important aspects are long-term versus short-term, and retrospective versus predictive. The long-term use of risk indicators implies a focus on surveillance of the plant risk with respect to the historical development (average risk – CDF_A). A short-term use of risk indicators is about getting an overview of the instantaneous risk at the plant – CDF_I, due to changes in operational conditions and incidents occurred. Overview and use of instantaneous risk require a continuous evaluation of the risk picture of the plant. The prerequisite for use of long-term indicators is a ‘Living PSA (LPSA)’. To calculate instantaneous risk, a ‘Risk Monitor’ is needed.

Retrospective use of risk indicators implies documenting and analysing the risk development at the plant, due to incidents occurred, component failures, human failures, unavailability of different systems, maintenance, etc. The purpose is to give a ‘true’ picture of the risk at the plant in a given time period. Predictive use means that the PSA models are connected to the planning of activities (e.g., maintenance), configuration changes, etc., to minimize the planned risk.

The PSA based risk indicators are related to the nuclear power plant risk, and may be used to monitor and follow-up changes in the risk. They presuppose that the current values for all parameters in the risk analysis model (LPSA model) are updated continuously so that the LPSA model can be used to calculate the risk. The risk can either be calculated as an average risk or instantaneous risk at a given point in time, depending on the current configuration and status at the facility.

Starting in the mid 1990s, SINTEF carried out the ‘Risk Indicator Project’ together with Statoil and the Norwegian Petroleum Directorate with the aim of developing a set of indicators to be used to monitor possible changes in the risk level (Øien et al., 1996, 1997, 1998; Øien and Sklet, 1999a,b, 2000). This project utilized a risk based approach, using the QRA as a basis (Øien, 2001b), and resembles the PSA based approach used within the nuclear power industry. The results from this research are described in the accompanying follow-up paper (Øien et al., 2010), which constitutes Part 2 of the research.

3.8. Accident sequence precursors

An Accident Sequence Precursor (ASP) may be considered as a ‘near incident’; an event consisting of (1) an initiating event, (2) an initiating event combined with system failure and unavailability, or (3) the occurrence of system failure and/or unavailability for a given time period (Johnsen and Rasmuson, 1996). Besides doing a common qualitative investigation of ‘near incidents’, a quantitative

---

[7] Note that ‘levels’ are used in two different ways, both for the level of the PSA and the level of indicators. Thus, level 1 PSA is different from level 1 indicators.

[8] Not included in Fig. 6, due to the assumption that Core Damage Frequency for Initiating Event and Safety System Unavailability are sufficient to identify conditions regarding the plant’s ability to deal with events that need further analyses or improvements.

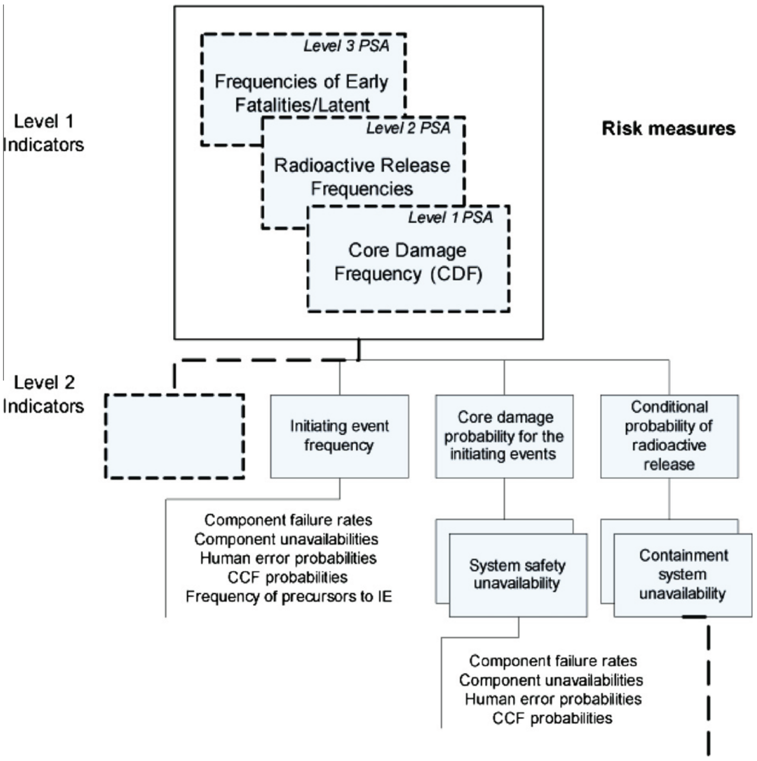


Fig. 6. Framework for establishing PSA based risk indicators (IAEA, 1999b).

analysis of the event based on the risk analysis (PSA/QRA) is carried out, in which the risk is estimated, given the event. For Nuclear Power Plants, the most common measure of risk is CDF – Core Damage Frequency (level 1 PSA). A CDF is estimated given the occurrence of a ‘precursor’, i.e., a Conditional Core Damage Probability (CCDP). Therefore, the estimated CCDP is a measure of how serious the ‘near incident’ was. US NRC has utilized this within its ‘Accident Sequence Precursor Program’ (ASP Program) (Minarick and Kulielka, 1982), and has supported research within the same area. The objectives of the US NRC’s ASP program (from the view of the regulators) are foremost to identify and rank the risk severity of operational events, thereafter to:

- Determine general implications of ASP events and describe risk knowledge (e.g., through trend analyses).
- Give supplementing information about plant specific performance.
- Give opportunity for control of the risk analyses.
- Give an empirical indication of industrial risk and corresponding trends.

For a single plant, the last mentioned objective means that the CCDP constitutes an empirical indicator of level 2 type, assuming that the event is an initiating event. In general terms, this is the only coupling between research on accident sequence precursors and research on risk indicators[9] (actual use of ASP, exceeding the original ‘ranking of events’ target, is to estimate annual CDF based on ASP information (Johnsen and Rasmuson, 1996), or use ASP to estimate the frequency of rare events (Bier and Yi, 1995).

There have been attempts to couple ASP and risk indicators in a project carried out by VEIKI[10] for Paks nuclear power plant in Hungary. In this project, the objectives were to estimate risk due to events (ASP) at lower levels (even lower than CDF) and to give early warnings of negative trends (Karsa, 1998). To enable early warning, risk indicators at lower levels than CDF had to be defined, so that estimations of conditional risk targets at lower levels than CCDP could be carried out. Karsa (1998) concludes that it is necessary to define and analyse more risk-based indicators.

3.9. The resilience engineering perspective on indicators

Resilience refers to the capability of recognizing, adapting to, and coping with the unexpected (Woods, 2006). Thus, resilience based indicators may be an aid in situations of incomplete knowledge (about what may go wrong). Resilience engineering, by acknowledging the fact that we do not have complete knowledge, insists that the previous approaches need to be complemented for the possibility of the unexpected.

EPRI (2001) focuses on leading indicators and refers to Reason’s Model of Organizational Accidents (1997). Wreathall’s model shows that unsafe actions are the results of local workplace factors, which in turn are influenced by the organizational factors. Proactive Assessment of Organizational and Workplace Factors (PAOWF) provides a tool to monitor the local workplace factors, whereas

---

[9] This is based on how NRC defines an ASP, but if a ‘precursor’ is allowed to be a precursor to the initiating event (not the event itself), then such a precursor may be used as an indicator if the number of initiating events occur too seldom to be used as an indicator (this is the meaning of ‘frequency of precursors of IE’ in Fig. 6).

[10] VEIKI Institute for Electric Power Research Co.

Table 3
Example of leading indicators (derived from EPRI, 2001).

Theme	Issues	Potential indicators
Management commitment	Personal commitment	Number of separate human performance (HP) meetings
Awareness	Knowledge seeking	Percentage of HP issues getting root cause analysis
Preparedness	Reactive	Ratio of unplanned to planned work orders
Flexibility	–	Average time to close a SmartForm
Just culture	Fault tolerance	Number and duration of temporary modifications
Learning culture	Responses to HP problems	Ratio of corrective actions involving discipline/counselling/retrain or change procedure/systematic changes
Opacity	Knowledge seeking	Number of quality management observations

Leading Indicators of Organizational Health (LIOH) provides a tool for monitoring the organizational factors.

The organizational factors – or themes – in LIOH are:

- Management commitment.
- Awareness.
- Preparedness.
- Flexibility.
- Just culture.
- Learning culture.
- Opacity.

Examples of leading indicators, derived from the use of LIOH, are shown in Table 3.

These leading indicators measure themes that are characteristics of a resilient organization, and may thus be seen as indicators within the framework of resilience engineering (Wreathall, 2006). The emerging practices of resilience engineering aim to provide organizations the means to track sources of organizational (system) resilience, to use the indicators to make better decisions in the face of production/safety trade-offs, and to create foresight so that organizations can anticipate opportunities and changing risks before failures and harm occur.

In the Building Safety project (SINTEF, 2010) we have developed a new method for the establishment of early warning indicators. It is based to some extent on the LIOH method. The new method has been adapted mainly in two ways. First, the factors seen as important to the management of safety (the seven 'themes') have been replaced by attributes of a resilient organization (eight 'contributing success factors'), (Størseth et al., 2009). Secondly, for each of the contributing success factors a set of general issues has been suggested and accompanied with proposals for early warning indicators, i.e., a list of general issues with proposed early warning indicators has been developed and included as part of the method, which is a deviation from the original LIOH method. In addition, there will be an option for including new general issues and early warning indicators during the workshop sessions in which the method is applied and indicators established/selected. This new method, called Resilience based Early Warning Indicators (REWI), is described in (Øien et al., 2010).

Woods (2006) argues that it may be possible to measure potential for resilience, rather than resilience itself. This is due to resilience being an agglomerated, rather than a single, quality. Factors identified that contribute to resilience, include buffering capacity, flexibility, margin, tolerance, and cross-scale interactions. Mendoça (2008) identifies and measures these factors affecting resilience by triangulation of observation using quantitative and qualitative data.

Most of the approaches identified for development of indicators within the field of resilience engineering have focused on organisational factors and human performance in a somewhat fragmented manner. One challenge is to apply a systemic approach, taking into account the interactions between human, organisations and technology with focus on integrating the socio-technical system, as a whole. Functional Resonance Accident Model (FRAM) (Hollnagel, 2004) explains failures and successes as a result of adaptations to cope with complexity. The model challenges the understanding of the functions of the systems and how performance variability is necessary. FRAM includes the identification of indicators that monitor performance variability of the system; however, further developments of FRAM are necessary.

4. Recent discussions about safety indicators

Hopkins (2009a) discusses two dimensions of safety indicators: personal safety versus process safety, and leading versus lagging indicators. Personal safety indicators are not, as previously mentioned, the topic of this paper. Personal safety is, for example, about avoiding cuttings, trips and falls among employees; hence it does not represent management of process hazards.

When the research on developing indicators or metrics for major hazards started, the focus was on direct or 'lagging' indicators, i.e., after-the-event type of indicators. This approach counts the number of accidents or incidents or near misses, however, these indicators are not very useful as pre-warnings or early warnings. For early warnings, one needs to look further back in the causal chain, at the underlying causes and the condition of the factors that leads to accidents. This has previously been termed indirect or proactive indicators, nowadays often referred to as 'leading' indicators,[11] which, according to Baker et al. (2007), provide performance feedback before an accident or incident occurs.

According to HSE and CIA (2006), performance measurements may be divided into reactive monitoring and active monitoring. The former means identifying and reporting on incidents, and learning from mistakes, whereas the latter provides feedback on performance before an accident or incident occurs.

Lagging indicators are related to reactive monitoring and show when a desired safety outcome has failed, or when it has not been achieved. Examples of lagging indicators are the number of unexpected loss-of-containment incidents and failures of safety critical instrumentation/alarms.

The leading indicators are a form of active monitoring used as inputs that are essential to achieve the desired safety outcome. These indicators require systematic checks if activities are carried out as intended. The information from the leading and lagging indicators should be used to follow-up findings as means to correct errors in the safety management system, and to review performance against all indicators to evaluate effectiveness of the safety management system on a regular basis. This means that performance indicators are not a replacement for an audit program, but is a complimentary activity contributing to more frequent and supplementary information on system performance (HSE and CIA, 2006).

---

[11] This does not mean that the terms 'direct and indirect' are identical and exchangeable with 'lagging and leading'. The term 'direct' has been closely related to the physical/technical part of the socio-technical system.

Leading and lagging indicators may be illustrated using Reason's accident model (1997), in which accidents are explained as a series of failings (holes) in the layers of defenses, barriers and safeguards (the 'Swiss Cheese Model'). Leading indicators identify the holes in the risk control system during routine checks, whereas the lagging indicators reveal the holes in the barriers as a result of an incident. The incident does not have to cause injuries or damages, but may be a near miss or a precursor event (HSE and CIA, 2006).

HSE and CIA (2006) emphasize the importance of utilizing both leading and lagging indicators and use the term 'dual assurance' approach. If performance is poor against a group of leading indicators, but the associated lagging indicator is satisfactory, it is likely that the leading indicators selected are too far removed from the critical control measure that delivers or maintains the desired outcome. If a group of leading indicators are on target and closely linked to the risk control system, but the associated lagging indicator shows poor performance, it is likely that the risk control system is ineffective in delivering the desired outcome (HSE and CIA, 2006). In this way the use of leading and lagging indicators helps to assure that the selected set of indicators are appropriate. This is quite opposite to Vinnem (2010) who claims that 'it is commonly accepted that 'leading' indicators are clearly to be preferred over 'lagging' indicators'.

Hopkins (2009a) concludes that Baker et al. (2007) and HSE and CIA (2006) are not using the terms leading and lagging in a consistent way, which has initiated a debate in the safety society about the definitions of leading and lagging indicators (Hale, 2009a).

Safety Science invited selected safety researchers and practitioners to respond to Hopkins paper (Hale, 2009a). A total of 20 contributions were received, including Hopkins reply to several responses (Hopkins, 2009b; Hale, 2009a). The debate highlights the diversity of understandings and a genuine confusion in this area (Hopkins, 2009b; Allford, 2009). Key points from the different contributions are presented and discussed in the following.

The main purposes of performance indicators are (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, and (3) to motivate those in a position to take the necessary action to actually do it. Leading indicators address the need to predict and act before a disastrous event (Hale, 2009b). Many people think that a distinction between lag and lead indicators is important. However, Hopkins argues that in the area of process safety, this distinction has no clear meaning and that it is of relative little value. He bases his argument on the fact that the bowtie model does not provide good basis for the distinction between lead and lag.

Regardless of the lead and lag distinction, the main contribution of the HSE document (HSE and CIA, 2006) relates to the measurement of the control systems' effectiveness, in order to identify measures of how well the process safety controls are functioning (Hopkins, 2009b; Erikson, 2009).

A common agreement is on the need for meaningful indicators for the state of the safety management system (Hopkins, 2009b). Interest is also on precursor events that are early warnings, and companies need to seek and use these warnings as a trigger for investigation and action (Erikson, 2009; Wreathall, 2009; Woods, 2009). Usually, such warnings are not acted upon, because the causal link can often only be established after the fact. Sound knowledge of cause-and-effect relations is therefore needed.

Accident and incident analyses teach us that it is usually not single precursor events, but particular patterns of events that lead to negative safety outcomes (Grote, 2009). Gaining more knowledge about these patterns and building competence within companies and regulatory bodies for recognizing relevant patterns are crucial. Ale (2009) supports this argument by pointing out that many accidents have happened, not because process variables were extremely out of range. Often they were inside the designed distribution, although the probability of the actual combination of extreme values was rare.

The indicators need to be identified as relevant for the group for which the performance is measured, and they need to be possible to influence by their management. The indicators should show responses within a convenient timeframe. In this way, the indicators will motivate to take necessary actions (Hale, 2009a). An alternative is selecting leading indicators based on an underlying safety model. The model will help to identify emergent patterns leading to expected and unexpected outcomes (Wreathall, 2009; Woods, 2009).

The safety community is multidisciplinary and the same meanings are not necessarily shared. Therefore, it is required to carefully define the concept of indicators every time we use the term (Hopkins, 2009b). The starting point could be to establish the purpose of indicators, describing the functions that they may have (Grote, 2009; Harms-Ringdahl, 2009). Kjellén (2009) proposes a combination of performance data, risk assessment and expert judgment. The challenge is to develop indicators with the ability to predict future safety performance. High Reliability Organisations (HROs) are interested in single events that provide weak signals of something being wrong. One suggestion is to treat the number of weak signals as an indicator; Hopkins (2009a) writes 'the more the better'.

Although the recent discussions, referred to above, include some interesting viewpoints, the main impression is that the confusion regarding leading versus lagging indicators (which initiated the debate) is even greater than before this 'dispute' was launched. The discussion show signs of lack of knowledge about the previous research (e.g., on direct versus indirect indicators) carried out in the 1980s and 1990s, which is closely linked to the debate on leading versus lagging indicators. With some few exceptions, it seems to be very little overlap between those safety researchers and practitioners developing this field in the 1980s and 1990s, and those invited to/taking part in the debate, which may explain the apparent lack of awareness of the historical development within the field of safety indicators (referred to in Section 3).

5. Discussions and conclusions

Leading indicators for major accidents as a research area is not new (even though the term 'leading' has been introduced recently), but it has been applied only to a limited extent in the industry, as revealed by, e.g., the Longford accident investigation (Hopkins, 2000) and the Texas City accident investigation (Baker et al., 2007; CSB, 2007). These accidents demonstrated the lack of knowledge in the industry on the usefulness of process safety indicators in general, and leading process safety indicators in particular.

Future developments in the field of safety indicators in general and on early warning indicators in particular, should rest on a sound theoretical foundation, including basic concepts, main perspectives and past developments, as well as an overview of the present status and ongoing research. The first part has been the main emphasis of this paper, whereas the latter is covered in the follow-up paper (Øien et al., 2010).

An indicator is a measurable representation of an aspect of reality. This aspect could be, e.g., safety or risk. Unless it is explicitly stated what we mean by a risk indicator, we should distinguish between safety indicators and risk indicators. They represent two different perspectives; one based on assumed relations or the use of correlation, and the other on causal connection through a risk model. In past developments, the latter is referred to as a probabilistic indicator, a PSA based risk indicator or simply a risk indicator.

One advantage of the risk indicators is that the effect on the aspect we want to measure – here risk – is known. We can then easily distinguish important (risk) influencing factors and the corresponding indicators, from the less important ones.

The major hazard industries can benefit substantially from increased utilization of existing methods for the development of risk or safety indicators. However, all development of indicators is context specific. There is no such thing as a universal model or method for the development of indicators and perhaps the use of several different methods will provide the best result – the most appropriate set of indicators.

Although a lot can be gained by existing methods, there are challenges yet to be solved. For instance, methods for the development of proactive indicators, such as organizational indicators, still lack consensus, and the problem with lack of data (the controller dilemma) may call for methods that focus on positive factors and corresponding positive indicators. This is an area in which resilient engineering research may contribute. Also, resilience based indicators may be an aid in situations of incomplete knowledge (about what may go wrong), since we focus on being prepared for the unexpected.

Based on the review of safety indicator developments in the past and resent discussions about safety indicators, we can conclude that: (i) extensive work on indicators have been carried out in the past, and this could and should have been better utilized be industry, (ii) there exist a confusion in the definitions of indicators (e.g., as discussed by Hopkins (2009a) regarding leading and lagging indicators), which is explained by the multidisciplinary nature of the safety community and perhaps is also due to lack of knowledge or disagreement about past research, and (iii) although the distinction between leading and lagging indicators may be of theoretical interest it can be counterproductive in practice.

This disagreement should not be allowed to impede the development of early warning indicators to prevent major accidents. What we need is not a discussion on what is lead and what is lag, but to develop and implement useful indicators that can provide (early) warnings about potential major accidents, so that we can prevent disasters such as those in Texas City and Longford, as well as major accidents related to petroleum production in the northern regions, in the future.

Acknowledgements

The work in this paper has been carried out as part of the research project 'Building Safety in Petroleum Exploration and Production in the Northern Regions'. Financial support from the Research Council of Norway, Eni Norge AS and TrygVesta is gratefully acknowledged. We also appreciate the valuable comments we have received from Erik Hollnagel, John Wreathall, and an anonymous reviewer during the preparation process of this paper.

References

Ale, B., 2009. More thinking about process safety indicators. Safety Science 47, 470–471.

Allford, L., 2009. Process safety indicators: response to Andrew Hopkins. Safety Science 47, 466.

Baker, III J.A., Leveson, N., Bowman, F.L., Priest, S., Erwin, G., Rosenthal, I., Gorton, S., Tebo, P.V., Hendershot, D., Wiegmann, D.A., Wilson, L.D., 2007. The Report of the BP US Refineries Independent Safety Review Panel.

Baranowsky, P.W., Mays, S.E., Wolf, T.R., 1999. Development of risk-based performance indicators. In: Proceedings of the Probabilistic Safety Assessment International Topical Meeting, American Nuclear Society, Washington, DC, USA, pp. 414–421.

Belles, R.J, Cletcher, J.W., Copinger, D.A., Dolan, B.W., Minarick, J.W., Muhlheim, M.D., 1998. Precursor to Potential Severe Core Damage Accidents. NUREG/CR-4674, ORNL/NOAC-232, vols. 1–26. US Nuclear Regulatory Commission, Washington, DC, USA.

Bier, V.M., Yi, W., 1995. The performance of precursor-based estimators for rare event frequencies. Reliability Engineering and System Safety 50, 241–251.

Blackman, H.S., Gertman, D.I., Hallbert, B.P., Schurman, D.L., Thompson, C., 1998. Integrating safety culture. In: Probabilistic Safety Assessment and Management. Proceedings from PSAM 4. Springer, New York City, USA.

Boccio, J.L., Vesely, W.E., Azarm, M.A., Carbonaro, J.F., Usher, J.L., Oden, N., 1989. Validation of Risk-Based Performance Indicators: Safety System Function Trends. NUREG/CR-5323, BNL-NUREG-52186. US Nuclear Regulatory Commission, Washington, DC, USA.

CSB, 2007. US Chemical Safety and Hazard Investigation Board, 2005. Investigation Report, Refinery Explosion and Fire. BP. Texas City, Texas. Report No. 2005-04-I-TX 2007.

Davoudian, K., Wu, J.-S., Apostolakis, G., 1994a. Incorporating organizational factors into risk assessment through the analysis of work processes. Reliability Engineering and System Safety 45, 85–105.

Davoudian, K., Wu, J.-S., Apostolakis, G., 1994b. The work process analysis model (WPAM). Reliability Engineering and System Safety 45, 107–125.

Duijm, N.J., Fievez, C., Gerbec, M., Hauptmanns, U., Konstandinidou, M., 2008. Management of health, safety and the environment in process industry. Safety Science 46 (6), 908–920.

Embrey, D.E, Humphreys, P.C, Rosa, E.A., Kirwan, B., Rea, K., 1984 SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment. NUREG/CR-3518. US Nuclear Regulatory Commission, Washington, DC, USA.

EPRI, 2001. The US Department of Energy. Final report on Leading Indicators of Human Performance. Palo Alto, CA and Washington, DC, 1003033.

Erikson, S.G., 2009. Performance indicators. Safety Science 47, 468.

Gertman, D.I., Hallbert, B.P., Blackman, H., Schurman, D., Thompson, C., 1998. Management and organizational factors research: the socio-organizational contribution to risk assessment and the technical evaluation of systems (SOCRATES). In: Mosleh, A., Bari, R.A. (Eds.), Probabilistic Safety Assessment and Management, Proceedings from PSAM 4. Springer, New York City, USA.

Grabowski, M., Ayyalasomayajula, P., Merrick, J., Harrald, J.R., Roberts, K., 2007. Leading indicators of safety in virtual organizations. Safety Science 45, 1013–1043.

Gray, P.C.R., Wiedemann, P.M., 1999. Risk management and sustainable development: mutual lessons from approaches to the use of indicators. Journal of Risk Research 2, 201–218.

Grenness, T., 1997. Introduction to Scientific Theory and Methodology. Tano Aschehoug, Oslo (in Norwegian).

Grote, G., 2009. Response to Andrew Hopkins. Safety Science 47, 478.

Haber, S.B., O'Brien, J.N., Metlay, D.S., Crouch, D.A., 1991. Influence of organizational factors on performance reliability. In: Overview and Detailed Methodological Development. NUREG/CR-5538, BNL-NUREG-52301, vol. 1. US Nuclear Regulatory Commission, Washington, DC, USA.

Hale, A., 2009a. Editorial special issue on process safety indicators. Safety Science 47, 459.

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

Harms-Ringdahl, L., 2009. Dimensions in safety indicators. Safety Science 47, 481–482.

Hellevik, O., 1999. Research Methodology in Sociology and Political Science. Universitetsforlaget, Oslo, Norway (In Norwegian).

Herrera, I., Woltjer, R., 2008. Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis. In: Safety, Reliability and Risk Analysis: Theory, Methods and Applications. Taylor & Francis.

Hollnagel, E., 2004. Barrier and Accident Prevention. Ashgate, Hampshire, England.

Holmberg, J., Laakso, K., Lehtinen, E., Johanson, G., 1994. Safety evaluation by living probabilistic safety assessment and safety indicators. In: TemaNord 1994:614, The Nordic Council of Ministers, Copenhagen, Denmark.

Hopkins, A., 2000. Lessons from Longford. The Esso Gas Plant Explosion. CCH Australia Limited, Sydney.

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

Hopkins, A., 2009b. Reply to comments. Safety Science 47, 508–510.

HSE and CIA (Chemical Industries Association), 2006. Developing process safety indicators. A step-by-step guide for chemical and major hazard industries. Health and Safety Executive.

IAEA (International Atomic Energy Agency), 1988. Basic Safety Principles for Nuclear Power Plants. Safety Series 75-INSAG-3. Vienna, Austria.

IAEA, 1991. Safety Culture. IAEA Safety Series No. 75-INSAG-4. International Nuclear Safety Advisory Group, Vienna, Austria.

IAEA, 1999a. Management of Operational Safety in Nuclear Power Plant. INSAG-13. International Nuclear Safety Advisory Group, Vienna, Austria.

IAEA, 1999b. Indicators to Monitor NPP Operational Safety Performance. IAEA-J4-CT-2883 Draft 15 January. Vienna, Austria.

Johnsen, J.W., Rasmuson, D.M., 1996. The US NRC's accident sequence precursor program: an overview and development of a Bayesian approach to estimate core damage frequency using precursor information. Reliability Engineering and System Safety 53 (2), 205–216.

Karsa, Z., 1998. Analysis of events to support risk evaluation of nuclear power plants. In: Hansen, S. (Ed.), Safety and Reliability, ESREL. Balkema, Rotterdam.

Kjellén, U., 2000. Prevention of Accidents Through Experience Feedback. Taylor & Francis, London and New York.

Kjellén, U., 2009. The safety measurement problem revisited. Safety Science 47, 486–489.

Körvers, P.M.W., Sonnemans, P.J.M., 2008. Accidents: a discrepancy between indicators and facts! Safety Science 46 (7), 1067–1077.
Leveson, N., 2004. A new accident model for engineering safer systems. Safety Science 42, 237–270.
Long, S.M., O'Reilly, P.D., Rodrick, E.G., Sattison, M.B., 1998. Current status of the SAPHIRE models for ASP evaluations. In: Mosleh, A., Bari, R.A. (Eds.), Probabilistic Safety Assessment and Management, Proceedings From PSAM 4. Springer, New York, p. 195.
Marcus, A.A., Nichols, M.L., Bromiley, P., Olson, J., Osborn, R.N., Scott, W., Pelto, P., Thurber, J., 1990. Organization and Safety in Nuclear Power Plants. NUREG/CR-5437. US Nuclear Regulatory Commission, Washington, DC, USA.
Mearns, K., 2009. From reactive to proactive – can LPIs deliver? Safety Science 47, 491–492.
Mendoça, D., 2008. Measures of resilient performance. In: Hollnagel, E., Nemeth, C.P., Dekker, S. (Eds.), Resilient Engineering Perspectives, Remaining Sensitive to the Possibility of Failure, vol. 1. Ashgate, Aldershot, USA.
Minarick, J.W., Kukielka, C.A., 1982. Precursors to Potential Severe Core Damage Accidents: 1969–1979: A Status Report. NUREG/CR-2497. US Nuclear Regulatory Commission, Washington, DC, USA.
Mintzberg, H., 1979. The Structure of Organizations. Englewood Cliffs, Prentice-Hall, Inc., NJ, USA.
Mintzberg, H., 1983. Power In and Around Organizations. Englewood Cliffs, Prentice-Hall, Inc., NJ, USA.
Mintzberg, H., 1988. Mintzberg on Management: Inside our Strange World of Organizations. The Free Press, New York, USA.
Moreño, M., Correa, M.A., Sola, R., 1998. Strategy for the development of operational safety indicators in the chemical industry. In: 9th International Symposium on Loss Prevention and Safety Promotion in the Process Industries, Barcelona, Spain, pp. 205–215.
OECD (Organisation for Economic Co-operation and Development), 2003. Guidance on safety performance indicators. Guidance for industry, public authorities and communities for developing SPI programmes related to chemical accident prevention, preparedness and response. OECD Environment, Health and Safety Publications. Series on Chemical Accidents 11.
Øien, K., 2001a. A framework for the establishment of organizational risk indicators. Reliability Engineering and System Safety 74, 147–167.
Øien, K., 2001b. Risk Control of Offshore Installations. A Framework for the Establishment of Risk Indicators. Department of Production and Quality Engineering, PhD thesis. Norwegian University of Science and Technology (NTNU), Trondheim, Norway.
Øien, K., 2001c. Risk indicators as a tool for risk control. Reliability Engineering and System Safety 74, 129–145.
Øien, K., 2008. Development of early warning indicators based on accident investigation. In: International Probabilistic Safety Assessment and Management Conference PSAM9. Hong Kong, China, 18–23 May 2008.
Øien, K., Sklet, S., 1999a. Risk control during operation of offshore petroleum installations. In: Schuëller, G.I., Kafka, P. (Eds.), Proceedings of the European Conference on Safety and Reliability (ESREL' 99). Balkema, Munich, Germany, pp. 1297–1302.
Øien, K., Sklet, S., 1999b. Application of risk analyses in the operating phase, establishment of safety indicators and modelling of organizational factors' effects on the risk level—a "state-of-the-art" description (In Norwegian: Bruk av risikoanalyser i driftsfasen, etablering av sikkerhetsindikatorer og modellering av organisatoriske faktorers effekt på risikonivået- en "state-of-the-art" beskrivelse). STF38 A99416. SINTEF Technology and Society, Safety Research, Trondheim, Norway.
Øien, K., Sklet, S., 2000. A structure for the evaluation and development of organizational factor frameworks. In: Kondo, S., Furuta, K. (Eds.), Proceedings of the International Conference on Probabilistic Safety Assessment and Management (PSAM 5). Universal Academy Press, Inc., Osaka, Japan, pp. 1711–1717.
Øien, K., Sklet, S., Nielsen, L., 1996. Use of risk analysis in the regulation of the Norwegian Petroleum Industry. In: Proceedings of the International Topical Meeting on Probabilistic Safety Assessment (PSA 1996). American Nuclear Society, Park City, IL, USA, pp. 756–762.
Øien, K., Sklet, S., Nielsen, L., 1997. Risk level indicators for surveillance of changes in risk level. In: Proceedings of the European Conference on Safety and Reliability (ESREL' 97). Pergamon, Lisbon, Portugal, pp. 1809–1816.
Øien, K., Sklet, S., Nielsen, L., 1998. Development of risk level indicators for a petroleum production platform. In: Proceedings of the 9th International Symposium of Loss Prevention and Safety Promotion in the Process Industries. Springer, Spain, pp. 382–393.
Øien, K., Massaiu, S., Tinmannsvik, R.K., Størseth, F, 2010b. Development of early warning indicators based on resilience engineering. Submitted to PSAM10, In: International Probabilistic Safety Assessment and Management Conference, Seattle, USA, 7–11 June 2010.
Øien, K., Utne, I.B., Tinmannsvik, R.K., Massaiu, S., 2010. Building Safety Indicators. Part 2—Application, practices and results. Safety Science 49 (2), 162–171.
Olson, J., McLaughlin, S.D., Osborn, R.N., Jackson, D.H., 1984. An Initial Empirical Analysis of Nuclear Power Plant Organization and its Effect on Safety Performance. NUREG/CR-3737, PNL-5102, BHARC-400/84/007. US Nuclear Regulatory Commission, Washington, DC, USA.
Olson, J., Osborn, R.N., Jackson, D.H., Shikiar, R., 1985. Objective Indicators of Organizational Performance at Nuclear Power Plants. NUREG/CR-4378, PNL-5576, BHARC-400/85/013. US Nuclear Regulatory Commission, Washington, DC, USA.
Olson, J., Chockie, A.D., Geisendorfer, C.L., Vallario, R.W., Mullen, M.F., 1988. Development of Programmatic Performance Indicators. NUREG/CR-5241, PNL-6680, BHARC-700/88/022. US Nuclear Regulatory Commission, Washington, DC, USA.
Osborn, R.N., Olson, J., Sommers, P.E., McLaughlin, S.D., Jackson, J.S., Scott, W.G., Connor, P.E., 1983a. Organizational Analysis and Safety for Utilities with Nuclear Power Plants Volume 1. An Organizational Overview. NUREG/CR-3215, PNL-4655, BHARC-400/83/011. US Nuclear Regulatory Commission, Washington, DC, USA.
Osborn, R.N., Olson, J., Sommers, P.E., McLaughlin, S.D., Jackson, M.S., Nadel, M.V., Scott, W.G., Connor, P.E., Kerwin, N., Kennedy, J.J.K., 1983b. Organizational Analysis and Safety for Utilities with Nuclear Power Plants Volume 2 – Perspectives for Organizational Assessment. NUREG/CR-3215, PNL-4655, BHARC-400/83/012. US Nuclear Regulatory Commission, Washington, DC, USA.
Osmundsen, P., Aven, T., Vinnem, J.E., 2008. Safety, economic incentives and insurance in the Norwegian petroleum industry. Reliability Engineering and System Safety 93, 137–143.
Reason, J., 1997. Managing the Risks of Organizational Accidents. Ashgate Publishing Limited, England.
Saqib, N., Siddiqi, M.T., 2008. Aggregation of safety performance indicators to higher-level indicators. Reliability Engineering and System Safety 93, 307–315.
SINTEF, 2010. Building Safety in Petroleum Exploration and Production in the Northern Regions. <http://www.sintef.no/buildingsafety>.
Størseth, F., Tinmannsvik, R.K., Øien, K., 2009. Building safety by resilient organization – a case specific approach. Paper at The European Safety and Reliability Association Annual Conference (ESREL), 7–18 September, 2009, Prague, Czech Republic.
Tarrants, W.E., 1980. The Measurement of Safety Performance. Garland STPM Press, New York, USA.
Vinnem, J.E., 2010. Risk indicators for major hazards on offshore installations. Safety Science 48, 770–787.
Wilpert, B., 2000. Organizational factors in nuclear safety. In: Kondo, S., Furuta, K. (Eds.), PSAM, vol. 5. Universal Academy Press, Inc., Tokyo, Japan, pp. 1251–1265.
Woods, D.D., 2006. Essential characteristics of resilience. In: Resilience Engineering: Concepts and Precepts. Ashgate, Aldershot.
Woods, D.D., 2009. Escaping failures of foresight. Safety Science 47, 498–501.
Wreathall, J., 2006. Properties of resilient organizations: an initial view. In: Resilience Engineering: Concepts and Precepts. Ashgate, Aldershot.
Wreathall, J., 2009. Leading? Lagging? Whatever! Safety Science 47, 493–494.
Wreathall, J., Fragola, J., Appignani, P., Burlile, G., Shen, Y., 1990. The Development and Evaluation of Programmatic Performance Indicators Associated with Maintenance at Nuclear Power Plants: Main Report. NUREG/CR-5436, SAIC-90/1130, vol. 1. US Nuclear Regulatory Commission, Washington, DC, USA.
Wreathall, J., Schurman, D.L., Modarres, M., Anderson, N., Roush, M.L., Mosleh, A., 1992. US Nuclear Regulatory Commission: A Framework and Method for the Amalgamation of Performance Indicators at Nuclear Power Plants NUREG/CR-5610, vols. 1 and 2. US Nuclear Regulatory Commission, Washington, DC, USA.
Wu, J.S., Apostolakis, G., Okrent, D., 1991. On the inclusion of organization and management factors into probabilistic safety assessments of nuclear power plants. In: Proceedings of the International Conference on Probabilistic Safety Assessment and Management (PSAM). Elsevier Science Publishing Co., NY, USA, pp. 619–624.
Youngblood, R.W., Hunt, R.N.M., Schmidt, E.R., Bolin, J., Dombek, F., Prochnow D., 1999. Elements of an Approach to Performance-Based Regulatory Oversight. NUREG/CR-5392, SCIE-NRC-373-98. US Nuclear Regulatory Commission, Washington, DC, USA.

Paper III

Leading indicators applied to maintenance in the framework of resilience engineering: A conceptual approach

Proceedings of the third Resilience Engineering Symposium,
held in Antibes-Juan Les Pins, France, on 28 - 30 October 2008

Paper presented at *The 3rd Resilience Engineering Symposium*, 28 – 30 October 2008, Antibes-Juan Les Pins, France.

Leading indicators applied to maintenance in the framework of resilience engineering: A conceptual approach

I.A. Herrera[1] and J. Hovden[2]
[1] Department of Production and Quality Engineering, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway
Ivonne.a.herrera@ntnu.no
[2] Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway
Jan.hovden@ntnu.no

Abstract. This paper explores the identification of leading indicators in aviation maintenance. Traditionally, improvements in aviation have been based on incident reporting and analyses and learning from failures. This traditional approach measures safety performance based on lagging indicators. However, there is a growing concern that this information does not provide the requisite information and insight to prevent future accidents. Resilience has been identified as the ability of the system to adjust prior, during and after a major mishap. This implies a need to recognize early signals to be able to anticipate and act properly.

The objective of the paper is to understand leading indicators and the "why" behind the leading indicator as basis to look forward in monitoring safety performance. With regard to the leading indicator definition applied to aviation and maintenance in the context of resilience, we focus on learning from success and failures and propose: Leading indicators are precursors based on a model of safety implying a significant possibility of a subsequent event that has an impact on safety and performance. Leading indicators can therefore provide information about changes in risk before traditional risk analyses are able to capture this change.

1 INTRODUCTION

Measuring safety performance is traditionally based on lagging indicators such as the accident rates. However, there is a growing concern that this information does not provide the required insight for the prevention of future accidents. This paper explores the identification of leading indicators that are of relevance for aviation maintenance. The focus is to understand leading indicators in the framework of resilience engineering (RE) and the "why" behind indicators as basis to look forward in the monitoring of safety. The RE framework emphasises aspects that are difficult to capture in traditional risk analyses. Hence, we discuss indicators that are hard to link to the risk analyses.

Research has shown that risk can exist in a system either because of components in the system or risky interaction between components (Perrow, 1984). There are several examples where maintenance activities were relevant to major accidents and extensive losses. These include the DC9 accident over Florida (1996), the Alaska Airline accident in 2000 and most recently, the Texas BP refinery explosion in USA (Baker, 2007). In many settings, the likelihood of an accident is low and preventing accidents requires continuous vigilance. The absence of an accident is not necessarily an indication that everything is going well (Van Steen, 1996). Organizations tend to attend to what is being measured rather to what is not (Hopkins, 2000). In the Longford accident the focus on Loss of Time Injuries (LTI) allowed the management to become complacent about the management of major hazards. The problem as Hopkins (2008) pointed out is that BP relied on lagging personnel indicators as a measure of performance of the process.

The paper starts with a description of the context of a maintenance organization. Then, a focused review of new approaches in the development of leading indicators is presented and discussed in the context of RE and maintenance. Finally, conclusions are presented regarding leading indicators, their characteristics and resilience approach to the identification of indicators in the maintenance management.

2 CHALLENGES AND OPPORTUNITIES IN MAINTENANCE

Maintenance is seen as a set of activities ensuring that the system continuously performs its intended function at its design level of reliability, performance and safety. Aviation maintenance comprises high technology and a group of interactive people concerned with safety and efficiency. The particular situation in aviation maintenance is that it depends on few aircraft manufacturers, prescriptive regulations influencing maintenance in order to follow specific requirements. The industry works together and provides feedback at all levels to improve continuous and safe operation. Recently, the International Civil Aviation Organization recommended establishing an effective Safety Management System (SMS). However, despite the benefit of an SMS, in general the aviation industry is still focused on reactive part of safety management. Learning is mainly based on failures recorded in the system and is a reactive approach to maintenance improvements. These measures are based on after-the-event information providing information of the past status of the system and they provide little information of the current state of performance for day to day decision making (Wreathall, 2006).

We need to analyse failures and successes and how they might be identified. In the case of successes, the literature provides little theoretical guidance. High Reliability Organizations (HRO) theory provides some examples of successes. The observations in the HRO theory imply two components, the high reliability of the organizations and individual excellence (Reason, 2003). The mindset in HRO expects surprises and has the flexibility to cope with them (Weick and Sutcliffe, 2007). Recently, RE has been defined by Hollnagel (2006) as “the ability of the system to adjust its functioning prior to or following changes and disturbances, so it can sustain operations in presence of continuous stress”. This ability is particularly interesting in a maintenance organization as RE implies that the organization is able to respond to threats and opportunities, to monitor risks and to anticipate potential disruptions in order to continue safe operations.

The challenges in aviation maintenance include some specific trends regarding the increase of subcontracting, maintenance activities spread over multiple locations, variation in task complexities, introduction of new technologies, decrease in available competence, reduction of available resources and time. The opportunities could include aircraft sensors providing continuous monitoring of aircraft systems, the trend towards openness, structured and detail planned maintenance at in the short and long periods, a dedicated group that follow current aircraft fleet status and operational advantages from early recognition and action with regard to maintenance problems enhancing market recognition for a proactive approach.

3 SEARCHING FOR LEADING INDICATORS

We looked into previous research on leading indicators to understand how they might be identified in the management of maintenance.

3.1 Different concepts

The objective of performance measures is to provide management required information for decision making. The function of a safety performance measure is to reveal the level of safety effectiveness in the organization with respect to the accident control desired (Kjellen, 2000). A safety indicator is an observable characteristic of an operational unit, presumed to bear a positive correlation with the safety of the system (Adapted from Holmberg, 1994).

Different definitions of leading indicators were found: *i)* Type of accident precursors, conditions, events or measures that precede an undesirable event and have some value in predicting the arrival of an event (Construction Owners Association of Alberta, 2004); *ii)* A form of active monitoring focused on few control systems (HSE, 2006); *iii)* "Activity" indicators that show if the organization is taking actions believed to lower risk (OECD, 2003); *iv)* Indicators that measure variables that are believed to be indicators or precursors of safety performance so that safety outcome is achieved (Baker, 2007). In contrast lagging indicators are measurements of a system that are taken after events, which measures outcomes and occurrences.

3.2 The ideal characteristics of indicators

Previous work on indicators reveals the key characteristics. None of the documents reviewed discussed the reasons behind the selection of the characteristics. As a general conclusion the following characteristics seem to be repeated across the literature: objective measure, easy to understand, indicate improvement or deterioration and collected from existing data (Kjellen 2000; Wreathall, 2006). Other characteristics that could be considered are diverse and complementary, interpreted by different groups in the same way, owned by the group whose performance is measured (Sefton, 1997). Specific, Measurable, Achievable, Relevant and Timed (SMART) characteristics are also mentioned for leading indicators including reasons behind indicators and benefit easy to understand, provide information that guide future actions, related to activities that are important for future performance, reinforce willing to intervention and provide clear indication of a means to improve performance (Blackmore, 1997). There is not a single measure that will meet all the characteristics. A combination of measures can provide a reasonable compromise (Kjellen, 2000).

3.4 What can we learn from other studies?

We will give a short introduction to different approaches focus on identification of leading indicators identified in selected studies. Wreathall (2006) identified themes in highly resilient organizations. Leading indicators sets should be based on: Management commitment, Just culture, Learning culture, Opacity, Awareness, Preparedness and Flexibility. Examples of indicators related to preparedness is "crisis training beyond minimum requirements" and to management commitment is "percentage of overtime". Graboski et al. (2007) identify leading indicators at sharp end (Empowerment, Individual responsibility, Anonymous reporting, Individual feedback, Problem identification, Vessels responsibility) and at organizational level (Organizational structure, Prioritizing for safety, Effective communication).

Another approach based on organizational resilience focuses on Commitment, Competence, and Cognizance ("three C's" in Reason and Hobbs, 2003). These three C's' are combined with "four P's": Principles, Policies, Procedures, and Practice. A matrix combines interaction of these concepts to indicate the organization's position in the safety space from increasing resistance to an accident to increasing vulnerability. The New Zealand Approach proposes organizational resilience as a function of vulnerability (likelihood and criticality of a failure), adaptive capacity (apply existing responses to problems and generate innovative responses to new problems), and situation awareness (understanding interdependencies and complexities within the system, knowing when environment are changing and systems response needs to change) (Seville et al., 2007).

Woods (2006) argues that it is possible to measure potential for resilience rather than resilience itself. Factors identified that contribute to resilience include buffering capacity, flexibility, margin, and tolerance and cross-scale interactions. Mendoça (2008) identifies and measures these factors affecting resilience by triangularization of observation using quantitative and qualitative data.

Most of the approaches identified for indicators within resilience framework have focus on organizational factors and human performance. The challenge is to integrate a systemic approach taking into account the interactions between human, organizations and technology. The Functional Resonance Accident Model (FRAM, Hollnagel, 2004) explains failures and successes as result of adaptations to cope with complexity. Two forms of monitoring have been identified the monitoring of performance variability at function level and the utilization of the "FRAModel" to understand systems status in relation to resilient characteristics at system level. Functions are defined by six aspects time, control, preconditions, resources, input and output. A change in one these aspects will have impact in the performance of the functions. At sharp end indicators could be related to time to execute a maintenance task, this aspect could be essential in case of maintenance of safety critical systems.

3.5 Lessons from the Ecology perspective

We would like to make an analogy to the ecology perspective on resilience. Figure 1 illustrates the potential ability for change and the resilience of the systems in respect to vulnerability to regular, irregular and unexpected events. Changes in the system are influenced by external and internal processes. Performance conditions are conditions that influence socio-technical system responses i.e. competence, team collaboration, quality of maintenance procedures, communication and system's complexity. Four possible states are illustrated "*K, conservation*"; "*Ω release*"; "*α, reorganization*" and "*r, growth or exploitation*". In aviation maintenance, we have a multiplicity of stable states dealing with scheduled and unscheduled activities.

In state "*Ω release*" tightly coupled and fragile, this state could be related to unscheduled maintenance and the possibility to disrupt operations. This state is followed by α reorganization including innovation and leading to a new stable state. This case represents the gathering of expertise, analyse and act including reorganization of maintenance. The next phase is r restructuring/exploitation and then K conservation characterized by slow changes. This could be related to maintenance solving a specific situation and returning back to a new stable state.

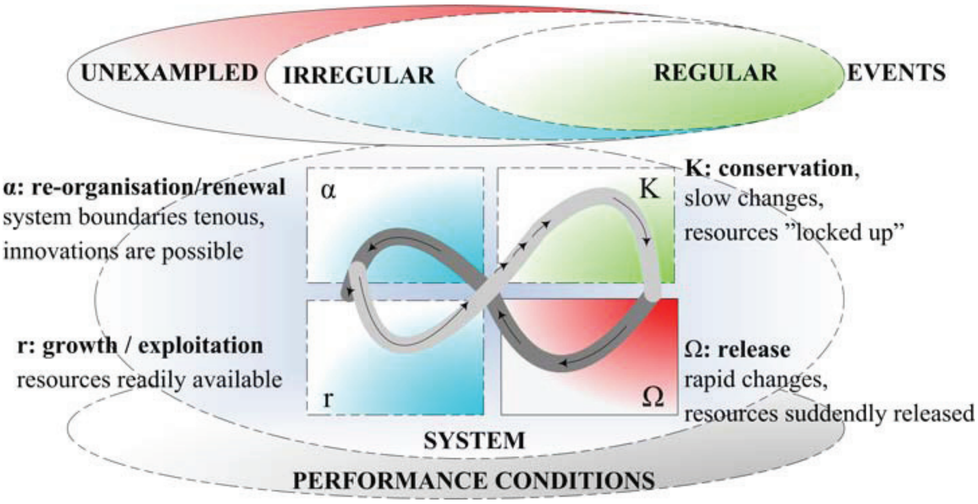


Fig. 1. Possible states in the adaptive cycle adapted from Gunderson (2002)

We argue that RE should provide alternative ways to cope with irregular and unexampled events in maintenance organizations. The model in the figure 1 introduces two dimensions potential and connectedness that influence transitions between the states. "Potential" is related to available resources; whereas "connectedness" is related to the influence of performance conditions and variability. Low connectedness is associated with variables that are dominated by outside factors and external variability. High connectedness represents those variables that are dominated by inside factors and influence the external variability. In relation to the management of maintenance, it would be convenient to identify which maintenance functions are more sensitive to external stresses and which maintenance functions are dominated by internal variability. An example of maintenance activities dominated by internal variability is reporting of trechnical failures. Some maintenance records contain more detailed information than other maintenance records and this fact influences the quality of the organizational response. Performance conditions in this example are having the time available to write the record, reporting system, experience of the reporter. An example of maintenance activities influenced by outside factors is the airline's execution of maintenance influenced by the variability of the manufacture's recommendations and regulators oversight. In relation, to the development of leading indicators, it is necessary to understand how internal and external factors affect the variability of the system.

At the same time the system is exposed to regular, irregular and unexampled events (Westrum, 2006). System response varies. Some systems are not able to cope with regular events. Regular events can be seen as same or known maintenance problems that are experienced. Therefore, unscheduled maintenance and disruption of operation are often experienced. This issue could be related to socio-technical ability of the system to learn. Leading indicators for these systems should therefore provide information towards improvement related to regular events. Other systems that manage responses to regular events will as a consequence be exposed to irregular and unexpected events. Leading indicators for these systems should focus on maintenance changes for unknown situations, sharing the risk picture of current maintenance situation, promoting alternative solutions for different risky situations.

3 DISCUSSION AND CONCLUSION

Most of safety performance indicators are based on after-the-event and provide limited information on current performance. Leading indicators are proposed as a complement to improve current practices for monitoring safety performance. The focus in this discussion is the understanding of the "why" behind leading indicators. The first challenge is to draw the line between the leading and lagging indicators. Hopkins (2007) pointed out that lagging indicators measure failures regardless if the consequences are catastrophic or not. Based on a literature review, there is no consistency between the definition of leading indicators and their application. Regarding the leading indicators definition applied to aviation and maintenance in the context of resilience, we focus on learning from both success and failures, proposing: Leading indicators are precursors based on a model of safety implying a significant possibility of a subsequent event that has an impact on safety and performance of maintenance activities.

Leading indicators are interpreted differently in the various safety models. In domino models leading indicators consist of single elements. Upon failing, these may subsequently lead to catastrophic failures. Leading indicators in the Swiss cheese model monitor performance of safety barriers. Indicators monitor according to defence-in-depth such as status of barriers.

The RE perspective proposes a systemic approach looking into the dynamics of safety. This means that leading indicators should attempt to provide a signal of the unintended system interactions. Successes and failures are the result from the normal performance variability. Instead of looking at failures the focus is changed to the understanding of the normal operation of the system. In this context, leading indicators allow the monitoring of changes which normally are not picked up and reflected in traditional risk assessments. These assessments have been based on decomposition and linearity.

Lessons from RE and resilience in ecology that should be taken into account for the development of leading indicators are: i) consideration of resources for the potential of their future use, ii) shifting balance between internal and external forces and connectedness in the system and iii) resilience is dynamic and must generate and sustain novelty and persistence to survive, iv) awareness of present state and indication regarding transitions between states. Questions for leading indicators: which factors might push the system to critical thresholds, which kind of indicators are worth monitoring if the system follows a particular trajectory? We consider that leading indicators will vary in different states, different events and will also differ depending on whether they are perceived at the sharp end or at management level. In addition to adaptation that it is mainly reactive, another aspect in resilience that is valid for socio-technical systems is recognition, this helps to identify unwanted interactions and provide guidance to creative responses.

Based on a literature review, indicators should contain the following characteristics: an "objective" measure that is easy to understand and will indicate improvement or deterioration that can be collected from existing data. While an "objective" characteristic is very relevant for lagging indicators that can be observed, we argue that leading indicators are characterized by inter-subjectivity: transparency and perceived in the same manner by different people. These are subject to inter-subjectivity with consensus between experts and decision makers. We propose that it is necessary to go further with the single indicators and provide interpretation to differentiate unintended interactions. Core tasks like monitoring and reflecting, carrying out analyses of change or "interpretation of information at hand to reveal what it is believed to be important" (Oedewald and Reiman, 2003) are necessary for interpretation and action.

Indicators in the framework of resilience do not replace other approaches to safety performance monitoring, but increase the understanding of normal performance. Once this new understanding is gained, the model is renewed and the system looks for new indications. Using FRAM to understand normal work in a specified maintenance activity, we propose leading indicators are context specific. Examples of leading indicators for the monitoring the accomplishment of maintenance in other sites are: the resources available, the capacity to identify circumstances beyond the experience, the possibility to reflect-on-action, openness, communication, the current technical state of the aircraft (critical systems monitored by Minimum Equipment List), maintenance oversight, and implementation of preventive maintenance.

Finally, while it is impossible to predict all possible scenarios, it is possible to improve system resilience by making sense of unintended interactions and use creativity to take advantage of system dynamics. This paper summarized the "why" of leading indicators as means to look forward and to lead actions for improvement. Further work on this paper will include development on specific leading indicators at sharp and blunt end.

REFERENCES

Baker, J (2007). The Report of the BP U.S. Refineries Independent Safety Review Panel

Blackmore, G. A. (1997). Leading performance indicators. Paper presented at the International Association of Drilling Contractors Seminar, Aberdeen.

Graboski, M., Premnath, A. Merrik, J., Harrald, J., Roberts, K. (2007). Leading indicators of safety in virtual organizations. Safety Science, 45, 1013-1043. Elsevier

Gunderson, L.H., Holling, C.S. (2002). Panarchy Understanding Transformations in Human and Natural systems. Island Press. USA.

Hollnagel, E. (2004). Barriers and accident prevention. Aldershot, UK: Ashgate

Hollnagel, E., Leveson, N., Woods, D. (2006). Resilience Engineering Concepts and Precepts, Aldershoot: Ashgate (*)

Holmberg J, Laakso K, Lehtinen E, Johanson G. (1994) Safety evaluation by living probabilistic safety assessment and safety indicators. Copenhagen, Denmark

Hopkins, A. (2000). Lessons from Longford. The ESSO Gas Plant Explosion. CHC Australia Limited

Hopkins, A. (2008) Thinking about process safety indicators. Safety Science; In press

Kjellén, U. (2000). Prevention of Accidents through Experience Feedback, Taylor & Francis, London UK

HSE (UK Health and Safety Executive) (2006) Developing process safety indicators

Mendoça D. (2008) Measures of Resilient Performance. In: Hollnagel E, Nemeth CP, Dekker S, editors. In Resilient Engineering Perspectives, Ashgate, Aldershot, USA.

Oedewald, P, Reiman, T. (2003). Core task modelling in cultural assessment: A case study in nuclear power plant maintenance. Cognition, Technology and Work 5, 283-293

Organization for Economic Cooperation and Development (OECD) (2003). Guidance on Safety Performance Indicators. Paris, France.

Perrow, C. (1984). Normal Accidents. Leaving with high risks technologies. Princenton University Press., USA

Reason, J. & Hobbs, A. (2003). Managing Maintenance Error, Ashgate, Aldershot

Sefton, A. (1997) Leading Indicators. Safety Measurement in the Future. Opening address at the International Association of Drilling Contractors Seminar, Aberdeen

Seville, E., McManus, S. Brundsdon, D, Vargo, J. (2007). Resilience Management: A Framework for Assessing and Imporving the Resilience of Organizations.

Van Steen, J (1996) Safety Performance Measurement. U.K

Weick, K., Sutcliffe, M. (2007) Managing the unexpected. Resilient Performance in the Age of Uncertainty. Second Edition. John Wiley & Sons, Inc. USA.

Westrum, R. (2006). A typology of resilience situations (*)

Woods, D. (2006). Essential Characteristics of Resilience (*)

Wreathall, J. (2006). Property of Resilient Organization: An Initial View (*)

Paper IV

Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis

Reliability Engineering and System Safety (2010)
vol 95, pp. 1269-1275

Invited paper for the special issue is on papers presented at the joint
European Safety and Reliability (ESREL) and Society for Risk Analysis Europe (SRA-Europe) conference held in Valencia Spain, September 22–25, 2008

Reliability Engineering and System Safety 95 (2010) 1269–1275

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis

I.A. Herrera [a,*], R. Woltjer [b,1]

[a] Department of Production and Quality Engineering, Norwegian University of Science and Technology (NTNU), S.P. Andersens vei 5, NO 7491 Trondheim, Norway
[b] Department of Computer and Information Science, Cognitive Systems Engineering Lab, Linköping University, Linköping, Sweden

ARTICLE INFO

Available online 1 July 2010

Keywords:
Performance variability
Systemic models
Non-linear models
Functional resonance
Accident analysis
Accident modelling

ABSTRACT

Accident models and analysis methods affect what accident investigators look for, which contributory factors are found, and which recommendations are issued. This paper contrasts the Sequentially Timed Events Plotting (STEP) method and the Functional Resonance Analysis Method (FRAM) for accident analysis and modelling. The main issue addressed in this paper is the comparison of the established multi-linear method STEP with the new systemic method FRAM and which new insights the latter provides for accident analysis in comparison to the former established multi-linear method. Since STEP and FRAM are based on a different understandings of the nature of accidents, the comparison of the methods focuses on what we can learn from both methods, how, when, and why to apply them. The main finding is that STEP helps to illustrate what happened, involving which actors at what time, whereas FRAM illustrates the dynamic interactions within socio-technical systems and lets the analyst understand the how and why by describing non-linear dependencies, performance conditions, variability, and their resonance across functions.

1. Introduction

Analysing and attempting to understand accidents is an essential part of the safety management and accident prevention process. Many methods may be used for this purpose (see [1,2] for overviews), each reflecting a specific perspective on accidents and how they come about, which may be called an accident model [3,4]. Analysis methods and thus their underlying (implicit or explicitly articulated) accident models affect what investigators look for, which contributory factors are found, and which recommendations are made [5]. Two such methods with underlying models are the Sequentially Timed Events Plotting (STEP) method [6] and the Functional Resonance Accident Model with the associated Functional Resonance Analysis Method (FRAM) [3,7].

Multi-linear event sequence models and methods (such as STEP) have been used in accident analysis to overcome the limitations of simple linear cause-effect approaches to accident analysis. In STEP, an accident is a special class of process where a perturbation transforms a dynamically stable activity into unintended interacting changes of states with a harmful outcome. In this multi-linear approach, an accident is viewed as several sequences of events and the system is decomposed by its structure consisting of interacting events in sequences or in parallel.

Researchers have argued that linear approaches fail to represent the complex dynamics and interdependencies commonly observed in socio-technical systems [3,4,8–11]. Recently, systemic models and methods have been proposed that consider safety as an emergent property of the socio-technical system as a whole.

The Functional Resonance Accident Model with its associated Functional Resonance Analysis Method (FRAM; [3]) embodies such a systemic approach. Rather than physical components and sequences of events, functions and function performance are the units of analysis. A function may be defined as “a set of actions that a system performs or is used for, which are valuable for the achievement of a set of goals” [12].

FRAM is based on four principles [13]. First, the principle that both successes and failures result from the adaptations that organizations, groups, and individuals perform in order to cope

---

Abbreviations: A/C, aircraft; AIBN, Aircraft Investigation Board Norway; APP, Oslo approach; CRM, Crew Resource Management; EFIS, Electronic Flight Instrument System; FRAM, Functional Resonance Analysis Method; ft, feet; FRQ, frequency; GA, go-around; G/S, glide slope; GPWS, Ground Proximity Warning System; L, left; LLZ, localizer; NAX541, aircraft identification call sign; NPF, non-pilot flying; OSL, Oslo Gardermoen airport; PF, pilot flying; R, right; RWY, runway; STEP, Sequentially Timed Events Plotting Method; TWR, tower

[*] Corresponding author. Tel.: +47 90 68 06 34; fax: +47 73 59 28 96.
E-mail address: ivonne.a.herrera@ntnu.no (I.A. Herrera).

[1] Present address: Products & Services, LFV Air Navigation Services of Sweden, Norrköping, Sweden.

0951-8320/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2010.06.003

with complexity. Success depends on their ability to anticipate, recognise, and manage risk. Failure is due to the absence of that ability (temporarily or permanently), rather than to the (organizational, human or technical) inability of a system component to function normally. Second, complex socio-technical systems are by necessity underspecified and only partly predictable. Procedures and tools are adapted to the situation, to meet multiple, possibly conflicting goals, and hence, performance variability is both normal and necessary. The variability of one function is seldom large enough to result in an accident. However, the third principle states that the variability of multiple functions may combine in unexpected ways, leading to disproportionately large consequences. Successes and failures are therefore emergent phenomena that cannot be explained by looking solely at the performance of (organizational, human or technical) system components. Fourth, the variability of a number of functions may resonate, causing the variability of some functions to exceed normal limits, the consequence of which may be an accident. FRAM as a model emphasizes the dynamics and non-linearity of this functional resonance, but also its non-randomness. FRAM as a method therefore aims to support the analysis and prediction of functional resonance in order to understand and avoid accidents.

2. Research questions and approach

The main question addressed in this paper is which new insights this latter systemic method provides for the accident analysis in comparison to the former established multi-linear method. Since the accident analysis methods compared in this paper are based on a different understanding of the nature of accidents, the comparison of the methods focuses on what we can learn from both methods, how, when, and why to apply them, and which aspects of these methods may need improvement.

The paper compares STEP and FRAM in relation to a specific incident to illustrate the lessons learned from each method. The starting point of the study is the incident investigation report. A short description of STEP and FRAM is included. For a more comprehensive description, the reader is referred to references [6,3]. Since different methods invite for different questions to be asked, it was necessary to interview air traffic controllers, pilots, and accident investigators to acquire more information. The information in this paper was collected through interviews and workshops involving a total of 50 people. The analysis with STEP and FRAM was an iterative process between researchers and operative personnel.

3. Summary of the incident

A Norwegian Air Shuttle Boeing 737–36N with call sign NAX541 was en-route from Stavanger Sola airport to Oslo Gardermoen airport (OSL). The aircraft was close to Gardermoen and was controlled by Oslo Approach (APP). The runway in use at Gardermoen was 19R. The aircraft was cleared to descent to an altitude to 4000 ft. The approach and the landing were carried out by the co-pilot as ‘‘pilot-flying’’ (PF) and the captain as ‘‘pilot non-flying’’ (PNF). Shortly after clearance to 4000 ft, the crew was informed that runway 19R was closed because of sweeping and that the landing should take place on runway 19L. The position of the aircraft was instructed by air traffic control to land on 19L. Changing of the runway from 19R to 19L caused a change in the go-around-altitude from 4000 ft at 19R to 3000ft at 19L. The crew performed a quick briefing for a new final approach.

During the final approach, while the aircraft was established on the localizer (LLZ) and glide slope (G/S) for runway 19L, the glide slope signal failed. It took some time for the pilots to recognise G/S failure. At the same time APP instructed the pilots to switch to tower (TWR) frequency. The pilots acknowledged the new frequency but did not yet switch. Immediately after the glide path signal disappeared the aircraft increased its descent rate to 2200ft/min while being flown manually towards LLZ-minima. The aircraft followed a significantly lower approach than intended and was at its lowest only 460 ft over ground level at 4.8 DME. The altitude at this distance from the runway should have been 1100 ft higher. The crew initiated go-around (GA) because the aircraft was still in dense clouds and it drifted a little from the LLZ at OSL. However, the crew did not notice the below-normal altitude during approach. Later a new normal landing was carried out.

The executive summary of the Norwegian Accident Investigation Board (AIBN) [14] explains that the investigation was focused on the glide slope transmission, its technical status and information significance for the cockpit instrument systems combined with cockpit human factors. The AIBN understanding of the situation attributes the main cause of the incident to the pilots' incorrect mental picture of aircraft movements and position. The report concludes that the in-cockpit glide slope capture representation was inadequate. In addition, the report points to a deficiency in the procedure for transfer of responsibility between approach and tower air traffic control.

Five recommendations resulted from the AIBN investigation. The first recommendation is that the responsibility between controls centres should be transferred 8 NM before landing or at acceptance by radar hand over. The second recommendation is related to the certification of avionics displays, advising the verification of the information provided to pilots, with special attention to glide slope and auto-pilot status information. Third, training should take into account glide slope failures after glide slope capture under ILS approach. Fourth, Oslo airport should consider the possibility of providing radar information to the tower controller to be able to identify approach paths deviations. The last recommendation is for the airline to consider situational awareness aspects in the crew resource management (CRM) training.

4. Sequentially Timed Events Plotting

STEP provides a comprehensive framework for accident investigation from the description of the accident process, through the identification of safety problems, to the development of safety recommendations. The first key concept in STEP is the multi-linear event sequence, aimed at overcoming the limitations of the single linear description of events. This is implemented in a worksheet with a procedure to construct a flowchart to store and illustrate the accident process. The STEP worksheet is a simple matrix. The rows are labelled with the names of the actors on the left side. The columns are labelled with marks across a time line.

Second, the description of the accident is performed by universal events building blocks. An event is defined as one actor performing one action. To ensure that there is a clear description the events are broken down until it is possible to visualize the process and be able to understand its proper control. In addition, it is necessary to compare the actual accident events with what was expected to happen.

A third concept is that the events flow logically in a process. This concept is achieved by linking arrows to show proceed/follow and logical relations between events. The result of the third concept is a cascading flow of events representing the accident process from the beginning of the first unplanned change event to the last connected harmful event on the STEP worksheet.

The organization of the events is developed and visualized as a "mental motion picture". The completeness of the sequence is validated with three tests. The row test verifies that there is a complete picture of each actor's actions through the accident. The column test verifies that the events in the individual actor rows are placed correctly in relation to other actors' actions. The necessary and sufficient test verifies that the early action was indeed sufficient to produce the later event, otherwise more actions are necessary.

The STEP worksheet is used to have a link between the recommended actions and the accident. The events represented in STEP are related to normal work and help to predict future risks. The safety problems are identified by analysing the worksheet to find events sets that constitute the safety problem. The identified safety problems are marked as triangles in the worksheet. These problems are evaluated in terms of severity. Then, they are assessed as candidates for recommendations. A STEP change analysis procedure is proposed to evaluate recommendations. Five activities constitute this procedure. The identification of countermeasures to safety problems, the ranking of the safety effects, assessment of the trade-off involved the selection of the best recommendations, and a quality check.

5. Application of STEP to NAX541

The incident is illustrated by a STEP. Due to page and paper limitations, Fig. 1 illustrates a small part of the STEP diagram that was created based on the incident report. In Fig. 1, the time line is on along the *X*-axis and the actors are on the *Y*-axis. An event is considered to mean an actor performing one action. The events are described in event building blocks, for example "APP request to A/C to change to TWR frequency". An arrow is used to link events. Safety problems are illustrated on the top line by triangles in the incident process. Three such problems were identified: (1) no communication between aircraft 1 and tower (triangle 1 in Fig. 1); (2) changed roles between PF and PNF not coordinated; and (3) pilots not aware of low altitude (2 and 3 not shown in simplified figure).

6. Functional Resonance Analysis Method

FRAM promotes a systemic view for accident analysis. The purpose of the analysis is to understand the characteristics of system functions. This method takes into account the non-linear propagation of events based on the concepts of normal performance variability and functional resonance. The analysis consists of four steps (that may be iterated):

Step 1: Identifying essential system functions, and characterizing each function by six basic parameters. A function is defined as an action of a component of the system. The nature of the functions may be technological, human, organizational or a coupling between human, technology and/or organization. The functions are described through six aspects, in terms of their input (I, that which the function uses or transforms), output (O, that which the function produces), preconditions (P, conditions that must be fulfilled to perform a function), resources (R, that which the function needs or consumes), time (T, that which affects time availability), and control (C, that which supervises or adjusts the function), and may be described in a table and subsequently visualized in a hexagonal representation (FRAM module, Fig. 2). The main result from this step is a FRAM "model" with all basic functions identified.

Step 2: Characterizing the (context dependent) potential variability through common performance conditions. Eleven common performance conditions (CPCs) are identified in the FRAM method to be used to elicit the potential variability: (1) availability of personnel and equipment; (2) training, preparation, competence; (3) communication quality; (4) human–machine interaction, operational support; (5) availability of procedures; (6) work conditions; (7) goals, number, and conflicts; (8) available time; (9) circadian rhythm, stress; (10) team collaboration; and (11) organizational quality. These CPCs address the combined human, technological, and organizational aspects of each function. After identifying the CPCs, the variability needs to

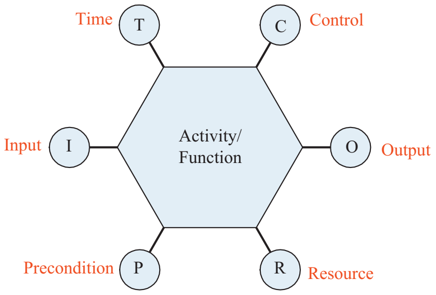


Fig. 2. A FRAM module.

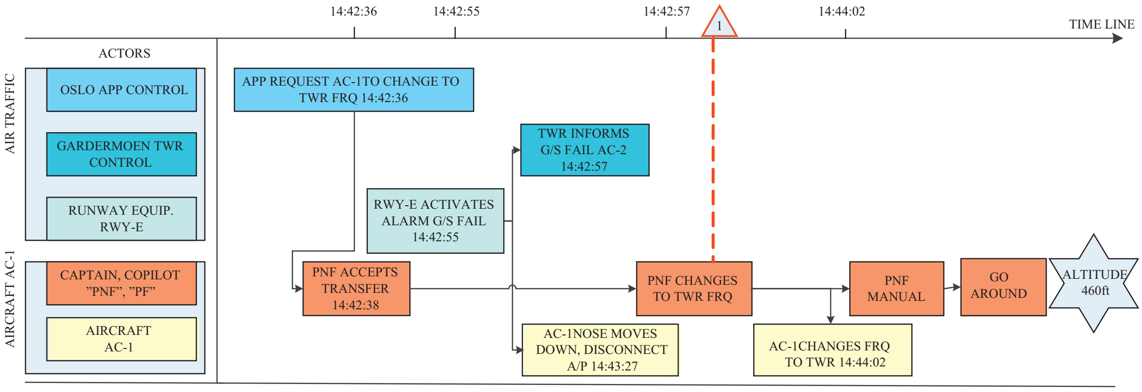


Fig. 1. STEP applied to NAX541 incident (simplified example).

be determined in a qualitative way in terms of stability, predictability, sufficiency, and boundaries of performance.

Step 3: Defining the functional resonance based on possible dependencies/couplings among functions and the potential for functional variability. The output of the functional description of step 1 is a list of functions each with their six aspects. Step 3 identifies instantiations, which are sets of couplings among functions for specified time intervals. The instantiations illustrate how different functions are active in a defined context. The description of the aspects defines the potential links among the functions. For example, the output of one function may be an input to another function, or produce a resource, fulfil a precondition, or enforce a control or time constraint. Depending on the conditions at a given point in time, potential links may become actual links; hence produce an instantiation of the model for those conditions. The potential links among functions may be combined with the results of step 2, the characterization of variability. That is, the links specify where the variability of one function may have an impact, or may propagate. This analysis thus determines how resonance can develop among functions in the system. For example, if the output of a function is unpredictably variable, another function that requires this output as a resource may be performed unpredictably as a consequence. Many such occurrences and propagations of variability may have the effect of resonance; the added variability under the normal detection threshold becomes a 'signal', a high risk or vulnerability.

Step 4: Identifying barriers for variability (damping factors) and specifying required performance monitoring. Barriers are hindrances that may either prevent an unwanted event to take place, or protect against the consequences of an unwanted event. Variability is materialised due to trade-offs in face of multiple conflicting goals within available time. In this context, it is necessary to have barriers that both damp the unwanted variability and facilitate desirable variability. Hence, barriers can be seen as both hindrances and enablers. On the one hand, barriers may either prevent an unwanted event from taking place, or protect against the consequences of an unwanted event. On the other hand, they may enhance the capabilities allowing the system to continue its operation. Barriers can be described in terms of barrier systems (the organizational and/or physical structure of the barrier) and barrier functions (the manner by which the barrier achieves its purpose). In FRAM, four categories of barrier systems are identified (each with their potential barrier functions):

(1) Physical barrier systems block the movement or transportation of mass, energy, or information. Examples include fuel tanks, safety belts, and filters.
(2) Functional barrier systems set up preconditions that need to be met before an action (by human and/or machine) can be undertaken. Examples include locks, passwords, and smoke detectors.
(3) Symbolic barrier systems are indications of constraints on action that are physically present. Examples include signs, checklists, alarms, and clearances. Potential functions encompass preventing, regulating, and authorizing actions.
(4) Incorporeal barrier systems are indications of constraints on action that are not physically present. Examples include ethical norms, group pressure, rules, and laws.

Besides recommendations for barriers, FRAM is aimed at specifying recommendations for the monitoring of performance variability, to be able to detect unwanted variability. Function definition and characterization allow understanding aspects that affect performance. These aspects are candidate for indicators. Instantiations can be used as a basis to consider the effect of the variability across and within functions. Relevant indicators of the spreading of variability may be related to beneficial or disadvantageous changes in potential, expected, and actual couplings. Functional modelling with FRAM aims identification indicators that provide information about the variability of normal performance of the system.

7. Application of FRAM to NAX541

Step 1 is related to the identification and characterization of functions: A total of 19 essential functions were identified and grouped in accordance to the area of operation. There are no specified rules for the 'level of granularity', instead functions are included or split up when the explanation of variability requires. In this particular analysis some higher level functions, e.g. 'Oslo APP control', and some lower level functions, e.g. 'Change frequency (frq) to TWR control'.

The operative areas and functions for this particular incident are:

- Crew operations: change runway (RWY) to 19L, new final approach briefing, auto-pilot approach (APP), change APP frq to TWR frq, manual approach, GO-AROUND, landing, approach, receiving radio communication, and transmitting radio communication
- Avionics functions: disconnect auto-pilot (A/P), Electronic Flight Instrument (EFIS), and Ground Proximity Warning System (GPWS)
- Air traffic control: Oslo APP control, RWY sweeping, glide slope transmission, and Gardermoen TWR control
- Aircraft in the vicinity: aircraft (A/C)-2 communication and A/C-3 communication

The NAX541 incident report contains information that helps to define aspects of functional performance. Essential functions are described with these aspects. Table 1 shows an example of the aspects of the function 'manual approach'. Similar tables were developed for 18 other functions.

In step 2 the potential for variability is described using a list of common performance conditions (CPCs). Table 2 presents an example of CPCs for the function 'manual approach'.

The description of variability is based on the information registered in the incident report combined with a set of questions based on the CPCs. Since little of this information regarding variability was available, it was necessary to interview operational personnel (air traffic controllers, pilots). An example is for CPC 'human–machine interface (HMI), operational support', a question was how aware pilots are of these EFIS, GPWS discrepancies, some stated "Boeing manuals explain which information is displayed, it is normal to have contradictory

Table 1
A FRAM module function description.

Function: manual approach	Aspect description
Input	GPWS alarms, pilot informed of G/S failure
Output	Altitude in accordance with approach path, Altitude lower/higher than flight path
Preconditions	A/P disconnected
Resources	Pilot flying, pilot non-flying
Time	Efficiency thoroughness trade-off, time available varies
Control	SOPs

Table 2
Manual flight approach CPCs.

Function: manual approach	Performance conditions	Rating
Availability of resources (personnel, equipment)		Adequate
Training, preparation, competence	PF little experience on type	Temporarily inadequate
Communication quality	Delay to contact tower	Inefficient
HMI operational support	Unclear alerts	Inadequate
Avail. procedures		Adequate
Work conditions	Interruptions?	Temporarily inadequate?
# Goals, conflicts	Overloaded	More than capacity
Available time	Task synchronisation	Temporarily inadequate
Circadian rhythm		Adjusted
Team collaboration	Switched roles	Inefficient
Org. quality		

information. In this case, understanding of the system as a whole is required. Pilots needs to judge relevant information for each situation." An additional example of questions for the function "Runway change", was if it is normal and correct to request runway change with such a short notice? The interviews identified that there are no formal operational limits for tower air traffic controllers, but for pilots there are. Thus an understanding of performance and its variability was obtained.

In step 3 links among functions are identified for certain time intervals. States are identified to be valid during specific time intervals, which define links among the aspects of functions, hence instantiate the model. An example instantiation is presented in Fig. 3, where links during the time interval 14:42:37–14:43:27 of the incident are described as an instantiation of the FRAM that resulted from step 1. Many more such instantiations may be generated, but here only one example can be shown.

To understand the events in relation to links and functions in this instantiation, numbers 1–5 and letters a–d have been used to illustrate two parallel processes. Following the numbers first, the APP controller communicates to the pilot that they should contact TWR at the TWR frequency (1). This is an output of 'Oslo APP control', and an input to 'receiving radio communication'. This latter function thus has as output the state that transfer is requested to the TWR frequency (2), which matches the preconditions of 'change APP frq to TWR frq', and 'transmitting radio communication'. The fulfilment of this precondition triggers the pilots to acknowledge the transfer to TWR to the APP controller (3), an output of transmitting function, input to 'Oslo APP control'. The pilots however do not switch immediately after the transfer is requested, hence the output is that the frequency still is set to APP, for a much longer time than would be intended (indicated by the red 'O'), and the pilots do not contact TWR (6) until much later. This has consequences for the precondition of receiving/transmitting (4), which is being on the same frequency with the control centre that has responsibility for the flight. With the delay in frequency change, the link that the pilot is in-formed of the G/S failure (5) is also delayed.

At about the same time, following the letters in Fig. 3, 'glide slope transmission' changes output to that there is no G/S signal at 14:42:55 (a), because of a failure of the G/S transmitting equipment (a resource, R in red). This makes the TWR controller in-form pilots on the TWR frequency of the G/S failure (b), excluding the incident aircraft crew because of the unfulfilled precondition because of link (4), de-laying the point that the pilot is informed of G/S failure (d). Concurrently, the loss of G/S no longer fulfils the precondition of the auto-pilot function, with the resulting output of A/P being disconnected (c) about half a minute after G/S loss. This in turn no longer fulfils the precondition of an auto-pilot approach and instead matches the precondition for a manual approach. All of this in turn results in variability on the manual approach, e.g. with decreased availability of time, inadequate control because of PF-PNF collaboration problems, and inadequate re-sources (e.g. displays unclear indications of A/P and G/S) resulting in highly variable performance (out-put) of the manual approach.

Step 4 addresses barriers to dampen unwanted variability and performance variability monitoring where variability should not be dampened. AIBN recommendations could be modelled as barrier systems and barrier functions, e.g. "responsibility between control centres should be transferred 8 NM before landing, or at acceptance by radar hand over." (AIBN, p. 31, our translation). In FRAM terminology this can be described as an incorporeal prescribing barrier. This barrier would have an effect on the variability of the APP and TWR control functions through the aspect of control and the links between input and output in various instantiations describing communication and transfer of responsibility. New suggestions for barriers also result from the FRAM. For example, a proactive communication from TWR to APP when a flight does not report on frequency would link their output and input (see link (*X*) in Fig. 3), triggering instantiations of links 1–6 so that control and contact is re-established. This barrier may be implemented in various systems and functions, such as through regulation, training, procedures, checklists, and display design, etc. The FRAM also points to the interconnectivity of air traffic control and pilot functions, suggesting joint training of these operators with a wide range of variability in the identified functions. As with any method, FRAM enables the suggestion of barriers (recommendations), which need to be evaluated by domain experts in terms of feasibility, acceptability, and cost effectiveness, among other factors.

The FRAM and the instantiations that were created here also point to the future development of indicators for matters such as overload and loss of control when cockpit crew has significant experience differences. Indicators may be identified by considering the variability of the functions and their couplings. In the case, indicators are related to proactive communication between TWR control and APP control or information between TWR and pilots.

8. Comparison

Accident models, implicitly underlying an analysis or explicitly modelling an adverse event, influence the elicitation, filtering, and aggregation of information. Then, what can we learn from the applications of STEP and FRAM to this incident?

STEP is relatively simple to understand and provides a clear picture of the course of the events. However, STEP only asks the question of which events happened in the specific sequence of events under analysis. This means that events mapped in STEP are separated from descriptions of the normal functioning of socio-technical systems and their contexts. For example, the STEP diagram illustrates that the PNF's switch to TWR frequency was delayed, but not why. Instead, STEP only looks for failures and safety problems, and highlights sequence and interaction between events. FRAM refrains from looking for human errors and safety problems but tries to understand why the incident happened. Since FRAM addresses both normal performance variability and the specifics of an adverse event, FRAM broadens data collection of the analysis compared to a STEP-driven analysis: Thus the development of the incident is contextualized

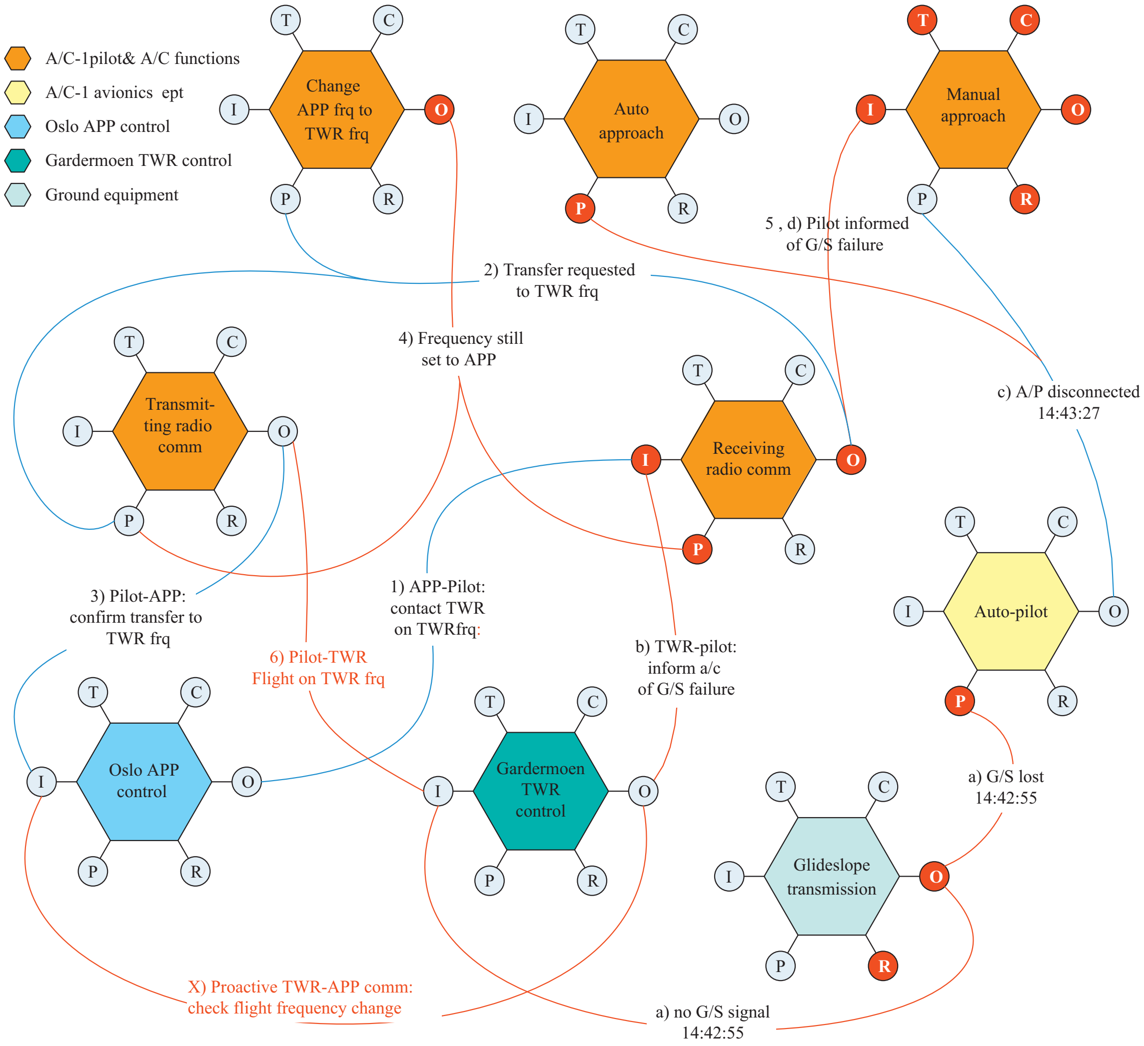


Fig. 3. A FRAM instantiation during time interval 14:42:37–14:43:27 with incident data.

in a normal socio-technical environment. Through asking questions based on the common performance conditions and linking functions in instantiations, FRAM identified additional factors and the context of why performance varied becomes apparent: For example, the operational limits for runway change for different operators were discussed; the question of why the frequency change was delayed gets answered based on the normal variability in pilot-first-officer-interaction patterns in cases of experience difference; the pilots' unawareness of the low altitude is understandable with regard to variability related to e.g. team collaboration and human–machine interface issues.

STEP provides a "mental motion picture" [6] illustrating sequences of events and interactions between processes, indicating *what* happened *when*. FRAM instead sketches a 'functional slide show' with its illustrations of functions, aspects, and emerging links between them in instances, indicating the *what* and *when*, and common performance conditions, variability, and functional resonance, indicating *why*. FRAM's qualitative descriptions of variability provide more gradations in the description of functions than the bimodal (success/failure) descriptions typical for STEP.

In relation to the question of when each method should be used, the type of incident and system to be analysed needs to be taken into account. STEP is suited to describe tractable systems, where it is possible to completely describe the system, the principles of functioning are known and there is sufficient knowledge of key parameters. FRAM is better suited for describing tightly coupled, intractable systems [15], of which the system described in this paper is an example. Because FRAM does not focus only on weaknesses but also on normal performance variability, this provides a more thorough understanding of the incident in relation to how work is normally performed. Therefore, FRAM may lead to a more accurate assessment of the impact of recommendations and the identification of factors left unexplored with STEP that may have a safety impact in the future. While the chain of events is suited for component failures or when one or more components failed, they are less adequate to satisfactorily explain system accidents [4]. This can be seen in the STEP–FRAM comparison here. The STEP diagram focuses on events and does not describe the systems aspects: the understanding of underlying systemic factors affecting performance is left to experts' interpretation. FRAM enables analysts to model these systemic factors explicitly.

9. Conclusions and practical implications

This paper presented two accident analysis methods: the multi-sequential STEP and systemic FRAM. The question of how to apply these methods was addressed by discussing the steps of the methods, illustrated by applying these methods to a missed approach incident. This paper concluded that FRAM provides a different explanation about how events are a result of normal variability and functional resonance, compared to STEP. The main finding is that STEP helps to illustrate what happened, whereas FRAM covers what happened and also illustrates the dynamic interactions within the socio-technical system and lets the analyst understand the how and why by describing non-linear dependencies, performance conditions and variability, and their resonance across functions. Another important finding is that it was possible to identify additional factors with FRAM. STEP interpretation and analysis depend extensively on investigator experience, FRAM guides the analyst more into asking questions for systemic factors and enables the explicit identification of relevant aspects of the accident based on day-to-day operational data. The example also illustrates how unwanted variability is propagated such as the information about G/S failure and the undesired resonance with the differences in pilots' experience. However, several incidents in different contexts would need to be analysed to validate and generalize these findings.

The variability of normal performance which most of the time is unproblematic and actually necessary for the underspecified system to work in practise, may suddenly and unexpectedly resonate with variability in other functions and escalate to a dangerous level. A theoretical implication is that FRAM modelling may violate the binary or sequential logic of other accident models. STEP is based on the description of the accident as a process where a failure for each of the components is described, and the occurrence of a system failure is determined by the state of the components. FRAM qualitative descriptions of variability provide more states in the descriptions of the functions than the bimodal (success/failure) descriptions typical for STEP. FRAM is one of the first methods that moves away from linear cause and effect models of thinking about safety providing a functional systemic approach.

Three practical implications are found. The first is that FRAM provides new ways of understanding failures and successes, which encourages investigators to look beyond the specifics of the time sequence and failure under analysis, moving the analysis into the conditions of normal work. The second is that FRAM models and analyses an intractable socio-technical system within a specific context. Third, since STEP and FRAM are based on different understandings of the nature of accidents, their combined application during accident analysis provides complementary perspectives and may contribute to a more comprehensive understanding of and more effective learning from an incident or accident.

While FRAM as a model has been accepted in the majority of discussions with practitioners in this study, and seems to fill a need for understanding intractable systems, FRAM as a method is still young and needs further development. This article contributes to the development of the method by outlining a way to illustrate instantiations of models for limited time intervals. Moreover, this article indicates the potential of and need for strategies to actively combine methods for incident/accident analysis as part of the analyst's toolbox in order to enable understanding, learning, and prevention. An additional need is the identification of normal/abnormal and desired/undesired variability which this paper has addressed briefly. Remaining challenges include a more structured approach to generating recommendations in terms of barriers, indicators, and redesign of functions, as well as evaluating how well FRAM is suited as a method to collect and organize data during early stages of accident investigation.

Acknowledgements

This work has benefited greatly from the help and support of several aviation experts and the participants in the 2nd FRAM workshop. We are particularly grateful to the investigators and managers of the Norwegian Accident Investigation Board who commented on a draft of the model. Thanks to Ranveig K. Tinmannsvik, Erik Jersin, Erik Hollnagel, Jørn Vatn, Kip Smith, Jan Hovden, Carl Rollenhagen, Arthur Dijkstra and the reviewers of the RESS journal and ESREL 2008 conference, for their insightful comments on our work.

References

[1] Harms-Ringdahl L. Safety analysis. Boca Raton, FL: CRC Press; 2001.
[2] Sklet S. Comparison of some selected methods for accident investigation. Journal of Hazardous Materials 2004;111:29–37.
[3] Hollnagel E. Barriers and accident prevention. Aldershot, UK: Ashgate; 2004.
[4] Leveson, N. Evaluating accident models using recent aerospace accidents. Technical Report, MIT Dept. of Aeronautics and Astronautics; 2001.
[5] Lundberg J, Rollenhagen C, Hollnagel E. What-You-Look-For-Is-What-You-Find—The consequences of underlying accident models in eight accident investigation manuals. Safety Science 2009;47:1297–311.
[6] Hendrick K, Benner L. Investigating accidents with STEP. New York: Marcel Dekker Inc.; 1987.
[7] Hollnagel, E. From FRAM to FRAM. Presentation at the 2nd FRAM Workshop, Sophia-Antipolis, France; 2008.
[8] Amalberti R. The paradoxes of almost totally safe transportation systems. Safety Science 2001;37:109–26.
[9] Dekker SWA. Ten questions about human error: A new view of human factors and system safety. Mahwah, NJ: Lawrence Erlbaum; 2004.
[10] Rochlin GI. Safe operation as a social construct. Ergonomics 1999;42: 1549–60.
[11] Woods DD, Cook RI. Nine steps to move forward from error. Cognition, Technology & Work 2002;4:137–44.
[12] Woltjer, R. Functional modeling of constraint management in aviation safety and command and control. Linköping Studies in Science and Technology, Dissertation No. 1249, Linköping University, Sweden; 2009.
[13] Hollnagel E, Pruchnicki S Woltjer R Etcher S. Analysis of Comair flight 5191 with the Functional Resonance Accident Model. In: Proceedings of the eighth international symposium of the Australian Aviation Psychology Association, Sydney, Australia; 2008.
[14] AIBN. Rapport etter alvorlig luftfartshendelse ved Oslo Lufthavn Gardermoen 9. February 2003 med Boeing 737-36N, NAX541, operert av Norwegian Air Shuttle. Aircraft Investigation Board Norway, SL RAP.:20/2004.
[15] Hollnagel E. The changing nature of risks. Ergonomics Australia 2008;22: 33–46.

Paper V

Aviation safety and maintenance under major organizational changes, investigating non-existing accidents

Journal of accident analysis and prevention (2009)
vol 41, iss. 6, pp 1155-1163

Invited paper for Special Issue 17th European Safety and Reliability Conference, ESREL2006, held in Estoril, Portugal, on 18–22 September 2006

Accident Analysis and Prevention 41 (2009) 1155–1163

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Aviation safety and maintenance under major organizational changes, investigating non-existing accidents

Ivonne A. Herrera[a,*], Arve O. Nordskag[a], Grete Myhre[b], Kåre Halvorsen[b]

[a] SINTEF Technology and Society, Department of Safety and Reliability, NO-7465 Trondheim, Norway
[b] Accident Investigation Board Norway, P.O. Box 213, NO-2003 Lillestrøm, Norway

ARTICLE INFO

Article history:
Received 14 June 2007
Received in revised form 18 April 2008
Accepted 6 June 2008

Keywords:
Safety
Indicators
Maintenance
Aviation
Organizational change

ABSTRACT

The objective of this paper is to discuss the following questions: Do concurrent organizational changes have a direct impact on aviation maintenance and safety, if so, how can this be measured? These questions were part of the investigation carried out by the Accident Investigation Board, Norway (AIBN). The AIBN investigated whether Norwegian aviation safety had been affected due to major organizational changes between 2000 and 2004. The main concern was the reduction in safety margins and its consequences.

This paper presents a summary of the techniques used and explains how they were applied in three airlines and by two offshore helicopter operators. The paper also discusses the development of safety related indicators in the aviation industry. In addition, there is a summary of the lessons learned and safety recommendations. The Norwegian Ministry of Transport has required all players in the aviation industry to follow up the findings and recommendations of the AIBN study.

1. Introduction

The situation in Norwegian aviation between 2000 and 2004 was that several major changes occurred at the same time in the organizations of the regulators, the airports, the navigation providers, and the operators. In addition, the effect of other developments in the aviation industry took place during this period. These included deregulation, liberation, privatization, cost reductions and the growth of the low cost carriers.

The objective of the study led, by the AIBN (2005), was to investigate how aviation safety is maintained in light of major change processes in Norwegian aviation. The safety status of the airlines and helicopter operators was established through an evaluation of the safety in the management of maintenance. In the study, it is assumed that changes affect safety in the operational and maintenance parts of the companies in similar ways. The main reason for selecting maintenance, and not the operational part of the organizations, was the opportunity to collect more quantitative data.

This paper describes a method that assesses how safety is maintained in the maintenance organizations of the airlines while there are concurrent ongoing internal and external organizational changes. The paper focuses on the development of safety related maintenance indicators. We discuss the application and results of the method applied in three airlines and two offshore helicopter operators.

2. Literature survey

The literature survey constituted the theoretical approach to the study. A combination of the gathered literature and empirical data were used in the development of a set of indicators, the analysis of results and the achievement of documented conclusions. No single theory has been identified that comprises many parallel changes, the management of safety and maintenance. It was therefore necessary to perform a review of existing approaches and methods that could be applied to the assessment of the management of safety under major organizational changes.

2.1. Maintenance and safety

Maintenance errors are estimated to contribute to 12% of major airline aircraft accidents and 50% of engine-related flight delays (Hobbs, 2000). Hobbs (2004) indicated that deficient maintenance and inspection could be considered the second largest safety threat after Controlled Flight into Terrain (CFIT). A study regarding

---

Abbreviations: AIBN, Accident Investigation Board Norway; ASR, Air Safety Report; CAA, Civil Aviation Authorities; DISP, Dispensation; EASA, European Aviation Safety Agency; FOR, Flight Occurrence Report; GOR, Ground Occurrence Report; HIL, Hold Item List; ICAO, International Civil Aviation Authority; JAA, Joint Aviation Authorities; MEL, Minimum Equipment List.

* Corresponding author. Tel.: +47 90680634; fax: +47 73592891.
E-mail address: ivonne.a.herrera@sintef.no (I.A. Herrera).

0001-4575/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2008.06.007

selected aviation accidents in the Unites States between 1996 and 2003 indicates that managerial issues and regulatory failures are classified as probable causes with 17% and 12%, respectively (Holloway and Johnson, 2003). The ICAO's working paper (2006) recommendation was to use a more proactive approach to prevent accidents. The proactive approach includes a more scientific approach to risk assessment, human factors and the development of means for collecting and analyzing data.

In order to develop a generic model to evaluate safety in maintenance activities, it was necessary to take into account experience from different industries including the aeronautical industry. Hale et al. (1998) presented a model to evaluate safety in maintenance management. This maintenance management model consisted of three levels (i) policy, (ii) planning and procedures and (iii) execution and feedback. The model enabled an evaluation to be made of how safety is managed at all levels from the formal establishment of a safety policy through its application on other levels. The model took into account risk identification and management for single maintenance and combined maintenance tasks. A theoretical model, an audit checklist and a questionnaire were developed to evaluate safety aspects in the management of maintenance. Some results from Hale's study that are relevant to AIBN maintenance study are:

- The model indicates maintenance levels and essential elements that should be taken into account to assess the management of safety. These elements are policy, corrective and preventive maintenance, modifications, maintenance tasks, engineering orders, inspection program, scheduling, planning and execution of maintenance work, reporting and analysis.
- General weaknesses regarding the translation of a safety policy to the other level in the middle maintenance concept, design, planning and resource management.
- A warning to management to pay attention to the complete line of communication before "surgery" reduces or hives off departments.

Sachon and Paté-Cornell (2000) developed a model to correlate delays and safety in airline maintenance. The model proposed the utilization of Probabilistic Risk Analysis (PRA) including management factors. Risk modeling is utilized to link the ground model and the flight model to management decisions. The decisions considered that are relevant to the AIBN study were (i) qualification of maintenance personnel, (ii) number of deferrals allowed and (iii) timing of maintenance operations. An interesting result was that the model supports the marginal trade-off between minimizing delay and maximizing safety. The limitations of this model include the completeness, the structure of the model and the estimation of model parameters.

Human factors in maintenance is a topic of increasing interest. Regulations now require human factor training for maintenance personnel. It has been stated that there still is a challenge in how human factors are integrated in all maintenance activities (Marx, 1998). In the Human Centred Systems for Aircraft Dispatch and Maintenance Safety (ADAMS) project, relevant findings to the AIBN study were related to:

- The ability of quality and safety systems to deal effectively with everyday performance of maintenance tasks.
- Quality systems dealing with maintenance procedures did not follow "short cuts" to do things "better" and "quicker".
- Quality systems stimulating effective solutions.

Baranzini and Cromie (2002) noted the importance of teamwork in aviation. In their paper, it is remarked that factors like design and operation of the entire workflows can be evaluated to ensure proficient operation across teams. This approach is interesting to the AIBN study since airlines have a tendency to subcontract their maintenance activities and it is therefore a need to coordinate activities between different companies. Subcontracting is demanding when it comes to communication between operator and maintenance facility.

2.2. *Safety indicators*

The development of safety indicators is often seen as an integral part of safety management. The safety indicators can be divided in two categories outcome-based indicators (reactive) and activity indicators (proactive). The outcome-based indicators will measure following an unwanted occurrence. The activity indicators were measures of efforts to prevent accidents or incidents. Studies from different industries related to development and use of indicators were reviewed; relevant findings are presented in the following paragraphs.

The nuclear industry proposed a set of safety performance indicators as a measure of safety performance (Dahlgren et al., 2001). Several potential benefits regarding the use of indicators were identified. Some of the benefits are the identification of an objective, auditable and non-disputable set of safety parameters and the procurement of insight in relation to what is important to safety. On the other hand, indicators cannot be used alone to draw conclusions about safety performance. Also indicators can be misused and manipulated. An indicator trend is not necessarily the sole indicative factor in safety performance, and finally indicators cannot replace qualitative engineering judgment.

The chemical industry prepared guidelines about how to develop a safety performance indicator program (OECD, 2003). An interesting aspect is that the use of safety indicators is recommended for both the industry as well as the regulator. In the guidelines the first step is the identification and adaptation of safety indicators. The types of indicators proposed are outcome and activity indicators. The second step includes the quantification and weighting of the indicators. As a last step, it is recommended to apply these indicators on a regular basis.

The aviation industry has focused on measuring reactive indicators as a safety measure. However, it is recognized that these indicators do not provide a complete picture of the safety level. Recently the American Department of Energy (US DOE, 2005) published a guide to performance indicators to be applied to government-used aircraft. The reason for using performance indicators was to have a structured approach to measure the performance of key processes to ensure a safety operation. Pilot competence, maintenance quality and management attitude were aspects which were focused on during the development of the indicators. The maintenance indicators focused on factors affecting aircraft availability. Indicators related to maintenance included mean time between failures, aircraft discrepancies, availability of competent technicians and maintenance scheduling effectiveness. The set of indicators includes a combination of reactive and activity indicators.

Reason (1991) introduced the concept of safety space, to indicate how organizations might move from a relatively safe dimension to an unsafe dimension. Reason claims that the driving forces which influence the drifting in the safety space are commitment, competence and cognizance. The importance of safety awareness in the organization is pointed out as being essential for controlling drifting. One tool to enhance safety awareness is the use of proactive safety state indicators.

2.3. Organizational change, safety and maintenance

The REACH Investigation report (Van der Geest et al., 2003) followed after the Swiss aviation accidents, analyzed the management of safety in Switzerland. The study used two processes, one for evaluation of safety management at a national level by using a Safety Policy Process model, and the other to analyze individual organizations by a Safety Management Process. Since the REACH investigation covers the complete aviation system in one country, the AIBN used this approach as the starting point in the planning phase of their investigation. The Swiss report was used in the AIBN study, for the maintenance case and the evaluation of the safety policy and its application within the management activities was analyzed.

The AIBN investigation of organizational change and safety introduced emotional reactions towards the process of major organizational change as an explanatory cause for incidents and accidents (Tveiten et al., 2006). It has been pointed out that it is necessary to include conflicts and alienation interacting with changes as factors that influence safety margins.

3. Research approach to link organizational changes, maintenance and safety

The research involved a multidisciplinary approach combining human, technological and organizational aspects on maintenance and safety. One of the main differences in the selected approach is that there was no limitation to identify the particular aspects affecting safety. The approach selected was focused on gathering as much data as possible to be analyzed by different methods. This is in line with what Hollnagel's arguments (2006) related to limitation in traditional investigation approaches "what you look for is what you find and what you find is what you fix".

The selected strategy was both theoretical and empirical. The theoretical part consisted of a literature survey regarding methods and lessons learned that were applicable to organizational changes, safety and maintenance. The aviation industry is highly regulated and demands companies to document all processes from policies, procedures to practices. Therefore the empirical part consisted of gathering company's management documentation, data and interviews. The purpose of this review was to identify gaps between management documentation, procedures and current practices affecting safety. These gaps should have a correlation to organizational changes.

Different sources of information and methods were used to provide a triangularization of data. A variety of methods were selected to collect as much relevant information as possible. The methods were (i) analysis of documentation, (ii) interviews, (iii) the use of safety related indicators and (iv) a questionnaire. The information was analyzed from different perspectives by a multidisciplinary team in order to achieve possible firm conclusions. It was a challenge to correlate organizational changes to maintenance and their impact on safety. The process presented in this paper is shown in Fig. 1.

3.1. Company documentation

The information needed was gathered from four different perspectives. These activities are outlined in Fig. 2, including a mapping of significant external, organizational and other changes supposed to affect the maintenance function, collection of company data and documentation regarding safety indicators and significant changes within the company. Significant changes such as level of qualification and training, safety management system and maintenance programs were taken into account.

The following information was collected:

- Significant changes affecting companies' maintenance activities.
- Trends in selected safety indicators.
- Companies' follow up and risk management from a safety perspective.
- Companies' risk management in the process of change.

The management documentation provided information regarding how the companies managed maintenance from the establishment of the policy until reporting after execution of maintenance activities (Hale et al., 1998). This information provided an overview of the company procedures to identify how the company manages safety.

3.2. Identification of safety indicators

It is recognized that accident and incident rates do not give a complete picture of the "health" of a system. To approach this challenge, a set of indicators for measuring safety in aviation was developed in cooperation with the Swedish and Norwegian Aviation Authorities (Tinmannsvik, 2005). These indicators were developed to identify the consequences of changes that could have a safety impact.

The safety indicators selected in the study were outcome-based activity indicators. The outcome-based indicators included were accident and incident rates, discrepancies reports and absence due to sickness. The activity indicators were defined in groups (i) external–internal audits, (ii) competence training and level of experience, (iii) maintenance, and (iv) financial investments. The indicators were classified in accordance with the importance in the monitoring of safety trends as high importance, average importance and minor importance (Herrera and Tinmannsvik, 2006). After selection of indicators, quantification is normalized by the amount of flight hours when required. A detailed list of indicators is provided in Appendix A.

The outcome and activity indicators are in line with the indicators proposed within the chemical industry. In relation to the Hale model (Hale, 1998), the indicators proposed a maintenance measuring program corrected by technical reports, preventive by changes in the maintenance schedule, inspection and monitoring by audits, the execution of maintenance program is evaluated by the Hold Item List indicator (backlog), reporting and analysis by technical and flight reports. Competence and training are also included in the indicators, this is supported by Reason and Hobbs (2002) and Hollnagel and Woods (2006) who noted the importance of the competence of personnel to support the ability of the organization to recover from unwanted situations.

3.3. Questionnaire

A questionnaire was prepared to identify inadequacies or safety impacts of newly executed and ongoing changes. The topics in the questionnaire were personal data, safety culture, management of safety, changes within the company (reporting, auditing, training, work practices), changes affecting safety (positive and negative). The questionnaire consisted of 175 questions for the maintenance personnel some questions were repeated in different formulations to ensure the consistency of answers which was confirmed by the statistical analysis.

The questionnaire was sent out in eight different versions, one for each of the following groups: pilots, managers, cabin crew, air

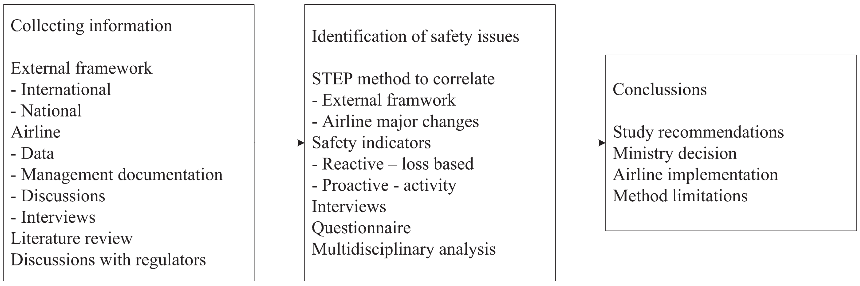


Fig. 1. Proposed approach to evaluate concurrent changes and safety impact.

traffic controllers, technicians, planners and engineers, employees in ground service and employees in the civil aviation authority.

The questionnaire for the maintenance personnel consisted of 175 questions. The questions were repeated in different formulations to ensure the consistency of answers which was confirmed by the statistical analysis. Since the AIBN investigation had a holistic view on the Norwegian aviation safety, not all questions where related to maintenance only. In spite of being a lot of questions, there were no significant change in the way they where answered throughout the questionnaire.

4. Results and principal findings

4.1. Linking organizational changes to safety

The baseline of the study started by mapping significant changes in external conditions was thought to affect aviation maintenance. The objective is identifying the changes that may have a safety impact. The information was collected by reviewing regulations, company management documents and having reviews with maintenance personnel managers and technicians.

The changes were illustrated by a Sequentially Timed Events Plot (STEP) Hendrick and Benner (1987). STEP is a method to systematically process accident investigation based on sequences of multi-linear events. STEP was selected because it fitted to the purpose of the study, was relative simple to use and provided a clear picture of simultaneous events and different relationships between the events. Sklet (2004) and Johnson (2003) present more comprehensive information about methods for accident investigation. STEP is adapted in order to the study to link organizational changes and their impact in the organizations. STEP enabled an illustration to be made of the interaction among different levels from the international organizations, the Ministry of Transport, the national regulators and then finally the operators. An example of STEP is given in Fig. 3 where international aviation organizations' training recommendations change had an impact on the operators training procedures and practices.

The safety related issue is illustrated on the top; in this example being “Conversion to B1, B2 certificates”. Due to unclear new requirements, the conversion to a new certificate was identified as safety related due to risk of having a certificate to perform tasks without ensuring relevant competence.

STEP provided an overall picture of changes that occurred in the period 2000–2004. This diagram linked the originator of the change and the actors (operators and regulators). It was possible to identify both the changes that had a positive impact and the changes that had a negative impact. A major result of this method was the identification of many changes occurring at the same time and that the speed of the changes were high (e.g. changes within the Civil Aviation Authorities, Norway CAA-NO). The following central changes were identified: merging of airlines, changes in the regulatory body, changes in the regulations, changing in the competition situation and increased focus on cost efficiency. Changes in the regulations have a direct impact in the organization and management activities within the airline.

4.2. What did we learn from indicators?

The activity indicators applied to airlines included internal audits and findings, findings older than 6 months, number of

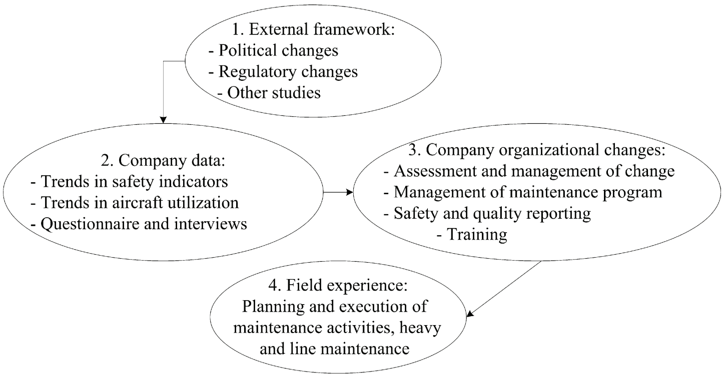


Fig. 2. Data collection from different perspectives.

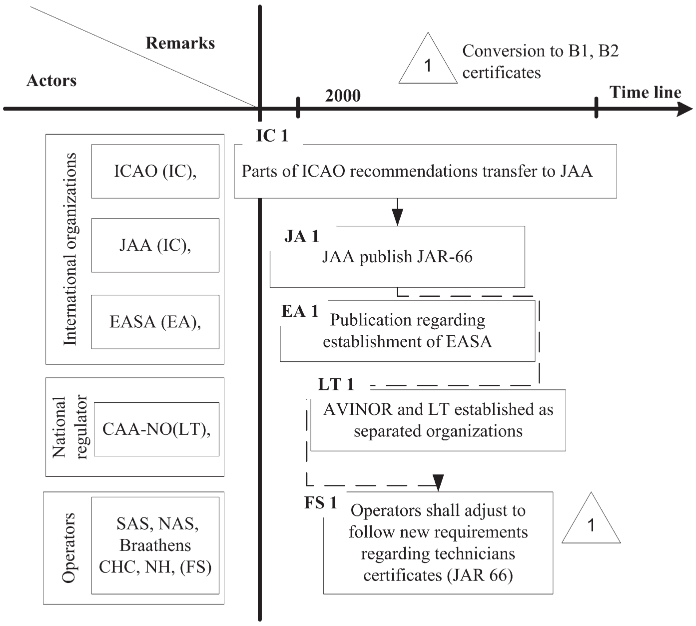


Fig. 3. Using STEP to link changes in the organizations and their safety impact.

requested and approved dispensations per year, number of maintenance certified staff, experience, number of maintenance training (technical and human factors), number of technicians participating in maintenance courses, number of working hours and shift changes, percentage of maintenance program based on company experience (maintenance program optimization).

An extensive list requesting data was presented to the operators. Based on data available to the majority operators, a set of indicators was selected. These were related to airline activities and aircraft utilization, maintenance costs, number of aircraft, flight hours, cycles, maintenance certified and non-certified staff, in-house maintenance and subcontract maintenance. In order to correlate safety and maintenance, it was necessary to have an additional set of indicators. Fig. 4 gives an example of safety indicators.

The figure illustrates how a combination of reactive (outcome) and proactive (activity) indicators could provide sign of the safety health of the organization. In addition, Fig. 4 shows the known phenomenon of a few accidents and a large amount of errors and recoveries (operation with no errors). An interesting aspect is while strong focus is on learning from rare accidents; there is no tradition to analyze successes (normal operations with no delays when the organization recovers from a failure that could have a safety impact).

The main results from the analysis of the indicators were:

- Increase in reporting incidents at the same time the companies attribute this to an increased approval of the importance of reporting.
- Technical dispensations, this indicator could be related to the company's ability to correct faults. Over the last period the use of dispensations was normalized.
- Minimum Equipment List (MEL) is a list of instruments and equipment on an aircraft allowing it to be operated with some of those instruments or equipment inoperative. Number of inoperative MEL items is an indicator saying something about the technical condition of the aircraft in relation to safety critical systems. The airlines with aircraft of the same type showed a slightly increase in the number of MEL.items over the period. One of the reasons for the increase of MELs was lack of available parts.
- Open items in the Hold Item List (HIL) are related to maintenance tasks that have not been accomplished as planned. This indicator is related to non-critical equipment and could say something related to the ability to carry out preventive maintenance. HIL for some operators demonstrates a stable trend and with an increase towards the end of the period.
- A reduction in line personnel and reduction in in-house training departments.

An important topic in this paper is the relation between safety indicators and safety level. Based on the statistical data, it is possible to argue that the increase of MEL has a direct relation to a diminution in the technical condition of the aircraft and may represent a diminution in the safety margins. From the theory, Reason and Hobbs (2002) argue that a reduction in the competence affects the ability of the organization to recover from an anomaly and it is therefore argued in this paper that there might be a relation between reduction of personnel and the impact in the increase of MEL and HIL.

The results of the analysis of indicators were discussed with the airline maintenance staff to assess the meaning of the trends. The focus of the indicators was on safety; a side effect was that monitoring of proactive indicators also enabled increased regularity. A major finding is that the use of the proactive indicator provided an additional source of information to the company in order to improve operations. An example is the use of indicators like Minimum Equipment List (MEL) provided an early warning to the organization of the health of the aircraft. This indicator confirms what it is argued that making boundaries visible may also increase system effectiveness (Rasmussen, 1997). In this case the MEL indicator is an operational boundary since it is linked to the

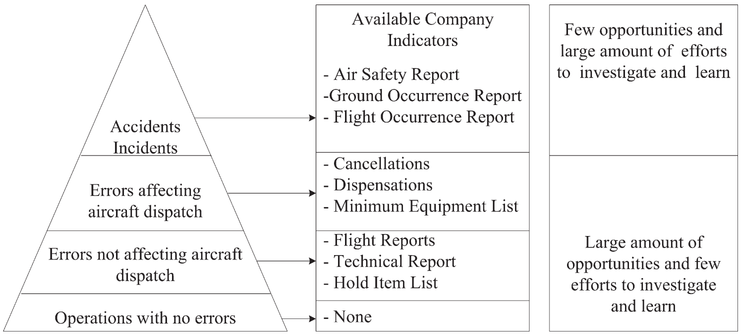


Fig. 4. Company safety indicators as navigation aid in the safety space.

airworthiness of the aircraft. Indicators related to training activities indicated downsizing in the training staff.

The statistical material covers a relatively short period 2000–2004 and it is difficult to make long-term conclusions. The identified short-term trends showed increase of utilization of the aircraft and changes in the maintenance program to follow recommendations from the manufacturers.

4.3. Questionnaire study

A major reason for performing an investigation on “how or if the major changes have affected aviation safety” was, besides resent incidents investigated by the AIBN, the statement from people working in the aviation industry; “we feel that the safety is reduced”. Personnel in leading positions from all 5 investigated operators (including the unions) were interviewed and in addition a questionnaire was sent to 9500 people, an average of 42% answered. The response was most positive among the air traffic controllers 66% and least positive among the largest group, the cabin crew, where only 36% returned the questionnaires.

There are many examples, like Challenger, Piper Alpha and Tjernobyl where organizations with excellent safety systems experiences tragic accidents. Despite these good systems the safety culture might be lacking. Safety culture is a part of a generic organizational culture and is a measure on how safety is focused both by management and employees throughout the organization by acting, saying and thinking.

Rasmussen (1997) identified five important areas that would recognize an organization with a good safety culture:

- Culture of information, i.e. gather information connected to accidents and incidents and perform an analysis of these in order to improve together with safety audits.
- A culture that promote reporting of any deviation and manage to close the loop by taking action and keeping the reporter informed.
- A no punitive culture.
- A flexible culture, able to work with changes and thereby adjust praxis without losing organizational redundancy.
- A culture of learning.

A battery of questions was prepared in order to measuring the safety culture. These questions were taken from GAIN (2001) and result could thereby be compared with similar questions made to Australian pilots. The survey indicated that the safety culture was good among planners/engineers, pilots, cabin crew and managers. The other groups scored in the better section of “moderate”. The incident-reporting practice seems to have improved over the last 5 years. Safety focus was said to have improved among colleagues. However safety focus among top managers has declined, according to all groups, except the managers. All groups of employees feel that the work relations between managers and employees have deteriorated. The managers themselves are, however, of another opinion. The majority had the impression that outsourcing and merging of companies were detrimental to aviation safety.

One important question was whether the given answers imply that safety had been jeopardized, or if they merely expressed the increased frustration due to tougher working conditions. This was put into test by asking if they thought that other groups used safety as an argument in order to gain or maintain particular advantages for their own group. They were also asked to evaluate changes in safety caused by changes in other groups working conditions, i.e. pilots were asked to consider whether the quality of air traffic control had changed. The answers indicated that tactical answers were not widespread.

The question regarding whether aviation safety had deteriorated over the past 5 years gave equal numbers in both ways for most groups while, the technicians, the air traffic controllers and the pilots thought that safety had been reduced while the managers thought it had improved. Answers from CAA personnel was in line with the other major groups.

In general, two different processes with fundamentally different consequences for aviation safety seem to have been developing in parallel. On the one hand, the organizational changes and worsened working conditions are, according to the opinion of those working within the Norwegian aviation, jeopardizing safety. On the other hand there are steady, systematic improvements in technical systems, procedures, reporting of incidents taking place, which are supposed to improve aviation safety.

The reason why there was not a vast majority stating that aviation safety was reduced, given the widespread dissatisfaction, is probably that these two developments work in opposite direction, and that that the net effect is conceived as close to zero. There is reason to believe that the period of turbulence should be as short as possible in order to maintain aviation safety.

If this survey had been repeated several improvements would have been made:

- The number of questions would have been fewer.
- The questionnaire had several questions pointing towards “the leader” without specifying if this was the top management or your leader. This excluded many questions from the statistical analysis.
- More questions would have been focused towards safety indicators.

The group that had experienced most changes where those with most response to the questionnaire. As discussed above there is no indication of tactical answers. In another setting there might have been fewer from these groups responding to the questionnaire and the answers most probably would have been different, due to the setting being different.

4.4. Identification of safety issues

Based on company documentation review, interviews, international and SINTEF in-house research regarding Safety Management Systems, the management of safety within the airlines was assessed. Major changes in the companies during the relevant period were also identified; these changes are related to organization, personnel, levels of qualification and training, and the maintenance program.

The AIBN carried out interviews with technical and administrative personnel at the five selected operators. These interviews should provide relevant information to be compared with results from other parts of the study. The complete AIBN study included approximately 4000 completed questionnaires covering all groups of personnel to make the results reliable. All results from the interviews and questionnaires were made anonymous and all data were made confidential in the reports. The results were used by the AIBN in their main report giving safety recommendations to the industry.

In order to identify whether safety in maintenance activities was affected by concurrent internal and external organizational changes, it was necessary to combine a literature review, the STEP method, safety indicators, the results from documentation reviews and the results from the interviews. In addition all information was analyzed by a multidisciplinary team.

The combination of methods provided the possibility to confirm the conclusions from different perspectives. The safety issues identified had positive and negative impacts on safety. Major changes having a safety impact were the merging of companies, the splitting of operative and technical parts of the companies, reduction in line personnel, movement of personal between companies, the mixing of different cultures in an organization and the increased tendency to subcontract maintenance work. An important issue is the merging of companies at the same time that there are changes in the training departments. This issue affected the implementation of new routines in uniform way.

The review of the management documentation demonstrated a gap in the management of safety. Safety was explicitly mentioned in the policy and less explicitly revealed through the procedures. At the blunt end we had the safety and quality departments with strong focus and commitment to safety that were responsible for the production of policies and procedures. In the sharp end, the technicians are responsible for carrying out procedures and practices. These technicians are continuously exposed to the trade-off between aircraft regularity and safety. This result complies with Hale (1998) with the result of general weaknesses in the translation of safety policy through other levels in the management of safety.

Changes in safety related organizations induced by political decisions where made without any safety analysis. When interviewing politicians the thesis was “we assume safety is under control”.

5. Discussion

The combination of techniques and the use of a multidisciplinary team allowed the problem to be reviewed from different perspectives. There was not an explicit model to evaluate the management of safety; instead a set of techniques was selected to provide qualitative and quantitative data for the analysis. The completeness of the method could be argued by the combination of techniques analyzing all maintenance activities from the setting of maintenance policy, maintenance program until the execution of maintenance.

Returning to our original questions: Do concurrent organizational changes have a direct impact on an aviation maintenance and safety and how can this be measured? This study has revealed that the statistical material does not uncover an immediate reduction in the technical standards. There is however, a possibility of long-term effects and therefore it is recommended that the companies follow-up this issue. Consideration of the production data demonstrates a reduction in the established level of aviation safety after the changes over the past few years. The AIBN identified that benefit analyses have been carried out in advance of a change. However there is no documented assessment of the safety impact.

In relation to the STEP results it was found that this method allowed link changes to different organizations and the identification of the safety consequences. In addition, this method facilitated the presentation of results to non-experts. STEP brought to light the dependencies between changes in companies and events that had happened earlier.

Statistical material and other studies do not confirm that all major organizational changes represent threats to safety. However, we have been able to identify the conditions that affect safety. This is supported by Rosness et al. (2005) who pointed out the importance of an organization being aware of warning signals, being too hasty in the process of change where the result may be a reduction of actual safety barriers and an uncontrolled reduction of safety margins.

The use of a combination of proactive and reactive indicators provided the industry with new insights towards their operations. In relation to the set of indicators used in the present project, the major constraint was related to the possibility of obtaining data that could be compared across the different companies. This constrain affects the collection and interpretation of data related to economic indicators. During the study indicators that allow the monitoring of normal operations without delay have been identified. These operations include the monitoring of near misses and adaptations to changes without interrupting operations. This issue is supported by Van der Schaaf et al. (1991) who recommended that rather than having a heavy focus on rare events like accidents, it is necessary to focus both on the detection of problems that are new to the system as well as on the detection and correction of failures.

The interviews and the questionnaires should have been more thorough in focusing on the use or misuse of safety indicators. As a consequence it is not possible to conclude about whether there has been a widespread focus on or knowledge about safety indicators as a means of determining a level of safety.

6. Conclusions

Concerning the method, it is not possible to investigate whether organizational changes affect maintenance and safety from just one perspective. The challenge was to combine both quantitative and qualitative approaches. Although it was possible to analyze trends for the indicators separately, in order to conclude about the safety trends related to airline maintenance it was necessary to look at the indicators together with qualitative information to get a picture of the changes regarding safety during the period.

We would like to point out that the method does not identify all safety impacts related to organizational changes, but it helps to identify safety issues and improves safety in maintenance activities.

The project will help the industry to pay closer attention to the management of change and its impact on safety.

One of our major conclusions regarding the indicators is that the industry uses the indicators and has the statistical data but they are not used for safety indication and they are not considered in the holistic assessment of safety. All the different indicators that we would like to call safety indicators are being used within the limitations of rules and regulations. An example is the Minimum Equipment List (MEL), the companies operated aircraft with increase in MEL and extensive use of Hold Item Lists (HIL) without seeing this as a reduction in the level of safety. Appendix A contains a list of safety indicators, both reactive and proactive. By using these indicators, together with interviews, it might be possible to improve the monitoring of safety performance.

The AIBN investigation (SL rap 35/2005) issued 15 safety recommendations. The main recommendations regarding organizational changes and aviation maintenance from AIBN (2005) were:

- Traditionally the airlines have used reactive indicators to assess safety issues and activated corrective actions. The airlines should consider putting in place proactive indicators as a predictive approach in order to identify emerging risks. Examples of these indicators are MEL, dispensations, HIL, continuation training. These indicators should receive a holistic evaluation in order to give a complete picture of the maintenance standard during simultaneous changes.
- It was advised to survey cultural differences before and after airlines merged or were associated to integrate a common corporate culture in the companies.
- “Overall follow-up and administrative routines for the supervisory authorities and the airline operators should be developed and integrated, which will include systematic and documented protection of air safety matters that are associated with the processes of change. This should form a supplement to the regulated and event-based quality systems that exist, and are mainly used, today. New recruitment/development of associated safety expertise should be considered in this connection.”
- Safety was not given enough focus during the planning and execution of the change process, neither in the companies nor by the government. It was recommended that all investigated parties should extend their knowledge of safety assessment.

Regarding maintenance and organizational changes, it is our opinion that the major challenge for the established carriers is the introduction of subcontracting many of the maintenance activities. This calls for closer attention from the regulator and the airline operators, the latter having the responsibility of the airworthiness of the aircraft. The airline operators should pay careful attention to communication lines and the establishment of systematic safety criteria for all maintenance activities and by no mean lose organizational skill and redundancy. The regulators should create and ensure mechanisms that allow the airline to demonstrate their quality control system and that the maintenance activities have been carried out in accordance with defined safety criteria. Based on the experience gained in this study, it is recommended that the regulators pay closer attention to the operators’ management of change processes when one of the conditions mentioned exist in combination with major organizational changes.

Some of the Norwegian companies (all that were subject to this investigation) have started to implement these recommendations. However, it would be interesting to put in place a maintenance model that includes safety indicators at all levels. Such a model could be used as best practice in the industry. A standardized risk model will allow airlines to identify common issues. In addition it should be pointed out the limitations and benefits of such model should be pointed out.

It was argued that there is no tradition in the airlines for a holistic risk assessment, although certain risk assessment has been carried out for specific purposes. During the project it was considered if it was possible to develop an overall model for measuring the safety level. It is recommended to develop a risk model including indicators. The suitability of indicators for this purpose has to be discussed with airlines, service companies, regulators and other relevant parties within the industry. During the project differences were discovered in the quality of the data that was collected from the companies involved in the investigation. It should be possible to measure the development of a safety level by specifying indicators. The industry should select indicators that are possible to follow up and develop methodology and guidelines to ensure that the results will be based on the same background in all companies involved.

In order to meet the competition of low price carriers, existing Norwegian airlines seek to reduce their costs. One specific action is to outsource all or some of the maintenance function. This should be done without carefully considering what to keep in-house, in order to keep control of the technical condition of the aircraft. This is related to the fact that in-house maintenance personnel during planned maintenance activities usually perform minor corrective activities at the same time, while external personnel need to get approval before performing this type of activity since this has to be paid for by the customer. In-house maintenance personnel are very often multi-skilled personnel with knowledge of the different aircraft systems. By outsourcing the maintenance activities partially, even to highly competent specialists, there will always be a possibility of losing this overall knowledge of all the aircraft systems. In a worst case the technical condition of the aircraft can be degraded, without this have been planned by any of the involved parties. When out-sourcing maintenance the company should perform risk analysis to consider what parts of the function and competence are necessary to keep in their own organization to ensure that safety levels will remain adequate.

Acknowledgements

The authors would like to acknowledge the contribution from the Norwegian and Swedish Civil Aviation Authorities to develop the safety related indicators.

Finally, we would like to thank all technical and administrative personnel from the regulator, airlines, helicopter operators, maintenance organizations for their constructive comments and openness in the best interest to improve aviation safety.

Appendix A. Detailed list of indicators

The list presented is based on Herrera and Tinmannsvik (2006). Below, it is presented a set of reactive (R) and proactive (P) indicators. The indicators are divided into (1) very important to monitor safety trend and (2) average importance. An additional indicator is proposed under the term near misses. The definition of near miss is “a deviation which has clearly potential negative consequences”. It is recommended to use this indicator within the organization to learn on how the organization recovers from a deviation. A warning is provided in relation to the use of indicator which is related to the combination of quantitative statistical result and the qualitative analysis to provide insight in the operation.

Very important to monitor safety trends:

- (R) Accident rate: Number of accidents per 100,000 flight hours (FH).
- (R) Deviations rate: Number of reported deviations, disturbances per year. This indicator could say something about improvements related to reporting culture.
- (R) Number of near misses. This indicator needs to be interpreted in relation to the organization capability to recover to events. In addition, this indicator may say something about the reporting culture.
- (P) Number of internal and external audits per year.
- (P) Number of continuation courses or recurrent training per technician per year. This indicator need to be interpreted in relation to aviation requirements.
- (P) Back-log (Hold Item List) per aircraft type per 100,000 FH This indicator should be analyzed together with the amount of dispensations requested per year.
- (P) Minimum Equipment List (MEL) reports per aircraft type per 100,000 FH. This indicator should be analyzed together with the amount of dispensations requested per year.

Average importance

- (R) Serious incidents rate: Number of serious incidents per 100,000 (FH).
- (R) Loss time injury frequency rate (LTI rate): Number of injuries per 1 million working hours. It is recommended to divide per group of employees, Line Maintenance, Heavy Maintenance, Planning and Engineering, Logistics.
- (R) Sick leave (%): Number of days off (due to illness) per year in relation to total number of working days × 100%.
- Number of deviations identified during audits per year. This indicator should be careful interpreted, it could say something about:
  (a) Organization safety level.
  (b) Audit quality and effectively.
- (P) Number of dispensations requested to the authorities per year.
- (P) Number of certified personnel per type of certificate per station per year.
- (P) Part of maintenance program that is based on in service experience, internal company requirements (in addition to manufacturer recommended maintenance program). It includes collection of information regarding development of maintenance intervals.
- (P) Fleet age per aircraft type: Indicates where the airline is located in relation to the technological development and safety equipment installed.
- (P) Number of aircraft types: Aircraft type diversity implies more resources for equipment, spares, training and competence.
- (P) Number of implemented safety measures per year: Indicates the organization commitment to accomplish safety recommendations. This indicator should be careful evaluated in relation to the content of the safety recommendations.

References

AIBN, Accident Investigation Board/Norway, 2005. Safety in Norwegian Aviation during the Process of Change, Lillestrøm, Norway.

Baranzini, D., Cromie, S., 2002. Team Systems in Aviation Maintenance: Interaction and Co-ordination Across Work Teams. Department of Psychology-TCD Dublin, Ireland.

Dahlgren, K., Lederman, L., Palomo, J., Szikszai, T., 2001. Safety performance indicators. In: Proceedings of the International Conference on Topical Issues in Nuclear Safety, Vienna, Austria.

Global Aviation Information Network (GAIN), 2001. Operator's Flight Safety Handbook (Accessed as http://www.204.108.6.79/products/products_Aviation-Operator.cfm).

Hale, A.R., Heming, B.H.M., Smit, K., Rodenburg, F.G., Leeuwen, N.D., 1998. Evaluating safety in the management of maintenance activities in the chemical process industry, Netherlands. Safety Science 28 (1), 21–24.

Hendrick, K., Benner, L., 1987. Investigating Accidents with STEP. Marcel Decker, New Jersey.

Herrera, I.A., Tinmannsvik, R.K., 2006. Key elements to avoid drifting out the safety space. In: Proceedings of the Second Resilience Engineering Symposium, Antibes-Juan-les-Pins, France, pp. 141–148.

Hobbs, A., 2000. Maintenance 'error', Lessons from the ATSB Survey. Flight Safety, Australia, pp. 36–37.

Hobbs, A., 2004. Latent Failures in the Hangar: Uncovering Organizational Deficiencies in Maintenance Operations. SJSU/NASA-Ames Research Center, USA.

Hollnagel, E., 2006. The myth of human error in risk analysis and safety management. In: Risko og sikkerhet i transport sector (RISIT) Conference, Lillehammer, Norway.

Hollnagel, E., Woods, D., 2006. Resilience engineering precepts. In: Hollnagell, E., Woods, D., Leveson, N. (Eds.), Resilience Engineering Concepts and Precepts. Ashgate, pp. 347–358.

Holloway, C.M., Johnson, C.W., 2003. Distribution of Causes in Selected U.S. Aviation Accident Reports Between 1996 and 2003, USA.

International Civil Aviation Organization (ICAO), 2006. Worldwide and regional trends in aviation safety. Working Paper DGCA/06-WP/2. Montreal.

Johnson, C.W., 2003. A Handbook of Accident and Incident Reporting. Glasgow University Press, Glasgow.

Marx, D.A., 1998. Learning from our Mistakes: A Review of Maintenance Error and Investigation and Analysis Systems. Federal Aviation Administration, Galaxy Corporation, USA.

Organization for Economic Co-operation and Development (OECD), 2003. Guidance on Safety Performance Indicators. Paris, France.

Rasmussen, J., 1997. Risk management in a dynamic society: a modelling problem. Safety Science 27, 183–213.

Reason, J., 1991. Too little and too late: a complementary on accident and incident reporting systems. In: Van der Schaaf, T.W., Lucas, D.A., Hale, A.R. (Eds.), Near Miss Reporting as Safety Tool. Butterworth-Heinemann Ltd., Oxford, England, pp. 9–26.

Reason, J., Hobbs, A., 2002. Managing Maintenance Error. Ashgate, Aldershot, USA.

Rosness, R., Forseth, U., Herrera, I., Jersin, I., Johnsen, S.O., Tinmannsvik, R.K., Tveiten, C.K., 2005. Flysikkerhet under omstillingsprosesser. SINTEF, Trondheim, Norway.

Sachon, M., Paté-Cornell, E., 2000. Delays and safety in airline maintenance. Reliability Engineering and System Safety 67, 301–309.

Sklet, S., 2004. Comparison of some selected methods for accident investigation. Journal of Hazardous Materials 111, 29–37.

Tinmannsvik, R., 2005. Ytelsesindikatorer for flysikkerhet—noen resultater fra svensk luftfart. SINTEF, Trondheim, Norway.

Tveiten, C.T., Forseth, U., Rosness, R., 2006. Proceedings of the European Safety and Reliability Conference, Estoril, Portugal, pp. 451–455.

U.S. Department of Energy (DOE), 2005. Implementation Guide Aviation Program Performance Indicators (Metrics). DOR G 440.2B-1A. Washington, DC.

Van der Geest, P.J., Piers, M.A., de Jong, H.H., Finger, M., Slater, D.H., Van Es, G.W.H., Van der Nat, G.J., 2003. Aviation Safety Management in Switzerland. Recovering from the Myth of Perfection. National Aerospace Laboratory (NLR), Netherlands.

Van der Schaaf, T.W., Lucas, D.A., Hale, A.R., 1991. Near Miss Reporting as a Safety Tool. Butterworth-Heinemann Ltd., Oxford, England.

Paper VI

Key elements to avoid drifting out of the safety space

Proceedings of the second Resilience Engineering symposium,
held in Antibes-Juan les Pins, 8-10 November 2006

Key elements to avoid drifting out of the safety space

Ivonne Andrade Herrera[1] and Ranveig Kviseth Tinmannsvik[2]

[1] The Norwegian University of Science and Technology (NTNU), Department of Production and Quality Engineering, NO-7491 Trondheim, Norway
Ivonne.A.Herrera@ntnu.no
[2] SINTEF Technology and Society, Department of Safety and Reliability, NO-7465 Trondheim, Norway
Ranveig.K.Tinmannsvik@sintef.no

Abstract. The paper is based on the experience from the development and application of safety performance indicators made in cooperation with the Norwegian and Swedish Civil Aviation Authorities. The following, two main categories of safety performance indicators are discussed: outcome-based indicators (reactive indicators; measuring the outcome/ result after a loss has happened) and activity-based indicators (proactive indicators; measuring efforts to prevent accidents). The paper presents a summary of the application of safety performance indicators to assess the management of safety in aviation maintenance. Lessons learned regarding the utilization of these safety performance indicators are presented. The paper looks critically into how the indicators are used and the conclusions that may be achieved. The paper will look into safety performance indicators and how they can contribute as indicators for resilience.

1 INTRODUCTION

The Accident Investigation Board/Norway (AIBN) has presented a study regarding the relation between concurrent organizational changes and safety (AIBN, 2005). The paper discusses the following question: How do we identify when we are drifting out of the safety space? The main objective is to discuss how performance indicators can contribute to control safety and resilience.

Recent development of the aviation industry regarding deregulation, cost reduction and increase of low cost carriers, demands the industry to be more effective and save costs. Another aspect is the tendency to subcontract activities, and the airlines may face the challenge of having to take decisions based on fragmented information. Even if the aviation industry is very safe, there is a general concern regarding cost reduction and safety. Aviation safety records show a stable accident rate while there is a concurrent increase in the number of passengers. The forecast accident rate worldwide for commercial aviation is one aircraft accident per week by 2010. In this context, maintenance errors are estimated to contribute 12% to major airline aircraft accidents and 50% to engine-related flight delays (Patankar, 2004).

The AIBN study recommended to develop more risk based supervision sustained by personnel with the relevant expertise (AIBN, 2005). In aviation, risk analyses are performed mainly by the manufacturer during the development phase of the aircraft, and this is the basis for the initial aircraft maintenance program. The initial maintenance program is delivered to the operators who are responsible for the development and update of this program in close cooperation with the manufacturer. Even if experience from the operators is taken into account in the definition phase, aviation still suffers from little research regarding risk informed operations.

As a consequence of a continuous conflict between safety and production, resilience has emerged as a new field to develop tools that include human and organizations factors to manage risk proactively. In the context of resilience engineering, it has been pointed out a need to develop resilience indicators. The paper is based on the experience from the development and application of safety performance indicators made in cooperation with the Norwegian and Swedish Civil Aviation Authorities. Lessons learned regarding the utilization of these safety performance indicators is presented. The paper looks critically into how the indicators were used, on the conclusions that can be achieved and it ends with a discussion regarding safety performance indicators and their contribution to risk informed organizations and resilience.

2 MONITORING RESILIENCE

Resilience is defined as the ability of an organization (system) to keep, or recover quickly to, a stable state, allowing it to continue operations during and after a major mishap or in presence of continuous significant stresses (Wreathall, 2006). Another definition of resilience is the ability of the systems to prevent or adapt to changing conditions in order to maintain a system property (Leveson et al, 2006). Application of resilience definition to aviation: “The capacity of the airline to continue safe operations in the face of unexpected threats or hazards including the occurrence of human errors and violations” (Wood and Dannatt, 2006). Resilience definitions include looking into the past, looking into the present (learning about normal operations) and be mindful to be able to cope with the next hazard.

Characteristic of resilience relates to how the organization acts in relation to safety-production goal conflict. Monitoring resilience involves properties such as kinds of disruptions that the system can adapt without a breakdown, systems abilities to restrict itself to response to external changes or conditions or pressures, how closely the system operates to the performance boundary, and how the system behaves near such a boundary (Woods, 2006).

Aviation maintenance could illustrate these characteristics; we could have an aircraft on ground due to technical problems, then the maintenance organization expertise from different areas work together to solve this problem, together taking the appropriate decisions to return the aircraft into operation. Due to organizational changes, the maintenance organization sometimes comprises various subcontractors; in this case the decisions involve several actors in the decision-making process, which complicates the dynamics and affects the organizations ability to maintain a normal situation. So there is a risk that the decision making process is based on a fragmented picture.

The monitoring of resilience requires both reactive and proactive parameters (indicators), which help the decision makers to detect and monitor changes in an organization which experience continuous pressure for production and safety. These indicators should describe e.g. how the organization deals with safety/production conflict, its management commitment, reporting culture, learning culture, preparedness/anticipation and flexibility.

There are tools already developed that could be applied or are applied to assess resilience. The Accident Risk Assessment Methodology for Industries in the context of Seveso II (ARAMIS) project developed a method to audit the Safety Management System and a questionnaire to measure the safety culture of an organization. It has been expressed that these two subjects are the main contribution of ARAMIS to resilience (Hale, 2006). Reason & Hobbs (2003) developed a check list for Transport Canada based on Check List for Assessing Institutional Resilience (CAIR) to assess safety culture, and the Australian Safety Bureau performed an assessment of resilience in 12 airlines, and recommendations are provided to improve the assessment of institutional resilience. The audits and check lists provide “snapshots” of the status of the organization, while the use of indicators could provide monitoring of changes and trends in the organization. Thus, these two approaches complement each other.

To address the fact that there are changes in risk with time, the Organizational Risk Influence Model (ORIM) presents a framework for the establishment of risk indicators including a risk control tool that measure the risk level of an offshore installation (Øien, 2001). This tool covers the technical, operational and organizational factors important to risk.

3 INDICATORS AS A NAVIGATION AID FOR FLIGHT SAFETY

3.1 Development of safety performance indicators

The development of performance based indicators for flight safety was done as part of the AIBN study in Norway (AIBN, 2005). The main focus of the work was flight safety, i.e. safety for passengers. SHE (Safety, Health and Environment) conditions for the employees in aviation were considered as relevant only if they were supposed to have a direct influence on passenger safety. Neither was *security* problems included in the study.

Safety performance indicators are usually established in order to monitor changes in factors influencing safety over a specific period of time. Another use of performance indicators is to estimate changes in risk level. The present study had, however, no ambitions for the latter application. Kjellén (2000) presents an overview of different SHE performance indicators, based on a framework for accident analysis:

1. *Loss*-based SHE performance indicators (e.g. the lost-time injury frequency rate, LTI-rate)

2. *Process*-based SHE performance indicators (e.g. the number of near accidents per year)
3. *Causal factor*-based SHE performance indicators (e.g. indicators based on information about contributing factors and root causes; similar to questions in safety audits).

It may be difficult to distinguish between indicators of category 1 and 2, as well as between category 2 and 3. An alternative categorization is therefore between (a) *outcome*-based indicators and (b) *activity*-based indicators:

In the following, two main categories of safety indicators have been discussed:

a) *Outcome*-based indicators (***reactive indicators***; measuring the outcome/ result <u>after</u> a loss has happened)
b) *Activity*-based indicators (***proactive indicators***; measuring efforts to <u>prevent</u> accidents).

Outcome-based indicators measure the frequency of injuries/near accidents (injury frequency rate, FAR – fatal accident rate); while activity-based indicators measure efforts to reduce injuries/losses (e.g. backlog in implementing safety measures, frequency of emergency response drills). In the AIBN study, 43 performance indicators for flight safety were put forward; 5 outcome-based and 38 activity-based indicators, respectively. The activity-based indicators were defined within the following main groups: (1) external audits (by authorities); (2) internal audits (company level); (3) emergency; (4) competence, training and experience; (5) work load; (6) maintenance; and (7) economy/ investments (Tinmannsvik, 2005). Examples of safety performance indicators for maintenance operations are shown in chapter 3.2.

The full list of 43 performance indicators were too much to handle in the project, therefore there was a need to distinguish between indicators that were supposed to be (1) very important in monitoring trends in flight safety (*dark grey colored*), (2) of average importance (*light grey colored*) and (3 no color) of minor importance for flight safety monitoring (*no color*).

The development of indicators, as well as the splitting in three groups according to their expected importance for flight safety monitoring, were based on safety audit checklists and discussions with experienced people in the civil aviation authorities in Norway, as well as in Sweden.

3.2 Indicators in practice – a maintenance case

A combination of performance indicators (proactive and reactive) was applied (AIBN, 2005) to assess the management of safety in five maintenance organizations. Table 1 gives a subset of the reactive indicators (R) and the proactive (A) indicators that are very important and on average importance in monitoring maintenance trends related to flight safety.

Table 1. Indicators for monitoring trends in maintenance organizations

Nr	Indicator	Comment
Reactive indicators (R)		
R1	Accident rate: Number of accidents per 100.000 Flight Hours (FH)	This is in accordance to ICAO Annex 13 accident definition
R2	Serious incidents rate: Number of serious incidents per 100.000 (FH)	This is in accordance to ICAO Annex 13 accident definition
R3	Deviations rate: Number of reported deviations, disturbances per year	This indicator should be careful interpreted, it could say something about improvements related to reporting culture
R4	Loss time injury frequency rate (LTI-rate): Number of injuries per 1 million working hours	It is recommended to divide per group of employees, Line Maintenance, Heavy Maintenance, Planning, Engineering, Logistics etc
R5	Sick leave (%): Number of days off (due to illness) per year in relation to total number of working days * 100%	
Proactive indicators (A)		
Internal and external audits		
A1	Number of internal and external audits per year	Different types of audits, management audits, system audits, inspections
A2	Number of deviations identified during audits per year	This indicator should be careful interpreted, it could say something about: a) organization safety level b) audit quality and effectively
A3	Number of dispensations requested to the authorities per year	
Competence, training & experience		
A4	Number of continuation or recurrent training per technician per year	This indicator need to be interpreted in relation to aviation requirements
A5	Number of certified personal per type of certificate per station per year	
Maintenance program		
A6	Part of maintenance program that is based on in service experience, internal company requirements (in addition to manufacturer recommended maintenance program)	It includes collection of information regarding development of maintenance intervals
A7	Back-log (Hold Item List) per aircraft type per 100.000 FH	This indicator should be analyzed together with the amount of dispensations requested per year
Corrective Maintenance		
A8	Minimum Equipment List (MEL) reports per aircraft type per 100.000 FH	This indicator should be analyzed together with the amount of dispensations requested per year

Nr	Indicator	Comment
Economy		
A9	Fleet age per aircraft type	Indicates where the airline is located in relation to the technological development and safety equipment installed
A10	Number of aircraft types	Aircraft type diversity implies more resources for equipment, spares, training and competence
A11	Number of implemented safety measures per year	Indicates the organization commitment to accomplish safety recommendations

Information was gathered from three airlines and two helicopter operators for a 2000-2004 period (Herrera et al, 2006). One of the problems while collecting the information is that even some definitions are standard in the aviation industry, the different operators may have different interpretations for the same term; e.g. it was not possible to gather data regarding trends in maintenance costs. It was also noticed that the operators mainly collect and analyze information that is required by the regulators. Information not required by regulators, even if available, was not used proactively in the management of safety. Indicator A3 “Number of dispensations requested to the authorities per year” illustrates this aspect. The operator requests a dispensation to the Civil Aviation Authority to continue operations when an abnormal situation occurred and the operator have to prove that the airworthiness of the aircraft is not affected to continue operation. In our case operators archived information about dispensations but did not use the information to monitor trends.

After the information was gathered and analyzed, there were discussions with maintenance personnel to verify the validity of the results. Indicators showed changes in staff, movements between companies, changes in levels of qualification and training and changes related to the maintenance program.

The indicators in our case study showed that the recurrent maintenance training has been reduced. They also showed that the constant pressure to reduce costs without affecting safety had an impact in reducing the maintenance program to the minimum acceptable level in some cases.

Conclusions from the AIBN study confirmed that the operators have systems in place to follow-up and analyze reactive indicators, but there is still a need to gather information and analyze proactive indicators.

4 DISCUSSIONS AND FURTHER WORK

Based on experience from the current study, it is evident that aviation is very strong regarding reactive indicators. To achieve robust conclusions regarding trends in safety performance indicators, data collection should be run for a long period of time; AIBN study demonstrated that 5 years is not sufficient. Conclusions concerning a potential drifting towards safety boundaries should however not be based only on performance indicators, but on a combination of quantitative and qualitative approaches.

The innovative aspect of the indicators proposed in this paper is the use of reactive, as well as proactive indicators to monitor trends related to normal operations in the aviation industry. Regarding the relationship between indicators used in the AIBN study and resilience, we could conclude that indicators such as Minimum Equipment List and Backlog have direct relation to how the organization handles the conflicts between safety and production. Indicators related to economy and implementing safety measures have direct relation to the management commitment to safety. For the future development of resilience indicators, we suggest indicators measuring the organization's ability to recover from serious deviations into stable state. Information from near misses, incidents that were overcome and ended up successfully, would be valuable data for such indicators.

A further work could include adaptation of indicators into the maintenance process, ensuring that indicators are embedded in the maintenance management system; (1) maintenance policy; (2) maintenance concept; (3) corrective maintenance; (4) preventive maintenance; (5) maintenance tasks, engineering orders; (5) planning, resource allocation; (6) scheduling work; (7) execution of work (including safety job analysis prior to perform the task); (8) inspection; (9) reporting, analysis and improvements (adapted from Hale et al., 1998). This adaptation should take into account the safety boundaries and from a resilience perspective be able to identify small changes that could affect safety.

ABBREVIATIONS

AIBN	Accident Investigation Board / Norway
ARAMIS	Accident Risk Assessment Methodology for Industries in the context of Seveso II
CAIR	Checklist for Assessing Institutional Resilience
FAR	Fatal Accident Rate
ICAO	International Civil Aviation Organization
LTI	Lost-Time Injury
ORIM	Organizational Risk Influence Model
SHE	Safety, Health and Environment

ACKNOWLEDGEMENT

The authors would like to thank the contribution from the Norwegian and Swedish Aviation Authorities to develop the safety performance indicators.
The authors acknowledge senior research scientist Per Hokstad at SINTEF and professor Jørn Vatn at NTNU for their valuable comments during the elaboration of the paper.

REFERENCES

AIBN, Accident Investigation Board / Norway (2005). Safety in Nowegian Aviation during the Process of Change. Lillestrøm, Norway

ARAMIS Project – Fifth Framework Program of the European Community, Energy, Environment and Sustainable development. Contract number: EVG1-CT-2001-0039. December 2004 (pp 47-52) http://aramis.jrc.it

Hale, A., Guldenmund, F. & Goossens, L (2006). Auditing Resilience in Risk Control and Safety Management Systems Resilience Engineering Concepts and Precepts, edited by Hollnagel, E. et al. (pp. 199-325). Ashgate, Alsershot, USA

Hale, A., Heming, B., Smit, K., Rodenburg, F. & Leuwen, N. (1998). Evaluating safety in the management of maintenance activities in the chemical process industries, Delft University of Technology, Netherlands

Herrera, I.A. (2006). Aviation safety and maintenance under major organizational changes, investigating a non existing accident. ESREL conference 2006 September 18-21, Estoril, Portugal

Kjellén, U. (2000). Prevention of Accidents Through Experience Feedback (pp. 226-261). Taylor & Francis, London, UK

Leveson N. et al. (2006). Engineering Resilience into Safety-Critical Systems, Resilience Engineering Concepts and Precepts, edited by Hollnagel, E. et al. (pp. 95-122). Ashgate, Aldershot, USA

Reason, J. & Hobbs, A. (2003). Managing Maintenance Error (pp. 159-175) Ashgate, Aldershot, USA

Tinmannsvik, R.K. (2005). Ytelsesindikatorer for flysikkerhet - noen resultater fra svensk luftfart, (Performance based indicators for flight safety - some results from Swedish aviation). SINTEF, Trondheim, Norway

Patankar, S. & Taylor, J. (2004). Risk Management and Error Reduction in Aviation Maintenance (pp. 2-25) Ashgate, Aldershot, USA

Wood, M & Dannatt, R. (2006). Assessing Institutional Resilience, A useful guide for airline safety managers? (pp. 35-45). Canberra, Australia

Woods, D. (2006). Essential Characteristics of Resilience, Resilience Engineering Concepts and Precepts, edited by Hollnagel, E. et al. (pp. 21-35). Ashgate, Aldershot, USA

Wreathall, J. (2006). Property of Resilient Organization: An Initial View, Resilience Engineering Concepts and Precepts, edited by Hollnagel, E. et al. (pp. 275-285). Ashgate, Aldershot, USA

Øien, K. (2001). Risk Control of Offshore Installations. A Framework for the Establishment of Risk Indicators, NTNU report 200104, Norway (pp.1-48)

Paper VII

Proposing safety performance indicators for helicopter offshore on the Norwegian Continental Shelf

Proceedings of the 10th International Probabilistic Safety Assessment & Management Conference, PSAM10, held in Seattle, USA, on 7-11 June 2010

10th International Probabilistic Safety Assessment & Management Conference, 7-11 June 2010, Seattle, USA.

Proposing safety performance indicators for helicopter offshore on the Norwegian Continental Shelf

Ivonne A. Herrera[a*] , **Erik Hollnagel**[b] **and Solfrid Håbrekke**[a]
[a]SINTEF Technology and Society, Safety Research, Trondheim, Norway
[b]MINES ParisTech, Sophia Antipolis, France

Abstract: Over last 10-years period there has been just one helicopter accident (with no fatalities) in the Norwegian sector of helicopter offshore operations. In this case, safety monitoring cannot be based on the absence of accidents. The main objective of this paper is to suggest a combination of leading and lagging indicators to monitor safety performance for helicopter offshore operations. An approach is described to identify indicators using different perspectives: a Risk Influence Model, the Functional Resonance Analysis Method (FRAM), and lessons learned from previous studies. The approach uses accident and incident data, as well as normal operations (when nothing goes wrong). The suggested indicators were evaluated through observations and interviews/workshop with helicopter operators, air traffic controllers, helicopter deck operators and regulators. The paper discusses the approach and proposes a set of domain specific safety performance indicators. The work was carried out under the Norwegian Helicopter Safety Study 3 (HSS-3).

Keywords: Resilience Engineering, Risk Analysis, Safety Management, Leading and Lagging Safety Indicators

---

1. INTRODUCTION

1.1. Background

Measurements of safety performance in aviation traditionally rely on lagging indicators such as accident rates, which may be further decomposed to identify particular safety issues. This categorization of accidents has enabled several improvements on specific issues. However, there is a growing concern that this information does not provide the required basis for the prevention of future accidents. The International Civil Aviation Organization (ICAO) recommended the establishment of an effective Safety Management System (SMS) [1]. Indicators are therefore needed to provide an adequate understanding of the current state of the system, and to predict possible future events or consequences of changes; i.e., leading rather than lagging indicators. Yet despite the benefit of a proactive SMS, the aviation industry generally still focuses on the reactive part of safety management.

Helicopter transport is essential for petroleum activities in the North Sea, since there is no other effective way to transport personnel. Over the last 10-years period there has been just one helicopter accident (with no fatalities) in the Norwegian sector of helicopter offshore operations. In this case, monitoring of safety cannot be based on the absence of accidents. This paper presents the results of work carried out under the Helicopter Safety Study -3 (HSS-3) [2], which had the overall objectives to contribute to improve safety and to set a reference standard for methodologies to analyse risk of offshore helicopter transportation. The HSS-3 project was a follow-up of previous studies: HSS-1 (period 1966-1990) [3] and HSS-2 (period 1990-1998) [4]. For the development of indicators, an important mandate for HSS-3 is to use experience from previous helicopter studies [3, 4, 5]. To complement this approach, HSS-3 incorporates development within safety thinking using a resilience engineering perspective to identify safety indicators.

---

[*] ivonne.a.herrera@sintef.no

1.2. Purpose of the paper

The main objective of this paper is to suggest a combination of leading and lagging indicators to monitor safety performance for helicopter offshore operations. An approach is described that identifies indicators, using different perspectives: (1) the Functional Resonance Analysis Method (FRAM), (2) a Risk Influence Model, and (3) lessons learned from previous studies. Data from normal operations (when nothing goes wrong), were used together with accident and incident data. The paper discusses the approach and proposes a set of domain specific safety performance indicators.

1.3. Delimitations

The main focus is the indicators within aviation safety in relation to major accidents, hence excludes occupational accidents. The FRAM method was used for the identification of indicators in relation to a specific scenario landing on helicopter deck. Several publications have described the use of risk influence models [2, 3, 4]. The paper emphasized indicators identified through monitoring normal operations.

2. APPROACH

2.1. Combining perspectives to identify indicators

Different perspectives were used to identify safety indicators as illustrated in Figure 1. The literature survey enabled a theoretical understanding of safety indicators, identification of relevant criteria and indicators from other studies. Resilience Engineering represents an alternative perspective on safety that takes into account successes and failures [6]. Resilience is defined as the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions. Resilience Engineering aims to develop theories, tools and methods to support resilient organizations. The FRAM method is based on resilience engineering principles and was used to identify leading indicators. The HSS-3 RIF model is an update of previous HSS-2 model. The RIF model was explored to identify lagging indicators. In combination with a literature survey, this provided candidates for indicators that were assessed in close consultation with the industry using indicators criteria, leading to a final set of leading and lagging indicators.

2.2. Data gathering

The identification of indicators was based on an iterative process:

- An initial set of indicators was identified based on literature review, application of RIF and FRAM method.
- A workshop assessed indicators against indicator criteria.
- Interviews with operational staff (pilots, engineers, training, helicopter deck, air traffic controllers, petroleum representatives and regulator) assessed indicators against indicator criteria.
- Observations of helicopter landing on helicopter deck during simulator session helped to improve modeling and improved analyst understanding of the context of operations.

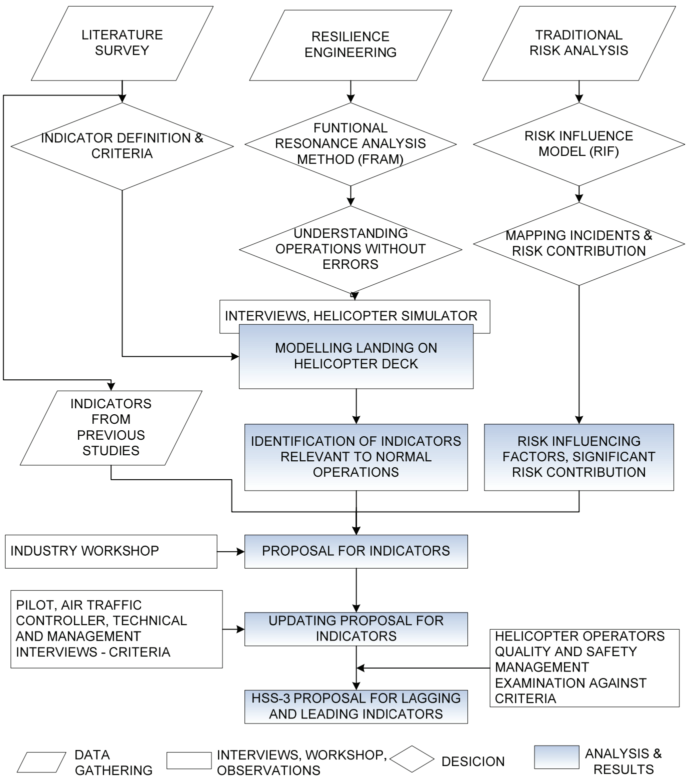
2.3. Important lessons from literature survey

Baseline for HSS-3 was the recommendations in the public report “Helicopter Safety on the on the Norwegian Continental Shelf, Part 2: Trends, objectives, risk influencing factors and recommended measures” [5]. The indicators that can be used to monitor risk were:

- number of deaths per million flight hours;
- number of accidents per million flight hours;

- number of deaths per year due to helicopter transport;
- number of serious accidents and incidents per year or million flight hours;
- number of occurrences per year or million flight hours,
- number of technical and operational reports per year or per million flight hours; and
- subjective risk (questionnaire)

Figure 1: HSS-3 overall approach to propose lagging and leading indicators



Over the last 10-years period there has been just one helicopter accident (with no fatalities) in the Norwegian sector. In addition, changes in regulations contribute to a reclassification of incidents and an increased of number of reports. An increased number of reports do not necessarily provide an indication of poor safety performance. Fatality rate and increased number of reports are therefore not suitable as sole indicators for safety performance. To complement this view, it is necessary to look for accident precursors to assess safety performance. In general, leading indicators are defined as conditions, events or measures that can be used to predict the future occurrence of an event, e.g., as accident precursors. The literature shows that there is no consistency between the definition of indicators and their application [9]. Special attention should therefore be given to the definition of indicators each time they are addressed.

Based on literature review, discussion in international forums and author's experience the following definitions are used:

- Lagging indicators measure results after unwanted events.
- Leading indicators refer to current system status and their interpretation may be used to say something about future performance

The literature presents an extensive list of characteristics for indicators. There is a need for a realistic approach for the selection. The following characteristics are adopted:

- Meaningful: the value can be correlated to accident frequency or consequence, a RIF for the risk model on accidents, or with FRAM functions for the risk model for normal operations.
- Available or affordable: it is possible to gather data with a reasonable cost.
- Reliable: The data should as far as possible be either objective or intersubjectively verifiable.
- Operational: It is possible to use the indicator to identify specific improvement measures in an operational context.
- Ownership: The indicators are “owned” by the personnel which performance is measured.

The Accident Investigation Board/Norway (AIBN) presents a study regarding the relation between concurrent organizational changes and safety [11]. In this study 5 outcome-based and 38 activity-based performance indicators for flight safety were proposed [11, 12]. The development of indicators and determination of importance for flight safety are based on safety audit checklists and discussions with experienced people from the Norwegian and Swedish civil aviation authorities. These indicators are considered as candidates for HSS-3 recommended indicators. Another significant finding in the study showed is that there is a strong focus on learning from rare accidents and failures. There is no tradition to analyze successes (normal operations with no delays) [13]. This trend has changed for flight operations and air traffic management with the introduction of Line Operations Safety Audit (LOSA) and Normal Operations Safety Survey (NOSS) respectively. These safety management tools are mainly based on managing errors and threats.

2.4. Functional Resonance Analysis Method (FRAM)

Resilience Engineering provides a practical basis for the development of systemic models in order to describe the characteristic performance of a system as a whole. It can therefore also be used as the starting point for developing a systemic or functional risk model (FRM). The purpose of a systemic model is to describe the dynamic and non-linear nature of what happens within a system. This should be seen as a complement to the traditional view where accidents are described either as sequences or as concatenation of latent conditions. Hollnagel presents a new method to perform accident investigation and safety assessment, called the Functional Resonance Analysis Method (FRAM) [7].

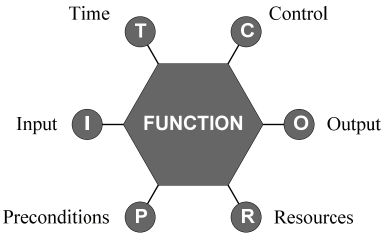
Resilience engineering sees success is as a consequence of the ability of groups, individuals, and organizations to anticipate the changing shape of risk before damage occurs; failure is simply the temporary or permanent absence of that. Adopting this view means that there is a need for models that can represent the variability of normal performance and methods that can use this both to provide more comprehensive explanations of accidents and to identify the possible risks. The helicopter safety study adopts this view to identify leading indicators.

In its present form, FRAM comprises the following five steps [8]:

Define the purpose of the analysis, since FRAM can be used for both accident investigation and safety assessment.

Identify and describe system functions. The result of the second step is the model. Every function can be characterized by six basic aspects: Input (I, that which the function uses or transforms), Output (O, that which the function produces), Preconditions (P, conditions that must be fulfilled to perform a function), resources (R, that which the function needs or consumes), Time (T, that which affects time availability), and Control (C, that which supervises or adjusts the function). A FRAM function is shown in Figure 2.

Figure 2: The six aspects of a FRAM function

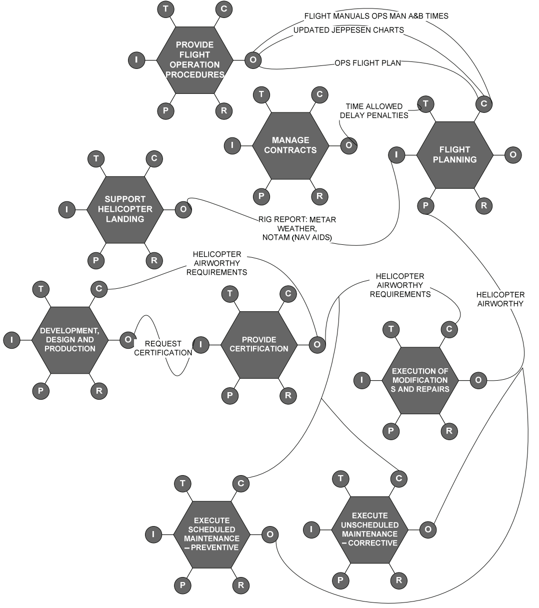


Assess and evaluate the potential variability of each function. This evaluation should be integrated with the retrospective information extracted from accident databases to the extent that data are available.

Identify functional resonance by means of instantiations. An instantiation illustrates aspects and the potential links among the functions in a defined context [10]. Figure 3 shows an instantiation for approach planning. The aim of this step is to determine the possible ways in which the variability from one function could spread in the system and how it may combine with the variability of other functions. This may result in situations where the system loses its capability safely to manage variability. The propagation may be both indirect via the effects that the variability may have on the general conditions or direct via the output from a function.

Identify effective countermeasures or barriers that can be introduced in the system. In FRAM, prospective countermeasures aim at dampening performance variability in order to maintain the system in a safe state. But it is consistent with the principle of Resilience Engineering to consider also measures that can sustain or amplify functional resonance that leads to desired or improved outcomes. Besides recommendations for countermeasures or barriers, FRAM can also be used to specify recommendations for the monitoring of performance and variability, in order to be able to detect undesired variability at an early stage. Performance indicators may thus be developed for individual functions and for the couplings among functions.

Figure 3: Instantiation landing on helicopter deck – flight planning (1 hour before departure)



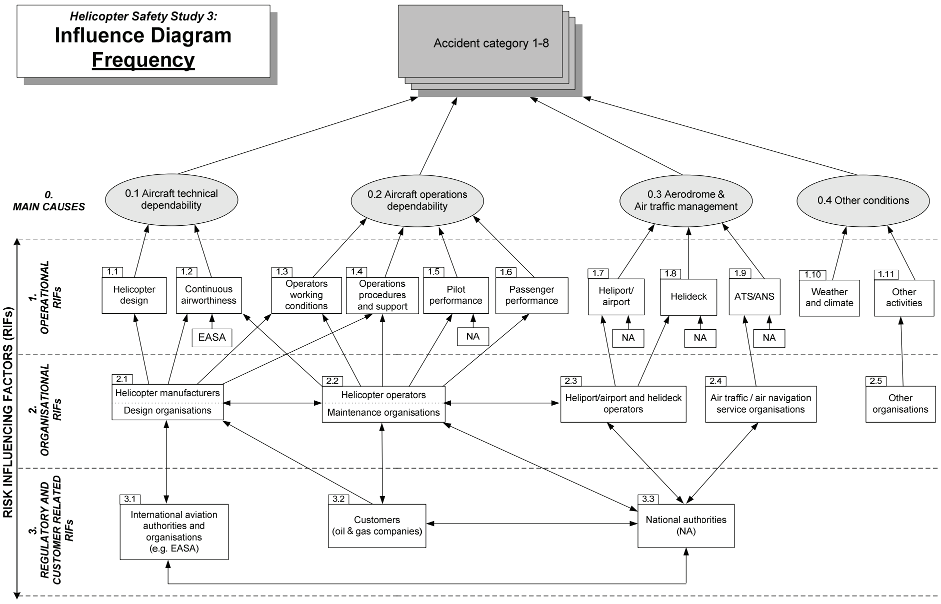
2.5. Risk Influence Modeling (RIF)

The risk influence model in the HSS-3 project is an update of the model developed in the previous helicopter safety study [4]. This approach assumes that accidents and incidents can be described as the result of cause-effect relations, sometimes as a single cause-effect chain but more often as a combination of multiple cause-effect chains.

The risk influence model is based on a number of Risk Influencing Factors (RIFs) arranged in influence diagrams. A RIF is defined as a set of conditions that influence the risk, either positively or negatively. The RIFs are likely to have a varying degree of importance for the different categories of accidents. Eight different accident categories are defined in the model. These categories are: accident by take-off or landing on heliport, accident by take-off or landing on helicopter deck, accident following critical aircraft system failure during flight, near miss or mid-air collision with other aircraft, collision with terrain, sea or building structure, accident exposing passengers inside the helicopter, accident exposing passengers outside the helicopter and other/unknown (i.e. lightening). The status of a RIF may be improved by specific actions or become worse due to changes and threats.

The RIFs are split into two categories; risk frequency influencing factors (as shown in Figure 4) and risk consequence influencing factors, and are organized in three levels. Operational RIFs (Level 1) are risk influencing factors related to activities directly influencing the risk and that are necessary to provide safe helicopter operations on a day-to-day basis. These activities include conditions related to technical dependability, operational dependability, provision of necessary external services and surroundings. Organisational RIFs (Level 2) are defined as risk influencing factors related to the organizational basis, support and control of running activities in the helicopter transport. These factors are related to helicopter manufacturers or design organizations, helicopter operators, maintenance organizations, air traffic and navigation services, heliport and helicopter deck operators. Regulatory and customer related RIFs (Level 3) are defined as risk influencing factors related to requirements and controlling activities from international organizations, authorities and customers.

Figure 4: Risk Influence Model HSS-3 for the frequency of accidents



3. MAIN RESULTS

3.1. Leading indicators identified applying FRAM

The scenario described in this paper is an approach to and landing on a floating platform during night with good visibility and no unusual events. Results from the application of FRAM referred to the five steps described in section 2.2. In the first step FRAM was used as safety assessment looking into normal operations to identify relevant indicators. In the second step a corpus of 21 functions was identified as relevant for the scenario landing on helicopter deck, e.g., Table 1.

Table 1: Example of FRAM functions for scenario landing on helicopter deck

Landing on helicopter deck FRAM functions	
Manage contracts	Manage competence
Perform weight & balance calculations	Manage procedures
Approach planning	Fix approach on GPS
Do pre-landing preparations	Arrive to minimum descend
Approach near by obstruction	Establish visual
Decide approach type (see Table 2)	Verify position
Land	Support helicopter landing

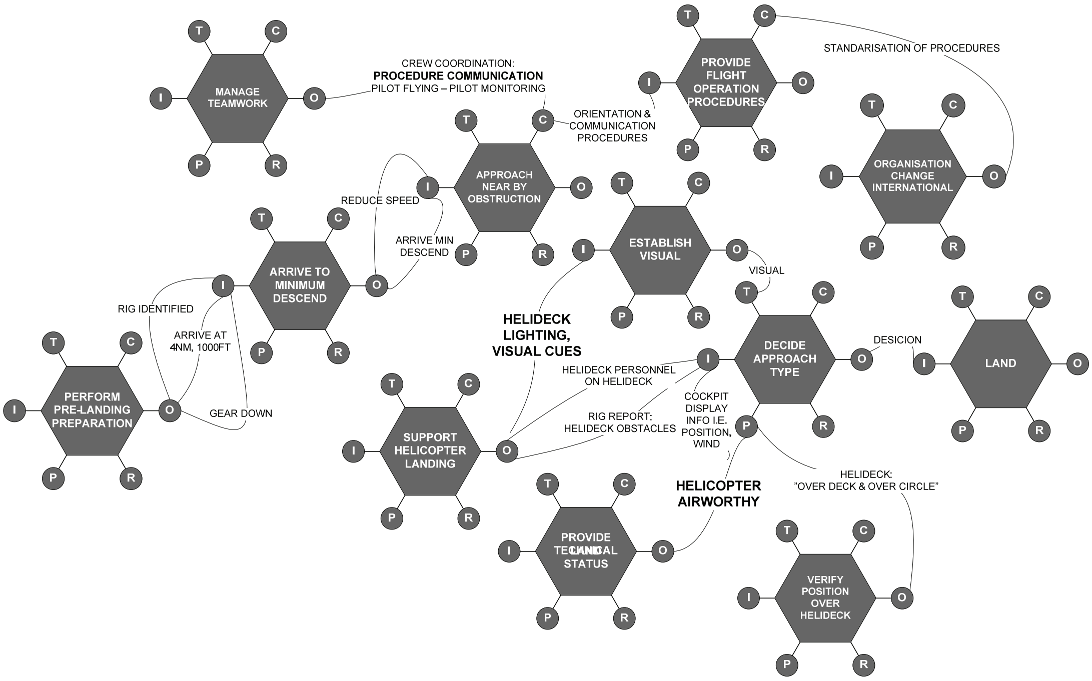
Each function was characterized in terms of six aspects (an example is shown on table 2). The granularity of the description of functions was based on iterative assessment of the scenario and the functions between the analyst and pilot.

Table 2: FRAM function characterization

FRAM Function	**Decide Approach type**
Input	Helicopter deck personnel on helicopter deck
Input	Helicopter deck obstacles report
Input	Cockpit display information: wind and position
Output	Decision
Precondition	Helicopter airworthy
Precondition	Company message “ Over deck & over circle”
Time	Visual
Control	Approach procedures

The third step was the assessment of the potential variability for each singular function. Landing on helicopter deck was analyzed in relation to landing on fixed and floating oil and gas installation during day and night. In this way, it was possible to determine variability related to normal operations. The fourth step is the determination of the ways in which variability is spread through the system. Instantiations were used to illustrate the combinations of variability. Then indicators were determined based on significant combinations of variability (bold letters illustrated on Figure 5).

Figure 5: FRAM instantiation for landing on helicopter deck during night



It was necessary to have operational indicators (shown in Table 3). This was achieved through discussions with operational and management personnel.

Table 3: Example of indicators identified for normal operations

Indicator	**Operationalization**
Helicopter airworthy	Indicators related to maintenance performance, Vibration Health Monitoring and Minimum Equipment List
Quality of communication helicopter crew and helicopter deck personnel	There are individual differences in relation to installation type i.e. fixed or floating. Use of observations to provide a qualitative evaluation.
Procedures quality and compliance	Use of audits and/or observations to provide an assessment of procedures revision and compliance
Manage contracts – use of penalties	Number of free days that have been negotiated, use of overtime
Visual references	Helicopter deck status in relation to regulation and recognized guidelines

3.2. Lagging indicators identified using RIF approach

The indicators are based on number of incidents and are mainly related to the operational level. Examples of identified indicators are:

- Technical RIF: Windshield cracking, chip warning, oil leakage detected by walk-around
- Operational RIF: Overload of cargo, incorrect marking or improper handling of dangerous goods, fuelling event, wrong charts in flight folder
- Helicopter deck RIF: Crane or other obstacles on rig near to helicopter deck, incorrect helicopter deck position, incorrect information of pitch/roll/heave from moving helicopter decks
- Weather and other RIF: Incorrect weather information, bird strikes

4. DISCUSSION AND CONCLUDING REMARKS

Helicopter safety is a result of something that the system does and not a passive property of the system. Safety is a dynamic characteristic and the result of interaction between several organizations. This view guides the indicators that are identified in the study. Each method represents different ways of understanding, it is necessary to be aware of their advantages and limitations. The paper demonstrates how a combination of several approaches provides a set of lagging and leading indicators. Since we are addressing a dynamic characteristic, it is recommended periodically to review the indicators to see if they are still relevant or whether new indicators should be considered.

The literature review show that the majority of indicators are selected from check lists or because they are easy to collect. This approach does not necessarily support indicators relevance towards safety. The FRAM modeling provides a more dynamic approach to helicopter operations. The use of instantiations enables to illustrate how variability spreads and which variability is significant to a successful landing. The main advantage of FRAM is that this approach considers the influence of the context on actual performance. Indicators identified using FRAM are leading, these indicators show a correlation to a successful operation. Indicators identified using RIF model are mainly lagging, these indicators are based on incidents and accidents information. The RIF model provides an static picture of the overall helicopter offshore operations. The recommended set of indicators represents a combination of quantitative and qualitative data. Experience from previous studied show that quantitative information does not provide enough information towards the quality aspect. This shortcoming is compensated emphasizing the importance of observations and use qualitative data.

This study represents a step forward from mainly learning from failures to consider also normal operations without failures. This approach has helped to identify alternative indicators. The identification of indicators using FRAM and the modeling enhanced understanding of the system. The indicator discussions with the industry helped to identify recommended measures to improve safety relevant to actual performance. The RIF approach allows a perspective of helicopter performance during the last 30 years. The FRAM approach represents a step forward in using new methods to improve aviation safety. While RIF method is widely recognized for risk assessment. The FRAM approach will require more applications to demonstrate its capability and have wider acceptance within the safety community.

Acknowledgements

The Helicopter Safety Study 3 is sponsored by the A/S Norske Shell, BP Norway, Civil Aviation Authority Norway, ConocoPhillips Norge, Eni Norge, GDF SUEZ E&P Norge AS, Marathon, Nexen Exploration Norge AS, Statoil and Total E&P Norge AS. We thank Norwegian technical, operative and administrative personnel from helicopter operators, air navigation providers, helicopter deck personnel and emergency preparedness organizations for their participation, as well as petroleum and civil aviation (Norway and UK) authorities for their constructive comments and openness in the best interest to improve helicopter safety. The authors acknowledge Eduardo Runte for his contribution in the initial phase of FRAM functional modelling. We thank the other members of the project team Lars Bodsberg, Per Hokstad, Erik Jersin and Tony Kråkenes for their valuable comments. The work presented in this paper has also benefit from discussions on indicators during the 3rd symposium on Resilience Engineering in 2008 and information from the European Helicopter Safety Team. We are grateful to John Wreathall and Teemu Reiman for sharing their knowledge and publications regarding the identification of leading safety indicators for the Nuclear Industry. Finally, we thank Andrew Hopkins for his views on safety indicators for the process industry presented during a visit to Norway (2009).

References

[1] International Civil Aviation Organisation (ICAO). *ICAO Safety Management Manual*. Second Edition - 2008 (Advance edition-unedited) Doc 9859, Montréal, Canada, (2009).

[2] I. A Herrera, S. Håbrekke, T. Kråkenes and P.R. Hokstad. *Helicopter Safety Study 3*. SINTEF report no. A14973. (Report in Norwegian / Executive Summary in English), Trondheim, Norway, (2010).

[3] O. Ingstad, R. Rosness, T. Sten, T. Ulleberg, M. Rausand and S. Lydersen. *Helicopter Safety Study. Main Report*. SINTEF Report no. STF75 A90008, Trondheim, Norway, (1990).

[4] P. Hokstad, E. Jersin, G.K. Hansen, J. Sneltvedt and T. Sten. *Helicopter Safety Study 2*. SINTEF Report no. STF38 A99423, Trondheim, Norway, (1999).

[5] Samferdselsdepartementet. *NOU 2002: 17 Helikoptersikkerheten på norsk kontinentalsokkel. Delutredning nr. 2: Utviklingstrekk, målsettinger, risikoinfluerende faktorer og prioriterte tiltak*, Ministry of Transport (Report in Norwegian / Executive Summary in English, Oslo, Norway, (2002).

[6] E. Hollnagel, D. Woods, N. Leveson. *Resilience Engineering: Concepts and precepts*. Ashgate, Aldershot, UK, (2006).

[7] E. Hollnagel. *Barriers and accident prevention.* Ashgate, Aldershot, UK, (2004).

[8] EUROCONTROL. *A White paper on Resilience Engineering for ATM.* EUROCONTROL, (2009).

[9] A. Hale et al. *Special Issue on Process Safety Indicators.* Safety Science Volume 47, Issue 4, Elsevier, doi:10/1016/j.ssci.2008.07.016, pp 459-510, (2009).

[10] I.A. Herrera, R. Woltjer. *Comparing a multi-linear (STEP) and systemic (FRAM) method for accident analysis.* Safety, Reliability and Risk Analysis: Theory, Methods and Applications. Proceedings of the European Safety and Reliability Association Annual Conference (ESREL), pp 19-26, Valencia, Spain, (2008).

[11] Accident Investigation Board / Norway (AIBN). "*Safety in Nowegian Aviation during the Process of Change"*. AIBN. Lillestrøm, Norway, (2005).

[12] R. K. Tinmannsvik. "*Ytelsesindikatorer for flysikkerhet - noen resultater fra svensk luftfart*", (Performance based indicators for flight safety - some results from Swedish aviation). SINTEF, (2005).

[13] I.A. Herrera, A.O. Nordskag, G. Myhre, K. Halvorsen. *Aviation safety and maintenance under major organizational changes, investigating non-existing accidents*. Journal of Accident Analysis & Prevention. Vol 41 pp. 1155-1163, (2009).

Paper VIII

Approaches to elaborate on the safety of offshore helicopter operations

Paper presented at the 5th International Working on Safety Conference, WOS 2010

held in Røros, Norway, on 7-10 September 2010

Approaches to elaborate on the safety of offshore helicopter operations

I.A. Herrera[a,b*] , P. Hokstad[b], U. Forseth[c] , S. Håbrekke[b] , T. Kråkenes[b]

[a] Department of Production and Quality Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[b] Safety Research, SINTEF Technology and Society, Trondheim, Norway
[c] Work Research, SINTEF Technology and Society, Trondheim, Norway

---

Abstract

Accidents during helicopter transportation represent a high risk for workers on offshore installations. This paper discusses the application, achievements and limitations of two approaches, risk influence modelling and storytelling, which are used in a joint industry project known as the Helicopter Safety Study 3 (HSS-3). This study represents a collaboration between scientists in the fields of sociology, human factors engineering and scientists in the industry. The two methods represent different scientific paradigms, that each has its own perspective for understanding safety and risk. The HSS-3 risk model is based on the use of risk influencing factors (RIFs) and accident categories. The HSS-3 risk model represents a quite simple influence modelling, and is not designed to account for interdependency between RIFs, and does not describe the influence from the operational context on risk. As a consequence, the storytelling is combined with the risk modelling approach to identify effects of organizational changes on work practices and safety. The application of these two methods leads to the following question: To what extent do the methods complement each other, and how can they be combined to provide a better understanding of helicopter operations. Shifting the level of abstraction from linearity and decomposition (i.e. RIFs) to storytelling and richer data sets triggers new ways of understanding, in addition to providing new knowledge about the “internal life” of the system. The combined use of these approaches provided a better way to identify changes in risk, safety threats and also safety improvement measures. While the identified safety measures and findings are domain specific, the combined use of these two approaches may be of interest also to other industries.

Keywords: Safety, Risk influence modelling, Storytelling, Organisational change, Aviation

---

1. Introduction

This paper discusses the application, achievements and limitations of two approaches, risk influence modelling and storytelling, applied in a joint industry project, the Helicopter Safety Study 3 (HSS-3; Herrera et al., 2010). This is a follow-up of previous helicopter studies, that have used statistical data and expert judgement to support a risk influence modelling to assess risk of offshore helicopter transportation, and identify risk reducing measures, (see HSS-1, Ingstad et al., 1990; HSS-2, Hokstad et al., 1999).

---

[*] Corresponding author: NTNU, IPK, S.P. Andersens vei 5,
NO 7491 Trondheim, Norway - Tel.: +47 90 68 06 34; fax: +47 73 59 28 96.
E-mail address: ivonne.a.herrera@ ntnu.no (I.A. Herrera)

There have not been any fatal helicopter accidents on the Norwegian Continental Shelf (NCS) during the period from 1999 to 2009, which is the main period of investigation for the HSS-3 study. At the same time, there has been a change in the event reporting, which has resulted in an increased amount of incident reports (occurrence data). Moreover, the study had to consider challenges caused by major organizational changes within the internal framework conditions[1] of the two major helicopter companies.

Since 1990, significant studies have been carried out to assess risk and identify risk reducing measures (Ingstad et al., 1990; Hokstad et al., 1999; Herrera et al., 2010; Ministry of Communications, 2001, 2002).

The industry as a whole has been committed to improving safety, and several technical and operational improvements have been implemented using the experience from these previous studies. Special requirements and guidelines for offshore helicopter operations have been developed (CAA-N, 2009; OLF, 2009; 2010]. Previous helicopter safety studies, such as HSS-1 and HSS-2, cover the periods from 1966 to 1990 and 1990 to 2008, respectively. HSS-3 covers the period from 1999 to 2019. The overall objective of HSS-3 is to contribute to improved safety in the offshore helicopter transport of personnel. It is the ambition of HSS-3 to create a reference standard methodology for the analysis of risk in helicopter transport, as well as the identification and assessment of risk reducing measures in this mode of transport.

The following main delimitations apply for results achieved in HSS-3:

- The quantification and assessment of various trends was carried out within the RIF model. Subjectivity in expert judgement and possible interaction effects between RIFs are main sources of uncertainty in the estimated contribution to risk from each RIF.
- The storytelling provided a richer data material, but due to time and resource limitations, it was only possible to present the essential elements, and no in-depth analysis was attempted.

The paper is organized as follows. Section 2 addresses paper objectives and research questions. Section 3 briefly introduces the context of the HSS-3 study, describing offshore helicopter operations on the NCS. The methods are described in section 4. For a more comprehensive description on the RIF method the reader is referred to Hokstad et al. (1999), and on storytelling to Gabriel (2000). The way that data is collected is described in section 5. Section 6 presents the analysis of data and the results. The findings are discussed on section 6. Finally, section 7 presents conclusions and implication for further research.

---

[1] Framework conditions refers to the conditions influencing practical possibilities an organization , organizational unit or individual has to control major and working environmental risk (Rosness, 2009)

2. Paper objectives and research questions

The main question addressed in this paper is to what extent a combined use of the helicopter risk influence modelling and storytelling will increase the basis for assessing risk of offshore helicopter transportation. Therefore it is an objective to discuss the application, achievements and limitations of two approaches, the HSS-3 risk influence modelling and storytelling, as used in a joint industry project known HSS-3.

To answer the main question a set of sub-questions are considered

- What are the strengths and weaknesses of two approaches?
- Which approach could be used to identify and assess safety challenges associated with organizational changes in the internal framework conditions?
- Does the study provide useful insights into required adjustments in the risk influence model, also accounting for findings from the storytelling?
- Can the two methods be combined into a single unified approach, or have storytelling and the helicopter risk influence modelling to be taken into account separately?

Note that in the HSS studies the risk to passenger on helicopter flights is quantified as the *mean number of fatalities per million person flight hours*. This is given as the product of the accident frequency (number of accidents per million person flight hours) and the consequence (number of fatalities per accident). Historical data can be applied to estimate the magnitude of this risk in the past and by combining these data with the RIF influence modelling and expert judgement (e.g. on changes in the RIFs), it is possible to estimate future risk level (i.e. mean number of fatalities per million flight hours). This particular risk measure is chosen because it corresponds precisely to the Fatal Accident Rate (FAR), being the preferred measure for accident risk used in the offshore industry.

3. Helicopter operations on the Norwegian Continental Shelf

Helicopter operations on the NCS are conducted under special weather conditions which are often harsh and demanding.

In addition to several technical and operational improvements, there have been important organizational changes in the ownership of the two major helicopter operators on the NCS in the period from 1999 to 2010:

- 1990–2000: Both helicopter operators were Norwegian owned, and had Norwegian management and administrative boards with much experience within the field of aviation safety.
- 2000–2009: Helicopter Service AS was reorganized into CHC Helicopter Service AS (the operating company) and CHC ASTEC (the maintenance organization), and then into CHC Norway AS and CHC Heli-One. Norsk Helikopter is now called Bristow Norway AS.

- From the beginning of 2010, both helicopter operators have been completely foreign owned. CHC Norway is now owned by Canadian Helicopter Corporation (CHC), which itself is owned by the venture company American First Reserve. Bristow Norway is now owned by the British company Bristow Group Inc.

Moreover, concurrent organizational changes have been experienced in Norwegian Aviation (AIBN, 2005).

4. The two approaches for elaborating on helicopter safety

It has been acknowledged that to a large extent the model that is selected for an analysis/evaluation will define what is to be considered and which evaluation criteria will be applied (Reiman and Odewald, 2009. A main premise for our work is that the HSS-3 risk model is a direct development of previous helicopter models, (HSS-1 and HSS-2), which illustrate and communicate the main influences on helicopter risk, and changes in the risk level during three consecutive time periods.

Both qualitative and quantitative evaluations of (changes in) the risk is performed using a risk influence model, based on a number of Risk Influencing Factors (RIFs) arranged in influence diagrams. A RIF is defined as a set of factors or conditions that influence risk associated with offshore helicopter transport. A RIF can be stable or vary over time as a function variable or a random variable. A RIF has a condition (status) which is either acceptable or poor/unacceptable, and which can change by introducing risk reducing measures, or as a result of other changes within the helicopter industry. The RIFs model comprises two influence diagrams: a frequency influencing diagram and a consequence influencing diagram. The influence diagram for frequency is shown in Figure 1, in which each box represents a RIF and the arrows indicate influences between RIFs or between RIFs and accident frequency/consequence. As illustrated in the figure, RIFs are organized into three levels which are defined as follows:

Level 1 – Operational RIFs: Conditions related to the necessary daily activity of achieving safe and efficient offshore helicopter transport, e.g. helicopter design and pilot performance. The operational RIFs for frequency are grouped within four main causes of accidents.

Level 2 – Organizational RIFs: Conditions related to organizations and their support and control of daily activities, e.g. helicopter operators and manufacturers.

Level 3 – Authority and customer-related RIFs: Conditions related to activities of national and international authorities and customers (i.e. oil and gas companies).

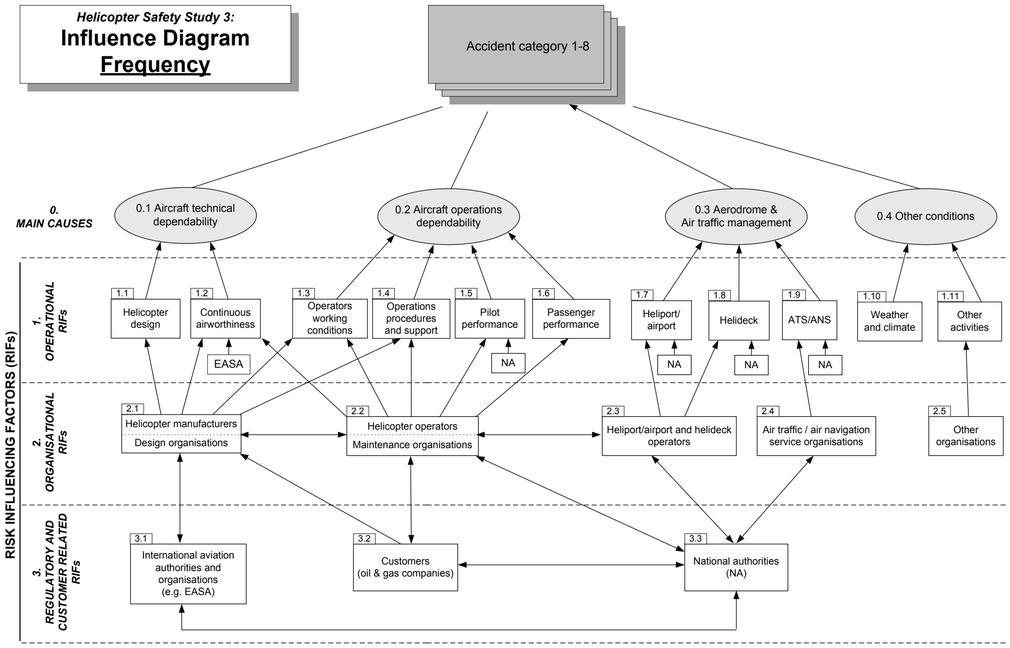


Figure 1. Influence Diagram for Accident Frequency (Herrera et al., 2010)

A similar influence diagram for consequence is organized in the same three levels, but comprises other RIFs. In addition the HSS-3 risk model splits the risk into contributions from eight different *accident categories*: an accident occurring at takeoff or landing on a helicopter deck; an accident following a critical helicopter system failure during flight and a collision with the ground, sea or a building.

The RIF model is applied in risk quantification to identify which RIFs contribute the most to risk, further to the identification of risk-based indicators, and also in a cost-benefit analysis to estimate achievable risk reduction for various measures (at various costs).

Expert judgements were used in combination with event data to perform a risk assessment, using the RIF model. Several sessions such as working meetings, telephone meetings and group interviews were carried out to assess the importance of RIFs and changes in the status of the RIFs over the time periods from 1999 to 2009 and from 2010 to 2019. The quantification had a focus on the operational RIFS (level 1). In relation to organizational changes (RIFs at level 2), four central topics were identified relevant for storytelling:

- Change of decision-making authority/management of resources and its significance for work practices;
- Changes and their significance for maintenance routines;
- Changes in competence and training;
- Changes in cooperation and communication with local management.

For each of these topics, an analysis of based on storytelling of central changes and the informants' interpretation of the situation are required in the form of examples. Then, a conclusion is presented in the form of organizational challenges.

Note that the HSS-3 RIF model is an overall, simple risk model, not designed to address the influences of specific detailed factors, and it accumulates the influence of several conditions under one risk influence factor (RIF). In particular, the effects of RIFs at level 1 (being used in the quantifications) are assumed to be additive, without interaction effects. Other risk models have evolved to more sophistication and are more complex. An advantage of the HSS-3 model, however, is that it has a simple structure, and gives an overall picture of how main factors affect the risk; being useful e.g. when the risk is discussed in an interdisciplinary group.

Thus, the current, simple RIF model is not designed for answering a question such as "What are the challenges due to organizational changes in the internal framework conditions?" So there could be introduced a RIF (at level 2) to account e.g. for such organizational changes. However, to keep the RIF model simple, an alternative approach should be introduced to address e.g. the influence of changes in the internal framework conditions, and *storytelling* was chosen to complement the helicopter risk influence model. One advantage of storytelling is that this approach facilitates the description of various interdependencies and influences of the operational context on risk. The storytelling perspective permeates a large part of organizational studies. It provides valuable insights into the nature of organizations, the power relations within them and the experiences of their members. Stories reveal how wider organizational issues are viewed, commented upon and worked on by their members. Stories are defined as "*narratives with plots and characters, generating emotion in narrator and audience, through a poetic elaboration of symbolic material. This material may be a product of fantasy or experience, including an experience from earlier narratives. Story plots entail conflicts, predicaments, trials, coincidences, and crises that call for choices, decisions, actions, interactions, whose actual outcomes are often at odds with the characters intentions and purposes* (Gabriel, 2000). Therefore, to gain access to how a situation is experienced and interpreted for the companies in question, a qualitative open storytelling approach is selected. Stories are used as a data source to achieve a broader perspective of how and what employees and managers talk about in regard to a specific topic. This data source communicates knowledge differently than short, fact-based descriptions or factors, as all the stories are written in a short quote version. The stories tell what took place, who and what was involved and what the result was. Then, an analysis is performed to evaluate the impact of the topic on practices and safety.

5.Data sources

The study represents close collaboration between scientists in the fields of sociology, human factors, engineering and experts in the industry. The HSS-3 study combined data collection from a literature review, interviews, group interviews, seminars, expert judgement sessions, discussions and some statistical observations of accidents/incidents. It included a comprehensive review of relevant literature and studies, accident and incidents reports as shown on Table 1 (a detailed list of relevant helicopter offshore literature is found Herrera et al., 2010). Historical data were used to assess the risk level in the past; experts were then used in combination with the RIF model to assess the current/future risk.

Table 1. Data sources for incidents and accident data

Sources	Information	Period	Relevance
Norwegian helicopter operators and Petroleum Safety Authorities	Traffic volume in the Norwegian sector	1999–2008	Risk quantification
Norwegian helicopter operators and Petroleum Safety Authorities	Reported incidents from the helicopter operators to the Norwegian Petroleum Safety Authorities	1999–2009	Accident categories, factors identified in reports, recommendations
Civil Aviation Authority Norway (CAA-N) and Aircraft Investigation Board, Norway (AIBN)	Reported accidents, incidents and investigation reports	1990–2009	Risk quantification, accident categories, factors identified in reports, recommendations
Oil and Gas Producers (OGP)	Annual statistics with an overview over accidents and the traffic volume in the North Sea	2000–2007	Accident categories, factors identified in reports, recommendations
Civil Aviation Authorities UK (CAA-UK) and Air Accident Investigation Branch (AAIB)	Accident data from CAA-UK Selected Aircraft Investigation reports from AAIB UK	2000–2009	Accident categories, factors identified in reports, recommendations
Accidents in the North Sea and Canada	Investigation reports, telephone interviews	1999–2009	To which degree these accidents could have occurred in Norway, accident categories, contributing factors, recommendations

Key participants in the industry were interviewed to map technical, operational and organizational development having an impact on helicopter operations offshore (e.g. pilots, engineers, managers, air traffic controllers, inspectors, national and international experts). A total of 14 individual and group interviews were effectuated to identify developments.

The identified developments formed the basis for expert judgments sessions to assess the change in risk due to these developments. Several expert judgements session were effectuated focusing on technical, operational, organizational aspects, factors affecting accident frequency or consequence. Pilots, maintenance personnel, managers, inspectors from the helicopter operators or for the regulation authorities, air traffic controllers interacted together with researchers. These sessions contributed to relate the developments to the different risk influence factors, e.g. how requirement and utilization of vibration health monitoring is related to RIF 1.2 “Continuous airworthiness”.

The first phase of the study allowed the identification of four central topics related to changes in the internal framework conditions. In relation to storytelling, changes in the internal framework conditions represented a sensitive area to address. Interdisciplinary discussions with known researchers, helicopter operators, managers, union's representatives and professors were effectuated to identify a proper way to assess data. A strategic selection of representatives from management, in addition to employee representatives from the operative and technical departments in CHC Norway, CHC Heli-One and Bristow Norway is performed. These persons were identified and contacted. They acted on behalf of the company or provided their individual opinion on the points addressed during the interviews. Interviews for storytelling were designed based on a list of specific topics (listed on section 5), as a dialogue in which the interviewees have the opportunity to offer their own stories on these central topics. A total of 9 individual interviews were performed each interview having a duration of 2 hours in average.

The quality of the data collected has been ensured by having several sources of information and a strategic combination of competences during the expert judgement sessions and interviews. Several expert judgment sessions and interviews were performed aiming to achieve some degree of convergence. The group of persons involved vary in relation to the themes explored on each session. A balance between operative, maintenance, customer and regulator was achieved. In addition, the project team was aware of the different interest represented in the groups. All sessions and interviews were documented to form the basis for the analysis and to maintain traceability.

6. Main results

The risk influence model of HSS-3 is used as a basis for discussing and quantifying risk in a structured manner. *Risk change* over a period of time is estimated, as well as *contributions to risk* from the various RIFs in the model. The HSS-3 study has identified and assessed technological, operational and organizational developments relevant to helicopter safety in the study period from 1999 to 2009, and for the next period (2010–2019). The list of developments is a result of a comprehensive review of relevant literature and studies, accident/incident reports and statistics, plus interviews with key participants within the industry (e.g. pilots, engineers, managers and ATC personnel).

For example, positive developments include the introduction of new helicopter types, improved simulator training, increased standardization and a focus on helicopter deck operations. There are, however, several developments that are also a cause for concern, i.e. the lack of a possibility to maintain additional requirements for offshore flights adapted to the conditions on the NCS.

The RIF model has together with predicted developmemts been used to assess changes in risk Accident frequency and consequence are quantified separately for the various time periods studied. Details of the quantification procedure are given in (Kråkenes and Håbrekke, 2010), and the main results are given in Table 2. All values given here are totals which cover all accident categories.

Table 2. Estimated change in risk (Kråkenes and Håbrekke, 2010)

Parameter	**1999–2009**	**2010–2019**
Frequency change	-11 %	-19 %
Consequence change	-10 %	-9 %
Risk change	**-20 %**	**-27 %**

The risk reduction is estimated at 20 % in the period from 1999 to 2009, and is estimated to be reduced by a further 27 % over the next decade.

Using statistics in combination with expert judgment, we can also assess the relative contributions to risk from the various RIFs. Knowledge of the main contributors to risk is very valuable in order to identify the areas in which risk reducing measures would be most effective. Figure 2 shows an example of results, presenting the relative contribution to risk from the 11 operational RIFs for frequency (Figure 1).

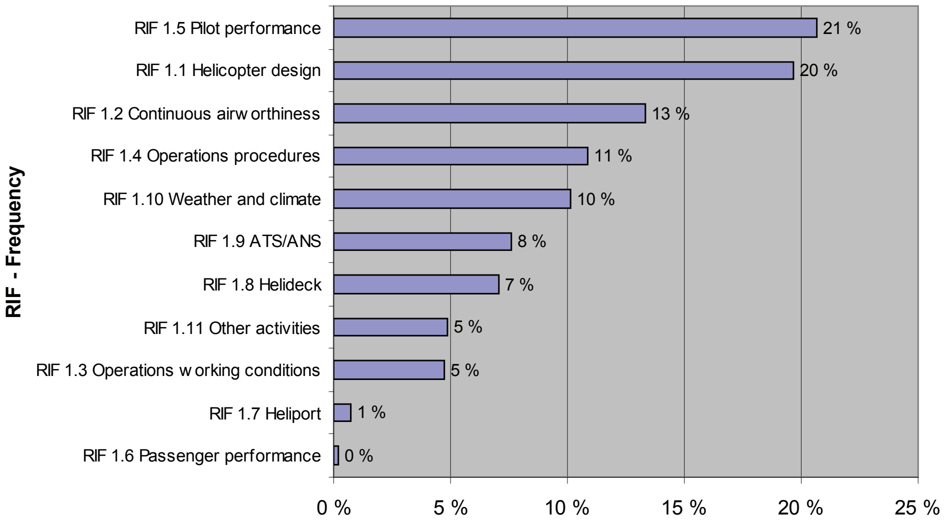


Figure 2. Contribution to risk from operational RIFs for frequency (Herrera et al., 2010)

Figure 2 reveals that the two RIFs which contribute most to the risk are RIF 1.5 *Pilot performance* and RIF 1.1 *Helicopter design, with* similar results found for the operational RIFS for consequence.

The period studied is characterized with a change of ownership from Norwegian owned to foreign owned, introduction of a new management culture and a transition from a traditional hierarchical structure to a matrix structure organization. The main focus of the storytelling was to investigate the effect of this. The interviewees were positive towards the increase in resources including the number of helicopters, and being part of a large organization in a competitive environment. Regarding the decision-making authority and management of resources, the interviews indicated the difficulty of knowing “where one stands” and “who one should relate to” in the wake of the organizational changes with the new owners. The stories documented that strategic decisions concerning investments are made abroad. Some examples are given below.

The matrix-structured organization may introduce ambiguities which need to be managed in a considerate way (Hopkins, 2009). Some stories illustrated how the decision-making structure leads to confusion and poses demands for a rapid response that may call for improvisation, whereas other stories dealt with the consequences of the centralization of tasks. The following story was told by several of the interviewees about a particular incident when they were preparing for take-off from a small helicopter base: *”Employee X had to leave to buy diesel, fill the tractor [which pulls the helicopter from the hangar] with diesel and pay with a private credit card - that’s how you get diesel on the tractor so that you can fly.”* Due to the introduction of new routines regarding purchases and the settlement of due payments from the accountant’s office abroad, we were told, an employee had to serve as a private bank for the company in order to re-fuel the engine. This story can be interpreted as an illustration of unintended consequences of the new organizational structure. It can also be seen as a critique of the new organization and the management system.

The following story is an example of how employees use humour as a disarming tool or as a counter-power strategy: “*An e-mail came from the management saying that if we could maintain over 90 per cent regularity for one week, they would buy cakes for all the bases. But then an employee answered in an e-mail saying that if the management could provide spare parts for the entire week, they would buy cakes for the entire management*.” The lack of spare parts is a large problem which leads to “cannibalism”. This means that the replacement and certification of a component needs to be performed twice. In cases of limited time, this represents stress for the organization. The problem has grown recently as a result of the “new” machine type being put into use on the Norwegian Continental Shelf, and many components had less performance than expected. Different examples were identified which support the fact that an organization can be pressured for a short period and have the ability to respond. Nevertheless, this pressure can represent a safety threat in the long run.

Training is becoming based more on information technology (IT). In some areas IT works fine, though it is not always possible to ask questions. The following comment illustrates the need for alternative means: *“If you are going through the entire helicopter with a pc, you end up “clicking” yourself forward, and you don’t even know where you are.”* Advance helicopter technology requires more training since it demands a greater understanding of the system. There is a need for an emphasis on being able to talk and reflect about the system and to relate to your and others practices and experience, especially in relation to operation conditions in the NCS. This and other stories like it highlight the need for an optimal balance between IT-based training and experience transfer in classrooms. In addition, the quality of the training should be improved, and not just the number of hours.

It has been asserted that changes in ownership have led to a more authoritarian management and control logic than people have been previously used to. In Norway, the norm is to work based on goals instead of directives, which was the explanation for this. The following quote illustrates cultural differences in management: "*In other countries when you say left, you go left. In Norway, if we say we are going left, our people would ask why? In Norway, we could rather have said: Today our goal is to go here, and the employees would achieve that by going left.*" For those who are used to "management by directives", this can seem time consuming and "slow" because the work is based on processes, delegation and guidelines. There appears to be a culture clash between two different management models: one is an authoritarian foreign model ("management by directives"), while the other is a Norwegian model in which management and employee representatives represent different interests, but cooperate.

So the HSS-3 introduces the use of storytelling to assess the impact of changes in framework conditions for helicopter operations offshore. The stories provide insights on conditions affecting the context of operations. For example, the story about lack of spare parts illustrates the effectiveness of maintenance and operative or inoperative equipment. The stories give knowledge about the interaction of different factors, how the maintenance repair activities carried out by the maintenance personnel (RIF 1.2 Continuous airworthiness) are affected by the availability of spares from helicopter manufacturers (RIF 2.1), the capacity to procure spare parts (RIF 2.2 Helicopter operators and maintenance organizations). Depending on which equipment is affected, it may also have an impact on the flight operation and relation to pilots working conditions, procedures and pilot performance (RIF 1.3, 1.4 and 1.5). Thus, the storytelling gives useful insight into important RIFs of the influence model (cf. Figure 2).

7. Discussion

The two approaches elaborate on the safety of helicopter operations by identifying changes in risk, important factors that influence risk, organizational challenges and improvement measures. The analyses identify and recommend a set of cost-effective risk reducing measures based on the combination of a quantitative and qualitative approach.

The combination of historical data, expert judgements and HSS-3 model provide an estimate on risk reduction. The modelling supports the identification of categories contributing to accident risk (e.g. controlled flight into terrain, critical system failure during flight, take-off/landing on helideck). The most important factors in relation to accident frequency or consequence contribute to propose relevant risk reducing measures for the next decade. One limitation of the HSS-3 modelling approach does not take into account the interaction effects between influencing factors and the influence of the operation's context. A critique of the quantitative risk approach is also related to the exclusion of organizational malfunctions and to the fact that the perception of undesirable effects depends on a person's values and preferences (Jaeger et al., 2001).

Storytelling can be seen to represent a view in which risk cannot be confined to separate and independent factors. This interpretive paradigm offers an intersubjective view of the world, searching for patterns of meaning represented by stories. The stories captured in the study constitute different images of an organization, are multifaceted and can be interpreted in more than one way. One important interpretation could be the resistance to carry out changes in the organization. In other cases, as shown in the example of spare parts, the utilization of humour in the stories used within the organization serves as organizational learning. It is possible to see how the stories are reflexive in the way they recreate the past in accordance with the present, providing inferences that have their own significance (Boje, 1991). The stories document relatively extensive changes in the internal framework conditions within the two helicopter operators. Moreover, it is possible to identify changes which are both positive and negative. As illustrated in the analysis, the stories have several functions in the organizations. The stories enable the sharing of knowledge between employees, and serve as sensemaking device (Weick, 1995, 2009). These stories can be seen as qualitative indicators of the status of the organization. Moreover, the stories also demonstrate a set of organisational trends that contribute to a shift of focus away from the primary, operational work tasks. For example the encountering of different management cultures, the change of the organization structure result on a demanding process. These changes have not yielded any concrete effects on safety in the form of incidents, but in a business based on alertness at every level, it is disconcerting to have unclear reporting lines, a lack of concurrence between power and authority, unrest and uncertainty in the organization. So in the long run, such conditions could potentially pose a threat to safety. The stories indicate that pilots and technicians have high standards for their performance, which has most likely contributed to everything working out well, despite the unrest within the organizations.

The stories give knowledge about a combination of framework conditions, conflicts and frustrations that are not reflected in the RIF model. We see difficulties in integrating rich organizational stories into the HSS model. Form one side, if the story is separated into several factors it losses its original meaning. On the other side, the integration of the stories into a single model would change the HSS simple structure into a more complex structure. Hence, the project study demonstrates how is it possible to have a dialogue between different approaches and use findings collectively without integrating these two approaches into a single approach.

8. Conclusions and implications

This study demonstrates the importance of using different perspectives to provide broader information about the system under evaluation. The socio-technical system addressed here is complex, dynamic and undergoing continuous change. As a consequence of this, several perspectives should be used to uncover different adaptations and trade-offs in an organization. The RIF model uses facts and factors, and is a very powerful tool for giving an overall picture of independent factors. Shifting the level of abstraction from linearity and decomposition to that of storytelling, and the use of richer material, yields new knowledge and triggers new ways of understanding the internal life of the system. The use of different perspectives reveals safety challenges that could otherwise remain unveiled, and it is demonstrated that a combined use of the helicopter risk influence modelling and storytelling will increase the basis for assessing risk of offshore helicopter transportation.

The two approaches have its strengths and weaknesses. Our study demonstrates that storytelling supplements the RIF modelling with information that encompasses more than “hard facts” and decomposition into risk factors. For example, we were able to identify the use of humour to illustrate how the organization adapts to coping with strict demands of flight regularity. Both approaches aim at informing about the status: are the situations getting better or worse? The risk model provides valuable information for risk reduction and factors that affect risk. The application of the RIF modelling confirms a risk reduction. The storytelling and the modelling confirm the significant importance of the helicopter operators, maintenance organizations and design organizations with respect to safety. Both approaches allowed the identification of conditions that represent threats to safety in the long run. One shortcoming in the modelling is the difficulty in addressing organizational changes that include complex interactions and dependability between various actors. The use of storytelling demonstrates that it is possible to inform about the status of organizational interactions and their impact in a rich format, which takes into account the context and demands of operations. Example of interactions captured by the storytelling are related to helicopter operator’s organizational changes introducing a matrix structure with ambiguities at the same time that the civil aviation authorities had a lack of personnel with expertise of offshore operations.

Storytelling is the preferred approach to identify and assess safety challenges associated with organizational changes in the internal framework conditions. The results could be related to various RIFs of the risk influence model and enable the identification of relevant risk reducing measures and interdependencies between these. Thus, some insights into required adjustments of the risk influence model were obtained by findings of the storytelling.

We see promise for the potential of a new dialogue between these perspectives to help provide new knowledge and ways to improve understanding in the organizations. However, we do not envisage merging both approaches into a single unified approach. The modelling provides a powerful overall picture and set of factors that are relevant. The storytelling may shed some light on specific topics and reveal the adaptations of the organization. One alternative not explored in this paper might be to use storytelling in the initial phase of a traditional risk analysis, thus providing new factors that need to be taken into account. A possible further development with regard to the application of storytelling is to ask the informants to tell what is important about the story and to reflect on their significance and meaning, and to compare the collected stories. Another possibility is to use indexing of the stories to identify patterns.

Finally, from the risk influencing factors, it may be possible to deduce risk indicators and use these indicators for the monitoring of risk. Most of the risk indicators identified in the RIF modelling are related to the “sharp end” and actual events (i.e. accidents). At the same time, stories can be seen as early warnings which identify a set of challenges that needs to be addressed. For this reason, stories can provide powerful insight, showing trends in an organization that may lead to unwanted consequences.

Acknowledgements

The Helicopter Safety Study 3 was sponsored by A/S Norske Shell, BP Norway, Civil Aviation Authority Norway, ConocoPhillips Norge, Eni Norge, GDF SUEZ E&P Norge AS, Marathon, Nexen Exploration Norge AS, Statoil and Total E&P Norge AS. We thank Norwegian technical, operative and administrative personnel from helicopter operators, air navigation providers, helicopter deck personnel and emergency preparedness organizations for their participation, as well as petroleum and civil aviation (Norway and UK) authorities for their constructive comments and openness in the best interest of improving helicopter safety.

References

Accident Investigation Board/Norway (AIBN), 2005. Safety in Norwegian Aviation during the Process of Change. AIBN. Lillestrøm, Norway.
Boje, D.M., 1991. The Storytelling Organization: A Study of Story Performance in an Office-supply Firm. Administrative Science Quarterly, 36: 106-126.
CAA-N, Civil Aviation Authority – Norway, 2008. BSL D 5-1, FOR 2007-10-26 nr 1181: Regulation governing continental shelf operations – commercial air traffic to and from helidecks on offshore installations and vessels, http://www.lovdata.no/cgi-wift/ldles?doc=/sf/sf/sf-20071026-1181.html,
Gabriel, Y., 2000. *Storytelling in organizations*, Oxford: Oxford University Press.
Herrera, I.A., Håbrekke, S., Kråkenes, T., Hokstad, P., Forseth, U., 2010. Helicopter Safety Study 3. Main report. SINTEF report no. A14973, Trondheim, Norway.
Hokstad, P., Jersin, E., Hansen, G.K., Sneltvedt, J., Sten, T., 1999. Helicopter Safety Study 2. Volume 1: Main Report. SINTEF report no. STF38 A99423, Trondheim, Norway.
Ingstad, O., Rosness, R., Sten, T., Ulleberg, T., Rausand, M., Lydersen, S., 1990. Helicopter Safety Study. Main Report. SINTEF Report no. STF75 A90008. Trondheim, Norway.
Hopkins, A., 2008. *Failure to Learn, the BP Texas City refinery disaster*. CHC Australia Limited.
Jaeger, C.C., Renn, O., Rosa, E.A., Webler, T., 2001. *Risk, Uncertainty and Rational Action*. Earthscan Publications Ltd. London.
Kråkenes, T., Håbrekke, S., 2010. Risk Change and Contributions to Risk in Offshore Helicopter Traffic on the Norwegian Continental Shelf. Proceedings 10th International Probabilistic Safety Assessment & Management Conference. PSAM 10. Seattle, USA.
Kråkenes, T., Håbrekke, S., Herrera, I.A., 2009. Risk influence modelling of recent developments in helicopter safety on the Norwegian continental shelf. In: Reliability, Risk and Safety: Theory and Applications: 1133-1140. The Netherlands: CRC Press.
Ministry of Transport and Communications, 2001. Helicopter safety on the Norwegian Continental Shelf. Part 1: Organising of the public authorities' involvement. Norges Offentlige Utredninger NOU 2001: 2. (Norwegian Official Report).
Ministry of Transport and Communications, 2002. Helicopter safety on the Norwegian Continental Shelf. Part 2: Trends, goals, Risk Influencing Factors and prioritised measures. Norges Offentlige Utredninger, NOU 2002: 17. (Norwegian Official Report).
OLF, Oljeindustriens Landsforening (The Norwegian Oil Industry Association), 2010. OLF Recommended guidelines for flights to petroleum installations OLF-066. Stavanger.
OLF, Oljeindustriens Landsforening (The Norwegian Oil Industry Association), 2009. Helideck manual. Stavanger.
Reiman, T., Oedewald, P. (2009). Evaluating safety critical organizations. Focus on the nuclear industry. Swedish Radiation Safety Authority, Research Report 2009:12.
Rosness, R., Blakstad, C. H., Forseth, U., 2009. Framework conditions' significance for major accident risk and work environment risk – A literature study. SINTEF report A11777. ISBN 978-82-14-04817-9. Trondheim. Norway
Van Steen, J., 1996. *Safety Performance Measurement*. Institution of Chemical Engineers U.K.
Weick, K., 1995 *Sensemaking in Organizations*. Thousand Oaks, CA; Sage.
Weick, K., 2009 *Making Sense of the Organization (vol. 2). Chichester: John Wiley & Sons.*