Enhancing Safety:
The Challenge of Foresight

ESReDA Project Group
Foresight in Safety

EUR 30441 EN

ESReDA
European Safety, Reliability & Data Association

How to cite this report: ESReDA Project Group Foresight in Safety, *Enhancing Safety: The Challenge of Foresight*, EUR 30441 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-25189-7, doi:10.2760/814452, JRC122252.

"Enhancing Safety: The Challenge of Foresight"

Edited by ESReDA Project Group *Foresight in Safety*
Available from: ESReDA website: www.esreda.org,
Joint Research Centre https://publications.jrc.ec.europa.eu/repository
and https://op.europa.eu/en/web/general-publications/publications

Preface by ESReDA President

Foresight in safety is a critical capability for high-risk industries. It is needed to achieve sustainable, reliable and safe performance: failures of foresight have been observed in most accidents. Foresight in safety requires a mindset within operating companies and regulators to think ahead and to remain critical about current risks that are not fully controlled or acceptable, and to develop vigilance to new threats. We should humbly acknowledge that risk anticipation is a difficult challenge and has no perfect, simple or definitive solution in complex systems. Indeed, systems are dynamic, continuously evolving with important changes, such as the current massive digitalisation, and external challenges - some with potential major crisis consequences such as the COVID19 pandemic that surprised many stakeholders.

The European Safety, Reliability and Data Association (ESReDA) was created almost thirty years ago to help stakeholders to cope with these kinds of challenges. Its principle is cooperative: across boarders in Europe; as well as between industries, institutions, and authorities. ESReDA is a non-profit making association of European industrial and academic organisations that promotes advances in the fields of safety and reliability. ESReDA is a forum for experts to exchange potential strategies and solutions to cope with these kinds of challenges. The association always welcomes comments and contributions concerning their publications and invites all to submit ideas for further developments in the field of new safety measures and reliability data as well.

The analyses and proposals presented here were written by an ESReDA project group. This report would not have been possible without substantial individual efforts by the ESReDA project group members who come from different companies, research institutes, universities and authorities. They have produced its contents without any financial support and have devoted considerable free time to the task. This publication is based on extensive experience from several industrial sectors (transportation, energy, chemical…) and countries in Europe. ESReDA is proud to present the results of their work and hopes it will benefit the many organisations and individuals worldwide concerned with foresight in safety challenges.

ESReDA would like to thank the authors for their contribution and also the member organisations for funding travel expenses for its members. In particular special thanks are due to those organisations that have allowed working group members to participate in this work including giving free access to their extensive in-house expertise and experience. We record our appreciation and grateful thanks to:

- Agenția de Investigare Feroviară Română (AGIFER), Romania
- Collectif Heuristique pour l'Analyse Organisationnelle de Sécurité (CHAOS), France
- Électricité de France, EDF R&D, France;
- Energias de Portugal (EDP) – Gestão da Produção de Energia, S.A., Portugal;
- European Commission, DG-Joint Research Centre, Energy, Transport & Climate, (JRC), the Netherlands; and DG-Joint Research Centre, Space, Security & Migration, (JRC), Italy
- Federal Public Service Employment, Labour and Social Dialogue, Belgium
- Fondation pour une Culture de Sécurité Industrielle (FonCSI), France
- Institut de Radioprotection et de Sûreté Nucléaire (IRSN), France;
- Kindunos Safety Consultancy Ltd, the Netherlands;
- Noordwijk Risk Initiative Foundation (NRI Foundation), The Netherlands,
- Societatea Nationala Nuclearelectrica SA, Romania,
- SRL Health, Safety, and Environment Consulting, Norway
- Starline, USA
- Tukes - Safety Technology Authority, Finland;
- University of Pardubicze, Czech Republic

In addition, we specifically thank the Joint Research Centre of the European Commission for their support on the publication and dissemination of this ESReDA deliverable. We hope this report meets the expectations of members of the public and organisations who have shown interest in the work of the group in this important field.

Porto, November 2020
Luis Andrade Ferreira
Universidade do Porto
Former President of ESReDA

Saclay, November 2020
Mohamed Eid
CEA
Current President of ESReDA

Preface by the ESReDA Project Group *Foresight in Safety*

This report is the result of a joint effort by experts, working in the fields of risk management, accident analysis, learning from experience and safety management. They come from 10 countries, mainly from Europe but also from the USA and Australia. Their expertise covers several industrial sectors.

The report aims to provide useful information, from both theoretical and practical viewpoints, about "Foresight in Safety", based on current practices and the state of scientific knowledge.

Safety is an area concerned by ongoing debate (e.g. is goal of safety to ensure that 'as few things as possible go wrong' or to ensure that 'as many things as possible go right'?), but we can assume that safety implies continued correct operation of the process such that errors and failures do not lead to major accidents.

The contents of this publication are summarised below.

Chapter 1 presents the case for a new approach to "foresight in safety", based in particular on the literature in future studies, and presents its theories, traditions, and challenges;

Chapter 2 shows how foresight is conceptually located between system resilience and listening to whistleblowers;

Chapter 3 aims to characterise some of the failures to foresee major accidents in order to foster the ongoing debate about the best strategies to enhance "foresight in safety";

Chapter 4 claims that organisational loss of memory is a topic to be taken into account, in particular the loss of knowledge regarding Early Warning Signs;

Chapter 5 demonstrates that scenarios describing potential event sequences are a practical tool for thinking about risk;

Chapter 6 points out that visualisation of early warning signs is critical to a clear understanding of the causes of existing weaknesses and to defining proper actions to prevent their escalation;

Chapter 7 shows how the ESReDA Cube[1] can be used to identify foresight potential for detection of Early Warning Signs;

Chapter 8 focuses on the investigation and diagnosis of dysfunctional organisational factors, in particular those than may lead to a safety level decrease with potential negative consequences;

Chapter 9 identifies areas of the asset management process that can be used to generate safety foresight, enabling the detection of process and systems/equipment deterioration and other anomalies before a serious accident can occur;

Chapter 10 explains how analysis of "big data" can help to detect Early Warning Signs of system failure, and can predict occurrence of very infrequent events affecting safety;

Chapter 11 focuses on the role of whistleblowers and argues that taking advantage of information provided by these persons could help avoid the occurrence of events;

Chapter 12 presents technologies, domains and applications that can be used to improve safety directly and by enabling use of foresight;

Chapter 13 describes some daily activities of regulators and discusses the ways in which regulators can generate and disseminate foresight during these activities.

This report can be read either as a whole (by downloading the full report) or you can read only single chapters (a hyperlink is available for every chapter on the ESReDA dedicated webpage – www.esreda.org).

---

[1] *The ESReDA Cube is a conceptual model focused on the "learning from accident" process. It was developed by a previous Project Group dealing with "Dynamic Learning as the Follow-up from Accident Investigation". (https://esreda.org/wp-content/uploads/2016/03/ESReDA-dynamic-learning-case-studies-180315-1.pdf).*

All members of the Project Group have been actively involved in preparing this report from 2015 to 2020.

An overview of the group's participating members with their name and affiliation is given below.

- Ludwig Benner (Starline, USA)
- Bastien Brocard (EDF, FRANCE)
- Nicolas Dechy (IRSN, FRANCE)
- Yves Dien (CHAOS, FRANCE)
- Antonio Felicio (ESReDA, PORTUGAL)
- Milos Ferjencik (University of Pardubicze, CZECH REPUBLIC)
- John Kingston (NRI Foundation, The NETHERLANDS)
- Paulo Maia (EDP, PORTUGAL)
- Eric Marsden (FonCSI, FRANCE)
- Sever Paul (AGIFER, ROMANIA)
- Sverre Røed-Larsen (SRL Health, Safety, and Environment Consulting, NORWAY)
- Dan Serbanescu (Nuclearelectrica SA, ROMANIA)
- Zdenko Šimić (EC JRC, The NETHERLANDS)
- John Stoop (Kindunos, The NETHERLANDS)
- Miodrag Stručić (EC JRC, The NETHERLANDS)
- Tuuli Tulonen (Tukes, FINLAND)
- Frank Verschueren (Ministry of Labour, BELGIUM)
- Ana Lisa Vetere Arellano (EC JRC, ITALY)

This PG was co-chaired by N. Dechy, Y. Dien and F. Verschueren. The treasurer responsibility was ensured by A. Felicio.

This report is edited by the ESReDA Project Group 'Foresight in Safety' as a whole.

The authors of each chapter are listed in each chapter.

The introduction and conclusion have been prepared by J. Kingston, A.-L. Vetere Arellano and Y. Dien on behalf of the project group.

A continuing odyssey[2]

Some members will continue in the new project group starting in November 2020 on "**Risks, Knowledge and Management**". New participants and volunteers are welcome!

The new project group and JRC will organise the 58th ESReDA seminar on this topic in the Netherlands in 2021 (postponed from 2020 due to Covid19 crisis).

Some members have been cooperating in the ESReDA framework since the beginning of 2000. Indeed, the Project Group "Foresight in Safety" follows on from other ESReDA Working / Project Groups.

The Working Group "**Accident Analysis**" (1993-2000) published two deliverables:

- "Directory of Accident Databases" (1997),
- "Guidance Document for Design, Operation, and Use of Safety, Health, and Environment (SHE) Databases" (2001).

Then the Working Group "**Accident Investigation**" (2001-2008) issued in 2003:

- a survey report entitled "Accident Investigation Practices – Results from a European Inquiry"
- a book in 2005 "Shaping Public Safety Investigations of Accidents in Europe".
- a report "Guidelines for Safety Investigations of Accidents" in 2009.

The Project Group "**Dynamic Learning as the Follow-up from Accident Investigations**" (2009-2015) issued three reports in 2015

- ("Case study Analysis on Dynamic Learning from Accidents — The ESReDA Cube, A Method and Metaphor for Exploring A Learning Space for Safety",
- "Barriers to Learning from Incidents and Accidents",
- "Training Toolkit: Learning from Incidents and Accidents").
- An essay was also issued by Pr. J. Stoop "Challenges to the investigation of occurrences".

---

[2] *As coined by Ana Lisa Vetere Arellano*

Since 2000, these Groups also organised a number of **ESReDA Seminars**:

- the 24th seminar in 2003 on 'Safety Investigation of Accidents' in Petten (the Netherlands), organised jointly with the JRC-Institute for Energy;
- the 33rd seminar in 2007 on 'Future challenges of Accident investigation' in Ispra (Italy), organised jointly with the JRC-Institute for the Protection and Security of the Citizen;
- the 36th in 2009 on 'Lessons learned from accident investigations' in Coimbra (Portugal), organised jointly with Energias de Portugal;
- the 45th in 2013 on 'Dynamic Learning from Incidents and Accidents' in Porto (Portugal), organised jointly with Energias de Portugal;
- the 53rd in 2017 on 'Enhancing Safety: the Challenge of Foresight' in Ispra, (Italy), organised jointly with the JRC-Institute for the Protection and Security of the Citizen;
- the 55th in 2018 on 'Accident Investigation and Learning to Improve Safety Management in Complex Systems' in Bucharest (Romania), organised jointly with AGIFER (Romanian Railway Investigating Agency).

The 53rd Seminar was hosted in 2017 at JRC, in Ispra, by A. L. Vetere Arellano and Z. Simic, with the support of M. Ioakeimidou. The proceedings were edited by A.-L. Vetere Arellano, Z. Simic and N. Dechy.

The 55th Seminar was hosted in 2018 at AGIFER, in Bucharest, by S. Paul with the support of V. Patrascu. The proceedings were edited by S. Paul, E. Marsden, F. Verschueren, T. Tulonen, M. Ferjencik, Y. Dien, K. Simola, V. Kopustinskas.

The reports as well as the seminar proceedings are available on the ESReDA website[3] either to order or as free downloads. On that aspect, the project group is also grateful to I. Šarūnienė (ESReDA Secretary) and K. Simola and V. Kopustinskas (JRC) for the proper publication of seminar proceedings on ESReDA website.

---

[3] *https://www.esreda.org/*

Introduction

ESReDA Project Group on Foresight in Safety[4]

This introduction gives an overview of topics in foresight contained in thirteen chapters written by members of ESReDA's project group on Foresight in Safety. The group did not attempt a comprehensive treatment of foresight. Instead, experts in safety have shared ideas about the role and challenges of foresight in their field—mainly industrial safety, and focussed on accident investigation and prevention. But the chapters span a broader scope; and address foresight in process safety, nuclear safety, transportation safety, occupational safety, consumer safety, and medical safety.

If you are a reader that is interested in foresight in socio-technical systems and safety, this is a book for you. In the main, the authors have concentrated on a pragmatic approach with a relatively short-term time horizon. Their aim is to enhance the level of safety in important, societal fields. However, there are also some observations that would apply equally well to 'futures research' and related fields. The reader is directed to the conclusions chapter in particular.

As you read, you will find links to the individual chapters. It is suggested that you read all the way through this introduction before reading the chapters. Please note that a related volume of seminar proceedings is also available for download.

The first part of this introduction presents signposts to ways of conceptualising foresight. These include how foresight is seen within different paradigms such as resilience engineering, social science, and accident investigation.

The second part of this introduction considers foresight alongside various aspects of socio-technical systems. There are three main headings: foresight as a property of organisations; foresight in the prediction and control of operations, and lastly; foresight as a subject for regulatory action.

---

[4] *The introduction has been prepared by John Kingston, Ana Lisa Vetere Arellano and Yves Dien on behalf of the project group.*

Conceptual views of "Foresight in Safety"

Foresight is based on knowledge of what has happened before and what has yet to happen. Chapter 5, Use of Scenarios, refers to this as retrospection and prospection, respectively.

Chapter 1, Foresight in safety. Theories, traditions, and challenges – a new approach, shows how the issue of foresight has been considered in the field of safety. It describes how some methodologies for foresight are already implemented in some sectors. The time concept in foresight and foresight traditions and futures research is described and analysed. The chapter also considers the relationships between safety, investigations and the modern system approach. The strategic triangle and resilience are also discussed.

Chapter 1 reviews several theoretical perspectives on foresight. The authors consider apply insights from futures research and resilience engineering to foresight in safety. Resilience engineering is further considered in Chapter 2: Future of safety, resilience. In a similar vein, Chapter 9, Asset Management, Monitoring and KPIs, discusses the role of foresight in asset management and makes comparisons to foresight in safety.

Different theoretical perspectives offer alternative approaches to improving foresight. However, it may not be easy to combine them. As Chapter 2 points out, engineering paradigms and socio-psychological paradigms construe foresight in different ways. Notwithstanding their distinctiveness, different perspectives can be complementary, and arguably this is essential in an inclusive approach to foresight.

Although views differ about what foresight is and how it works, there seems to be general agreement that foresight should be a trigger for learning and risk reduction. The connection between foreseeing and acting is sometimes implicit, such as identifying early warning signs through analysis of scenarios. This is explored in Chapter 5, Use of Scenarios. Scenarios are used as a practical tool for thinking about risk; they are relatively straightforward to create and have many uses. In contrast, the authors of Chapter 7, Utilizing the ESReDA Cube to detect EWS, look at the explicit steps that lead from analysis to action. These steps include identifying which actors are best placed to interpret findings, finding solutions, and

initiating change in the system. Similarly, Chapter 6, Visibility of Early Warning Signs, describes how meaningful signals need to be escalated for review and action. However, there are many situations when an actor—a foreseer as it were—cannot get the right kind of change to happen, or happen with sufficient urgency. Chapter 11, Whistleblowers, explores these situations.

Foresight and levels of system

There are many different perspectives on foresight. It can be seen as a technical question concerned with how people work with technology to better understand and control operational systems. Foresight can also be seen as an aspect of governance of the organisation that owns or has a role in the operation.

The organisation

Foresight as a function of organisations

As explained in Chapter 3, Failures of Foresight in Safety: Fantasy Risk Analysis and Blindness, investigations of major accidents often reveal problems of foresight. As well as questioning the technical adequacy of hazard identification and risk quantification, investigators are increasingly willing to consider structural problems—dysfunctions—in the organisation itself.

Chapter 8, Organizational dysfunctionalities, describes how these conditions are important features for safety which, when degraded, predispose operations to accidents. The chapter also describes how a competent authority assessed the organisational functions of major hazard sites, including those associated with foresight, before accidents. The authors report instances of important functions working badly in organisations, apparently without any awareness by top management. According to the competent authority's inspectors, the factors of special relevance to foresight are often located at a senior level, and include strategy, policy, structure, resources, roles, and responsibilities.

Blindness, deafness and whistle-blowers

The assessment scheme described in Chapter 8 can be seen as a normatively defined contrast between a functional and a dysfunctional organisation. Chapter 3 makes a comparable distinction; using the metaphor of physical ability and pathology, it notes that foresight problems can be considered as blindness and deafness. As noted in Chapter 11, in whistleblowing cases, management appear

insensible to the foresight voiced to them, with only subsequent accidents restoring their sight and hearing. Even if blind to the risks that whistle-blowers express, organisations tend to excessively fixate on the whistle-blower as an enemy in their midst, rather than acting constructively on their message. Chapter 11 gives examples of this and discusses ways that allow whistle-blowers to voice their concerns without the denials, suppression and punishment that invariably follow.

The operational system

Foresight, considered earlier as an organisational duty, can also be considered as a capability. As in most fields, tasks once seen as uniquely human are increasingly assisted by technology. And whilst technology can undoubtedly assist, foresight is largely an exercise in human cognition, professional knowledge, and human relationships within socio-technical systems.

Technology to support foresight

Chapter 12, Role of Technology, points out that people use foresight in a variety of situations, such as development, training, operation, monitoring, diagnoses, prediction, emergency response, and accident management. Technology needs to be designed to accommodate the distinctive requirements of these situations. Chapter 12 reviews the state of these developments. The authors consider the assistance provided by sensors, computing power, communication bandwidth, computer-aided hybrid development, real-time modelling analysis and artificial intelligence. Chapter 10, Big Data Analytics and Early Warning Signs, looks at how foresight by humans can be assisted by patterns detected using big data technology. It describes how big data approaches can help to detect new safety threats, improve the monitoring of safety barriers, facilitate structural health monitoring, and provide inputs to safety investigations. The promise of big data approaches is tempered by a number of issues, not least the inscrutability of the algorithms used and data adequacy.

Knowledge, visualisation and Early Warning Signs (EWS)

Visualisation and early warning signs (EWS) are a recurrent theme in safety foresight. Chapter 6, Visibility of Early Warning Signs, argues that foresight depends on clarity about the meaning of early warning signs. Chapter 5, Use of Scenarios, explains that knowledge of EWS is critical to foresight, and describes how scenarios can be used to identify them. Chapter 3, Failures of Foresight in

Safety: Fantasy Risk Analysis and Blindness, and Chapter 2, Future of Safety, Resilience, note that when EWS are weak signals; they present a special challenge to foresight. Chapter 3 also underlines that, whatever their strengths, EWS seldom trigger actions that are effective in terms of safety.

Knowledge of early warning signs, and their meaning, depend on memory. As Chapter 4, Loss of Memory, explains, memory can be personal, externalised in a database/repository or a combination of these. And in Chapter 5, Use of Scenarios, it is explained that foresight depends on some aspects of memory more than others. However, Chapter 4 points out that because memory is not intrinsic to organisations, risk management depends on deliberate efforts to create this capacity. Organisations suffer memory loss in a variety of ways, and retention of memory is a goal that needs to be managed. Chapter 4 describes the challenges of organisational forgetting and presents approaches that may be helpful in retaining memory. These approaches include using scenarios and the early warning signs associated with them.

Chapter 7, Utilizing the ESReDA Cube to detect EWS, discusses visualisation, not of EWS per se, but of systemic relationships. It presents the ESReDA Cube as a tool to assist stakeholders to make sense of situations, find solutions and improve foresight.

Role of regulators

Chapter 13, The Role of Safety Authorities in Providing Foresight, and Chapter 8, Organizational dysfunctionalities, remind of the unique role that regulators play in the governance of safety. As asserted in Chapter 3, Failures of Foresight in Safety: Fantasy Risk Analysis and Blindness, organisations may be foresight-blind and, as noted in Chapter 8, regulators are well-placed to discover this and challenge the organisation's management and culture.

Chapter 13 describes the unique opportunities that regulators have for foresight in safety, by virtue of their privileged acquisition of data from all levels, and their overview. It describes the role of regulators to generate and disseminate foresight of risks, and how this is discharged by competent authorities.

Table of contents

1   Theories, traditions, and challenges – A new approach

Sverre Røed-Larsen, SRL Health Safety Environment Consulting, Norway,
John Stoop, Kindunos Safety Consultancy Ltd, the Netherlands,
Jan Erik Karlsen, University of Stavanger, Norway.

Executive summary

In this chapter we explore historical relations between safety, foresight, innovation, and policymaking. We also look at how these relations got lost over the last two decades and how they can be restored. Socio-economic drivers, political philosophies, and social values shape foresight in safety. We have taken a top-down perspective to gain insights into these higher order forces.

The chapter outlines the historical background of foresight, reviews the evolution of foresight-theories, and lists the methods used. The time concept in foresight, foresight traditions, and futures research is described and analysed, as are the relationships between safety, investigations, and the modern system approach. The strategic triangle and resilience are also discussed.

Among the recurrent themes discussed is the role of safety in legacy and innovative systems, the full information paradigm in combining feedback and feed-forward control of safety, and the role of resilience engineering.

A discussion on the 'Foresight in a world at risk', illustrated by the 2020 coronavirus pandemic, stresses the need to be organised in order to safeguard resilience. In summary: sense and learn from the past, make-sense and act in the present, and prepare for the unexpected future.

The approach towards a safety foresight methodology and challenges is outlined, and examples given of foresight implementation in areas such as management, education, and learning.

Finally, we suggest integrating several notions as building blocks for a multidisciplinary activity in the domain of safety and foresight. Recommendations are made for a new holistic safety management based on feed-forward as well as on feedback information and insights.

1.1   Introduction

1.1.1   Combining foresight and safety

During the work of various ESReDA project groups on safety, the topic has been shifting. Starting in 1993 with exploring the early phases of the investigation process – as data collection and guidelines for investigation of accidents - to later phases like dynamic learning as the follow-up from accident investigations. The focus on foresight represents a shift from reactive to proactive approaches. What are the origins of foresight as a pro-active notion, and how are they related to safety?

1.1.2   Foresight, how it began

In addressing the concept of foresight from a historical perspective, Martin (2010) clarifies various early interpretations and perspectives, originating from the Science, Technology and Society (STS) debates in the 1970's. These STS debates aimed to foreseeable effects of innovations to societal developments and their problem-solving potential for practical problems. Such innovations affected a large-scale introduction of Nuclear Power Plants (NPP), Electronic Highway and DNA technology. Disasters in these areas were deemed to have unacceptable consequences and should be addressed proactively in order to make them socially acceptable.

According to Martin (2010) foresight is defined as:

"*a process by which one comes to a fuller understanding of the forces shaping the long-term future which should be taken into account in policy formation, planning*

and decision making. By clarifying input assumptions, one can come to a prediction of outputs which can be justified scientifically."

In this definition, the goal of foresight is to systematically survey all paths that could be developed and identify what options or alternatives are open. This process explicitly does not restrict itself to preferential options from a single actor perspective but covers all options from an evidence-based perspective. From this point of view, decisions of today create the future by taking actions (Martin 2010). Such foresight is based on an understanding of interrelations between science, technology, and society. It should help to stimulate public discussion of desirable futures and of the role of government in such futures (Steed and Tiffin 1986).

During its development, foresight has covered three domains of interest:

- technological innovations and their transition processes by industrial initiatives;
- policy making, assessing the impact of decisions and actions of governance control;
- foreseeable safety consequences of new technologies revealed by case-based learning.

1.1.3   A sensitivity to overarching philosophies
Projects on foresight were initiated in several highly industrialised countries: the UK, US, Canada, Japan, France, Germany, the Netherlands, and Australia. Martin observes that the development of foresight was dominated by different countries' political and socio-economic philosophy (Martin 2010).

In the USA and UK, the Reagan and Thatcherite 'New Economy' intended to 'roll back the state'. The aim was to reduce governmental responsibilities in selecting preferential priorities for policy making decisions on innovative developments. The selection of winners and prioritising was left to 'the market'. Foresight had no part of their privatisation and deregulation policy. In this political philosophy, there was no need to identify and select scientific and technological priorities. In the UK—and to a lesser extent the US—a convenient framework of scientific notions was developed by social and organisational scientists as Reason, Rasmussen, Perrow and Turner. Safety became 'emergent' and unforeseeable due to 'complexity', while accidents became 'normal' after a period of 'incubation' (Stoop, 2020). This framework disculpated those with governance responsibilities, and masked governmental failure of foresight in safety. It left foresight to those corporate

levels and gave them exclusive managerial responsibility for safety and risk control. Within the framework, safety was no longer a societal concern, but became an operational performance indicator at a corporate level, submitted to efficiency/thoroughness trade-off considerations, balanced against other process indicators such as costs and lead-times. To this purpose a toolkit with ALARP (As Low As Reasonably Practicable) criteria for accepting risks and safety cases was developed, in conjunction with Safety Management Systems with quantifiable safety performance indicators.

This neoliberal framework assumed confidence in a proper functioning of such delegated responsibility for safe operations and a fair-trade behaviour of each of the actors with respect to risk avoidance, liability, and expert knowledge (Pupulidy, 2019). The shift in responsibility, from governmental oversight and control to self-regulation in industry, was based on the assumptions that substantive safety knowledge was in the market and that governmental oversight of corporate safety management processes would suffice.

However, unforeseen vulnerabilities in these assumptions emerged over time, culminating in serious concerns about deregulation, privatisation, and a proper functioning of Safety Management Systems (Farrier, 2017; Pupulidy, 2019):

- in disconnecting content from process, a shift occurred from a factual and actual performance control to compliance with standard operating procedures. Regulatory on-site inspections were replaced by functional demands on managerial processes.
- This shift also hampered feedback from anomalies, empirical disclosure of deviations, incidents, and accidents. Lessons learned from safety investigation and recommendations at a sectoral level became detached from corporate Safety Management System input (Farrier 2017). At this corporate level, a new set of performance indicators had to be developed, such as Safety II as the expression of Best Practices and learning from successes.
- Recognition of that a degree of operator variability is normal also indicated differences between Work As Intended - by management - and Work As Done - by operators. This difference caused controversies about the acceptability of deviations and compliance with Operational Excellence (Winters 2017). Issues emerged on liability and accountability, culminating in legislation on Corporate Manslaughter and Corporate Homicide.

- Erosion of operator flexibility in task performance occurred, in particular in conditions deviating from optimal, and in crisis and disaster situations. In aviation, a simultaneous operator training was reduced to operate under standard situations, accommodating a more flexible and cheaper transition between configuration adaptations and software equipment versions in the operating environment. This shift from competence-based operator skills to compliance-based task performance eroded the notion of operator flexibility in dynamic operating environments and conditions to a great extent. It led to several catastrophic accidents. In the aviation and maritime industries, 'Good Airmanship' and 'Good Seamanship' came under pressure.
- A lack of agreement about operator performance, non-compliance with established safety standards, and exclusive managerial control created a stigmatising role as whistle blowers. This count in particular for substantive experts and experienced first line operators in assessing safety critical situations that were beyond control and awareness of corporate management.
- Several catastrophic events demonstrated that the neoliberal framework of delegating responsibilities tends to erode existing barriers and precautionary measures relied-on to prevent disaster. In particular, with the airplane model Boeing 737, disruptive developments were introduced, supported by Next Generation and MAX branding, while their certification was treated as only derivative. A decision-making tool for certification of derivative developments proved to be lacking. The Boeing 737MAX crashes have become the salutary example of unforeseen consequences of deregulation and privatisation of the civil aircraft certification regime with still unforeseeable global consequences for its revision and adaptation.

With the emergence of deficiencies of the neoliberal New Economy philosophy, a next generation of safety management philosophies is under development as a successor of what behavioural scientists called the 'old school of safety thinking'. With the acceptance of deviation as normal - inevitably manifesting itself by emergent properties - the safety debate shifted from the origin of deviations and causes of mishaps towards recovery from such deviations and mitigation of their potentially catastrophic consequences. Most prominent in this 'new school' of thinking at the organisational level is the notion of Resilience Engineering (Woods and Hollnagel, 2006). At the level of governmental oversight, retrospective independent safety investigations were institutionalised under the notion of 'Independent Investigations, a Citizen's Right and Society's Duty' (Van Vollenhoven, 2001).

Since the ability to foresee deviation and taking precautionary measures was denied due to the assumed impenetrable and inherent complexity of socio-technical systems, foresight as a notion was no longer incorporated in this safety debate. This has had far reaching implications for managing the (scientific) knowledge base for enhancing safety in complex socio-technical systems. After a seemingly stable situation of validating assumptions and expectations, these systems seem to have reached their third and final phase of development (Minsky, 1986). In this phase, a distinction between derivative and disruptive adaptations is lacking, while profit-taking is no longer covered by future developments due to a lagging investment in precautionary arrangements and scientific knowledge development (Minsky, 1986; Vincenti, 1990). According to Snowden (2007), such a final phase may trigger a transition from complex systems into chaotic systems. Such a chaotic system potentially creates catastrophic interdependencies due to its reliance on operational feedback.

As deficiencies of the New Economy philosophy become visible, a next generation of safety management philosophies is under development'. Most prominent in this new school of thinking is the notion of Resilience Engineering (Woods and Hollnagel, 2006). Foresight as a notion, however, has yet to be incorporated in this safety debate.

Outside the Anglo-Saxon world, northern European countries have seen the safety and risk debate take a different direction with respect to foresight. Based on the Rhineland governance model (Stoop, De Kroes, and Hale; 2017), the debate adhered to a concept of cooperation and deliberation. Examples of this include the Dutch consensual Polder model and the Scandinavian humanitarian philosophy of Vision Zero. These differences impacted the approach and development of foresight in safety as a societal and strategic value.

In the Netherlands, several major projects were initiated by the government in the public debate on the desirable future and the role of government (Martin 2010). The emphasis was on forecasting the consequences of policy making with respect to introducing nuclear power, the electronic highway, water management, land use planning and large railway infrastructure projects. Several Parliamentary Inquiries disclosed emergent market failures in realising these projects and invoked the justification of governmental initiatives and interventions in these

areas (TCI, 2004; Van Kleef, 2016). To facilitate technological transition strategies, network-based Public Private Partnerships were created. Contractual conditions for risk liability changed from DC (Design and Construct) to DBFMO (Design, Build, Finance, Maintenance and Operations), changing the financial accountability relations and risk management responsibilities between public and private partners in such networks (TCI 2004). Along with these new contracting forms, privatisation and liberalisation became the norm.

This development was in line with EU initiatives on R&D projects in Framework programmes, based on long-term planning, such as the EU Vision 2050. Such major projects were supported by the establishment of R&D institutes similar to the RAND Corporation in the USA, and the creation of research networks between academies, industry, and universities. Educational courses were established, such as a faculty of Technology, Policy and Management at Delft University of Technology. Leading multinationals, such as Philips, Shell, and Unilever, developed in-house methods for innovation and change management in cooperation with academia and research institutes (Berkhout, 2000).

In these developments, safety has competed poorly against other corporate priorities such as environment, sustainability, circular economy, and climate change challenges (TCI 2004). The decay of safety concerns coincided with complacency in government and industrial legacy sectors where safety had achieved an outstanding performance level. Transport, nuclear power, and the process industry were assumed to be Non-Plus Ultra-Safe, leaving room for only marginal safety enhancements at very high costs (Amalberti, 2001).

1.1.4 Two worlds drift apart

Foresight on safety in the abovementioned sectors became disconnected from their technological developments, while the scientific debates on safety shifted from substantive assessments to managing process control, risk perception and risk acceptance standards. Safety was assumed to approach a theoretical asymptote value of 10e-7 which would leave residual risks as highly unlikely and therefore, negligible. Consequently, there was no trigger to explore R&D needs and development in safety investigation methodology beyond accident modelling and Human Factors research. In the New Economy philosophy, process drives out content, market drives out knowledge. Even a question was raised whether safety science was superfluous to existing (social) disciplines or was a science at all (Safety Science 2014).

In the scientific debate on recognition of social sciences in the foresight domain, a wide variety of different terminologies, paradigms and notions emerged (Martin 2010). Simultaneously, a dialectic stall emerged in the safety debates, confronting safety notions and interpretations from a variety of perspectives (Safety Science 2014, Stoop, Hale and De Kroes 2017). The variety in terminology created confusion and controversies in both domains (Martin 2010, Safety Science 2014). As stated by Martin (2010): terminology is vitally important in the social sciences. 'The emergence of a new term often heralds the identification of some new phenomenon, or at least the recognition of an existing phenomenon that, until now, has laid undetected by social scientists. He identifies several threats to coining new and unique phrases: a particular choice of phrasing may either greatly enhance the prospect or ruin the chances of that research having any appreciable impact. It also may create problems in establishing intellectual property claims on intractable problems and cause loss of persuasive arguments to incorporate foresight in a political philosophy. Finally, while allocation a new and unique label may attribute newly discovered phenomena to the reputation of social scientists, it may give rise to priority disputes in their discipline and in such disputed cases, accusations of plagiarism among colleagues and discrediting or rejection of scientific schools of thinking by practitioners (Martin 2010, Zimmermann et al., 2011; Stoop, 2019).

A less virulent consequence of coining phrases is the gradual separation that occurs across scientific disciplines, in particular between engineering design and social sciences, where a 'debate of the deaf' occurred. Due to differences in language, contexts and operating conditions, separations can also occur between various industrial sectors, academic debates and safety investigation practices. Such a separation can be observed with respect to safety between resilience engineering, safety science and the aviation sector, each developing their own reference framework, paradigms, methods, and tools (Zimmermann et al., 2011). A striking example of such a difference across sectors is present between the process industry and aviation, questioning whether there is a distinction or not between process safety, occupational safety, external safety, rescue and emergency safety at a governance or corporate level (Stoop 2019). Such differences also created diverse problem definitions and problem-solving strategies across disciplines and application domains (Martin 2010). Such differences also raise doubts about the extent to which the theories and notions in foresight generalise to the field of safety, and vice versa.

1.1.5 Feedback from reality

Each of these disciplines and domains applied specific approaches, covering impact assessment studies, probabilistic risk assessment and safety management policies. Both legacy sectors and new technologies were submitted to such safety and risk assessments, focusing on perception and acceptance of either occupational safety, process safety, environmental safety, rescue and emergency, recovery and resilience, criticality and vulnerability issues (Van Kleef, 2016). Most of these debates were ad hoc and driven by events. Such managerial assessments considered residual risks. However, risks assessed as more remote than the 10e-7 frequency limit, would be deemed negligible and their potential catastrophic consequences expelled from the equation. Instead of understanding such events, the absence of investigating their nature and context caused ignorance about their complexity and dynamics. Devils in the details were not scrutinised. Furthermore, their social impact, public perception and acceptance were not considered in the decision making on their acceptability. This managerial safety and risk philosophy created a category of very low frequency/catastrophic consequence events which were not foreseen to their full extent but were nonetheless considered 'normal' (Perrow 1999). Only in the 1990's, after a series of iconic disasters, did their criticality and social impact became the subject of academic interest.

The 1990-2000 era revealed complacency in the governmental oversight of this category of catastrophic events; concerns were raised about foreseeability and acceptability. Notions of prevention, proaction, recovery, resilience and foresight became buzzwords in the academic and policy making debate. Safety 2 was coined as a proactive, complimentary 'new school' notion for the reactive 'old school' of safety 1, accompanied with a plea for a paradigm shift in safety thinking (Safety Science 2014, Stoop, De Kroes and Hale 2017).

At the same time, several iconic accidents in the 1990-2000 era in the high-tech industries of various industrialised countries raised concerns about the predictability and societal control over major safety and risk events. The main examples are noted, below.

- Several major air crashes occurred shortly after one another in the Netherlands: Bijlmer Boeing 747 (1992), Texel DC3 (1996), City Hopper, Schiphol (1994), Eindhoven Hercules (1996), while international TWA 800 (1996) and Concorde (2000) crashes shook public confidence in aviation.

- In the railways, train crashes occurred in the UK at Clapham Junction (1988), Channel Tunnel (1996), Ladbroke Grove (1999), Hatfield (2000), in the Netherlands near Hoofddorp (1992), in Germany at Eschede (1998), and in Norway (2000).
- Passenger ferries capsized in 1987 (Herald of Free Enterprise) and 1994 (Estonia), while severe oil spills occurred with the Exxon Valdez (1989), Braer (1993), Sea Empress (1996) and Erica (1999)
- Nightclub fires with large numbers of casualties occurred in Sweden (Gothenburg 1998) and the Netherlands (Volendam, 2000), while a firework explosion in Enschede (2000) and a nitrate explosion in Toulouse (2001) destroyed a complete neighbourhood
- In the process industry, in 1976 the Seveso disaster and in 1984 the Bhopal disaster occurred, while Harrisburg (1979) and Chernobyl (1986) disrupted the nuclear energy sector.

These accidents became iconic because they served as wake-up calls and triggers for change in these industries and in the prevention of such events. Learning from accidents to prevent recurrence of similar events became a political topic in order to restore public confidence in industrial sectors and regain governance control over disastrous events and their aftermaths. Recovery from industrial disaster became relevant, while the Hurricane Katrina flooding stimulated resilience engineering thinking in the public domain. In addition to already existing subjects, new policy domains were explored such as rescue and emergency, public governance and oversight, prevention and proaction. In 1997 the Swedish Riksdag (Parliament) adopted the concept of Vision Zero; no fatalities in road safety as a risk acceptance policy making goal, while several countries took initiatives for establishing independent safety investigation agencies. All across Europe, investigation agencies broadened their traditional perspective from the transport sector to other sectors of industry and public governance on either a single mode or multimodal and multisectoral basis. In 1993, the community of independent national transport safety boards established an international network, the ITSA (International Transportation Safety Association). This sharing of experiences and learning from each other by feedback from reality originated from their experiences with case based and evidence-based learning. Due to a series of major disasters in various domains, the Netherlands took a leading role in this development. Independent safety investigations became a governance role model

for industrialised countries across the world under the motto of 'Independent Safety Investigations, a Citizens' Right and Societies' Duty' (Van Vollenhoven, 2001). Safety investigations into specific events provided the necessary feedback for prevention and proaction. This investigative approach was acknowledged by the European Union (EU) by issuing a series of Directives, institutionalising independent safety investigation agencies in various sectors and domains. Foresight based on feedback from reality provides a powerful, plausible, and credible retrospective approach. However, prospective foresight, based on theoretical grounds and scientific methods, was not incorporated in this knowledge network development.

1.1.6   Three driving forces

This chapter identifies three higher order driving forces that govern relations between foresight and safety. These offer a means for long term development. Each of these three forces is embedded in a specific context of the science, technology, and society (STS) debate:

- in societal policy making, foresight reflects the acceptability and sustainability of the consequences of new technologies and their social benefits;
- in technological innovation, foresight in industries assists the change and transition management processes that introduce new industrial developments and deliver their economic benefits; but,
- in the scientific domain, safety and foresight have become separate disciplinary activities, both in feedback learning and in the feed-forward assessment of new technologies.

In conclusion, there seems to be a unique opportunity to re-unite and integrate safety and foresight by combining a feedback and feed-forward perspective on long term future developments.

1.2   Thinking about the future

1.2.1   Five different attitudes to future

Human beings have always been concerned about their place in existence: the past, present, or future. Many have been especially concerned about the future that lay in front of them individually, in front of their families, or in front of their group. Today, we also include the nation, major regions such as EU, and the global community.

An individual's point of view dictates their attitude towards the future. Almost any aspect of belief or identity is pertinent: religious, political, social, economic, demographic, commercial and other variables such as ethnicity, age, gender, status, and sexual orientation. Some general views:

- The future as fear and threat (religion, but as heaven in a new life!)
- The future as happiness and joy (ideology, religion, social engineering)
- The future as unimportant and immaterial (determinism)
- The future as characterised by risks, probabilities, and possibilities (science)
- The future as adaptive and prosperous (technology and socio-technical engineering)

The time horizon may for analytical reasons be divided into short term, middle and long term.

Each of these approaches have been described in religious literature (the Bible), in many philosophical books, in scientific works, in technical papers and books, in novels and poetry, in science fiction, etc. Many conceptions about our future destiny form part of our oral traditions. Famous persons, who have contributed to futuristic thinking, include i.e. Leonardo da Vinci, Jules Verne, H.G. Wells, Herman Kahn, Johan Galtung, Stephen Hawking, Aldous Huxley, Robert Jung, George Orwell, and Alvin Toffler.

Some recent examples of global threats include studies made by OECD, studies concerning opportunities and trends in technology (South by Southwest, 2016) and several climate reports.

Another example which highlights challenges more than threats is Samsung's SmartThings report about Future living. This is an example of a study with a very long-time horizon (a 100 years hence); it deals with the huge implications of the digital revolution on our lifestyles, homes, cities, and countries.

1.2.2   Theories and their scientific background

The scientific approach to foresight dates to the 1950s, with the start of Technology Assessment and Forecasting. Today, modern safety thinking has been elaborated in many directions and is applied to many different subjects. Foresight includes the use of a variety of methods and techniques (Popper 2008 a/b; Jackson

2013; Sager 2017). The specific notion of 'foresight in safety' is analysed and defined in a separate chapter.

The scientific approach labelled as 'foresight' is defined in contrast to another discipline called variously 'future research' or 'futures studies'. 'Future research/futures studies' were often disputed within scientific circles: could such an approach – which was not based on theories and hypotheses and tested against empirical data - be included as 'real scientific research'? Or was it an art? Although a final agreement has yet to be reached, it is clear that the study of futures (possible, probable, or preferable) has neither the traditional characteristics of natural sciences nor the methodology of some social sciences (Selin 2008). However, futures studies are now both an academic branch and a business. The academic use of futures studies can be found in the environmental/climate sector and dedicated research centres, often with scientific programmes. However, far more extensive are the semi-commercial (e.g. think tanks) or purely commercial consultancies offering a widely sought-after, broad repertoire of techniques, such as trend studies/trend analysis.

As a form of futures studies, strategic foresight studies had many early authors and scientists that initiated or anticipated the more systematic and knowledge-based understanding which were established after WWII. The use of strategic foresight studies grew mainly within defence planning and expanded later to the public sector (state/regional innovation), to large regional organisations (such as EU) to the private sector (such as multi-national companies).

1.3   Foresight as an object of research

1.3.1   The time concept in foresight

Safety implies change, and change – seen as a process – is embedded in time. The time concept represents a fundamental challenge in philosophy because our thoughts about the social world and time reside inside time itself. It may be debated whether there is anything that exists outside time. As foresight mostly deals with the temporal called 'future', it is vital to establish a kind of consensus within which foresight management can operate. In foresight, the time concept is reconstructed. Often, we divide the time span of the future into short, medium and long-term perspectives; short being 5-10 years, medium 10-20 and long-term 20-30 years and beyond. This time perspective is clearly socially constructed, but

for what purpose? A plausible explanation for the conventional use of time horizons in foresight may be found in the purpose of the foresight itself. Since most examples of foresight (like safety management) have a clear action orientation, they need a trustworthy perspective not stretching into the eternity, rather limit it to a few decades.

Implicitly, most examples of foresight apply an operational definition of time, not strictly linear, but still a chronological concept, Karlsen et al. (2010) claim:

"The past is seen as something that has ended, having no starting point but bordering the present, which in turn is defined as the state we experience now and actually live in. Now is consequently something which is there all the time, pushing the future to a state which is not actually here, other than in our minds."

The future is constantly approaching us but is reconstructed in the organisationally recognisable time horizons applied in foresight management. The reconstruction does not change the ontological characteristics of future, just make it easier for us to handle time as an embedded aspect of the changes we imagine when undertaking the practice of foresight management. Nordlund (2012) surveyed how well-known futurists considered timescales in their central works. Like Karlsen et al. (2010) on foresight, Nordlund concludes that 'the time-scale has not been given special attention', other than when specifying scale terms, like short, medium, and long (ibid, p. 413) in futures research and forecasting. Thus, these fields (i.e. foresight and safety management) do not have a theory of time, just the mentioning of time as a rather loose and boundary condition.

1.3.2   Brief outline of foresight traditions

The foresight approach is part of a wider scientific tradition: to use analyses about the past, about the present situation (diagnosis), to identify future objects and the possibility to reach them (prognosis), and how to reach the future goals (prescription). However, here again, the actual studies differ in many ways between the two extremes: on one side pure basic scientific research about the future, and on the other side pure business studies, e.g. in the context of strategic foresight management.

Georghiou (2001) has defined foresight as an approach that overlaps three other disciplines: future studies, strategic planning and policy analysis. Although 'foresight' has been connected to, or partly integrated in, other research fields, the foresight tradition as a whole has some unique elements.

Some characteristics of the foresight approach are:

- **Process:** cross-disciplinary and cross-sectoral participation, and action-oriented.
- **Time:** medium to long term perspectives (often 5 – 50 years) in contrast to 0 – 5 years for risk assessment (short perspective).
- **Goal:** aimed at present-day decisions and mobility/joint actions by identifying possible future developments, driving forces, emerging technologies, barriers, threats, and opportunities.
- **Results:** outlooks, proposals of future developments, scenarios, visions, roadmaps, and actions.
- **Prerequisite:** the world is multi-dimensional and basically uncertain and complex.

The importance of foresight studies and explanations can be illustrated by the multitude of actors who are using foresight theories and methods. Both individuals (researchers, authors, scientists etc.), university institutes and organisations (Foresight professional networks, public-sector foresight organisations, and non-governmental foresight organisations) have allocated resources in order to develop and implement foresight studies and results in many sectors.

As examples may be mentioned as networks World Future Society and World Futures Studies Federation, as organisations in the public sector National Intelligence Council and NASA /both US), The Institute for Prospective Technological Studies (EU), Government Office for Science (UK) and Norwegian Research Council (Norway), as NGOs Rand Corporation, Hudson Institute, Copenhagen Institute for Future Studies, Strategic Foresight Group and Project 2049 Institute. The reports and findings may be published in journals like Futures, Journal of Future Studies, Technological Forecasting and Change, and the Futurist magazine.

1.3.3   A promising future for a new discipline

Foresight has developed as a scientific research field during the last years and has theories, hypothesis and concepts that have been elaborated. Many universities around the world now have foresight research on their research agenda, and some have also established scientific degrees and education programmes. Outside of universities, the foresight approach has been used by several public and private institutions, enterprises (especially multinational companies) and consultancy firms, think tanks, etc. The main implementation is connected to change management, strategic analysis, and policy development.

The EC was an early adopter of foresight research (technological foresight, regional foresight etc.). The EU Commission supported in 2009 the development of a European platform in foresight (EFP - Project no.244895). The emphasis by the EU institutions stimulated and created innovation across the EEA, as national initiatives, and new research programmes. The goal was not only to develop a broad spectrum of methods nor to introduce a new kind of thinking or create valuable processes, but to direct the processes into action, which could enhance constructive changes in today's practices.

Some numbers can illustrate the focus the EU has had on foresight. A search of publications via the EU Science Hub produces 452 hits on foresight. Among them, four books and 180 articles. In addition, EU has organised several conferences, workshops, scientific programmes, and expert groups in the foresight field. The scope has been very wide, ranging from global perspective, as 'Vision of the world in 2035' – a foresight report issued by The Defence Technical Information Center (US) in 2016, via many environmental topics, such as climate change, land use, water usage, wind potentials, weather-related hazards and regional climate, to government, migration, employment, and big data in road transport policy, as well as nanotechnology, low carbon energy technologies, and sustainable food and nutrition security – all key words from reports published in 2018/19. Concerning today's organisation, EU has established a separate Unit for the Foresight, Behavioural Insights and Design for Policy at the Directorate General Joint Research Centre (JRC) in Brussels.

In the EU, however, programmes like Horizon 2050 are formulated in terms of values and goals and not in terms of quantified performance indicators with various options for adaptation and transition strategies. They do not indicate how to achieve and how to assess these goals and values. This is left to underlying scientific research and development programmes. Such programmes, however, frequently restrict themselves to the early phases of innovation and transition processes, as expressed in the notion of Technology Readiness Levels (covering TRL 1-3 on a scale of 1-10 discriminating 10 phases in the S-curve of system life cycles). Later phases of this TRL process are left to applied sciences (4-6), industry (7-8) and private entrepreneurs (9-10) which take the final steps to their market. In those latter stages, the information on the developments foreseen has become

private company confidential assets. By definition, foresight in early phases of development should enable democratic participation on the foreseeable consequences in the mid- and long term for society in general, based on knowledge and insights that are open to scrutiny from different perspectives, values and interests.

Defined in this perspective, foresight is:

- a process, discriminating several steps;
- a focus on predefined aspects, symptoms, and patterns; and,
- a judgement call, identifying values and decisions from a multi-actor perspective.

In general, foresight assessment can be built up by combining tools and techniques from different domains and disciplines, stakeholder perspectives and value judgements.

In foreseeing the acceptability of future performance, innovations and transitions suffer from a phase called 'Valley of Death'. After an initial start, setbacks occur that may oscillate into unforeseen stagnation and failure. Many promising socio-technological developments do not survive these setbacks and perish. Because such problems may emerge later than foreseeable on the short term, a (specific) time horizon should be identified in which foresight is a reliable, plausible, credible and feasible predictor for future performance.

Potential building blocks for such a foresight process are:

- iterative assessment of findings and change agents by the Cyclic Innovation Model (Berkhout, 2000);
- presumptive anomaly as expressed in the Variation Selection Model (Vincenti 1990);
- identification of showstoppers/stealers and disruptive factors in the innovation process;
- identification of societal changes, values, business models and risk awareness, perception and appreciation;
- decomposition of a systems architecture and dynamics with its safety critical decisions during design, development, introduction, midlife upgrade and demolition;

- identification of knowledge deficiencies, assumptions, simplifications and limitations of the scientific body of knowledge, available during several phases of assessment;
- feedback from reality across domains, disciplines from multiple perspectives;
- similarities with socio-technological projects in the past as a learning experience.

In addition to the dynamic role the European Commission and its departments have had in developing foresight as research and a tool for decision making (including the shortcomings), a growing national interest in the foresight discipline has, during few decades, fostered several research institutes, research programmes, books and reports, conferences, workshops, and education at university level throughout Europe (see also 1.3 above).

1.4 Safety: investigations and the 'modern' systems approach

In the Anglo-Saxon safety debate, a predominant and relatively pessimistic retrospection prevails. Systems are believed to be too complex for foresight and risk assessment to deal with. Taleb (2008) launched the metaphor of 'Black Swans' as the ultimate inability to explore and comprehend socio-technical systems. And, as Donald Rumsfeld (US Secretary of Defense, 1975-1977 and 2001-2006) suggested, 'unknown unknowns' may always limit our knowledge of the future. Safety science seems to be at the edge of a paradigm shift, both from a theoretical and a practical perspective. The European safety science community study a wide array of new approaches. Some challenge the validity of safety science as a science (Safety Science, 2014), while others proclaim new safety concepts and notions, such as Resilience Engineering, a 'New View on Human Error' or Safety I and Safety II. Such developments challenge and redefine commonly shared notions such as precaution, cause-consequence relations, human performance, cognition, and culture with sometimes far reaching consequences for their application. ESReDA advocates the generic value and applicability of safety investigations across industrial domains and scientific disciplines. ESReDA foresees a predictive foresight on safety and its integration in a system engineering perspective. In several industrial sectors with a high-tech nature, safety is considered a shared responsibility, superseding a single actor or mono-disciplinary perspective. Life Cycle Analysis seems indispensable for an assessing safety throughout the life cycle

of complex legacy systems, addressing specific characteristics of transport, process, and nuclear power applications.

Within safety, it may be useful to measure 'weak signals' and other indicators. There are various approaches that help: investigation, scenarios, risk analysis and assessment.

Future thinking may be in use in different industrial sectors (such as energy production, the production of chemical substances and products, consumable production, transportation and to some extent also in the consumer-/service sector), but often restricted to a short or medium-term time horizon.

1.4.1   Safety in legacy and modern systems

Such new thinking was accompanied by a change in moral and ethical values on safety. Traditionally, technical design has relied on notions such as failsafe and safe rational decision-making theories do not provide satisfactory explanations of abnormal life, crash worthiness, damage tolerance, compartmentation, redundancy, and reliability. But recent developments show that this is changing. With the introduction of ICT as a fundamental new technology, new ethical notions such as Value Sensitive design and Responsible Innovation principles have been developed. They deal with complexity, system design and integration of safety assessment by Encompassing Design and Multidisciplinary Design Optimisation methods, Knowledge Based Engineering and Value Engineering. New legal definitions dealing with safety assessment and liability have been introduced such as Corporate Manslaughter and Corporate Homicide; shifting social responsibilities for unanticipated consequences back to manufacturers and designers.

The consequences of application of new materials such as composites, technological innovations in ICT, food, system-of-systems networks, and Internet of Things cannot be predicted and assessed by today's evaluation methods. A new combination of learning from feedback and feed-forward is not yet developed and validated. New thinking, as illustrated by the ESReDA Cube (see chapter 7), has indicated several opportunities to tackle such quests.

Since safety of innovative complex and dynamic systems cannot be assessed based on their past performance, new approaches and notions should be developed. A distinction between socio-organisational and socio-technical system categories becomes inevitable, dealing with their intrinsic, inherent, and emergent properties as specific classes of hazard, threats, and consequences. A distinction between

high energy density systems and dynamic network concepts is necessary to deal with massive instantaneous outbursts of energy of a mechanical, chemical, or nuclear nature and the way consequences propagate through networks. A new distinction should be made between normal, undisrupted performance which is highly predictable and controllable, and non-normal situations, emerging from drift, natural growth, aging and exceedance of designed performance envelopes. New mental representations of human performance become necessary, since Tayloristic models of compliant behaviour and behaviour in normal situations or normal behaviour in abnormal situations. A Good Operatorship notion dealing with competence rather than compliance is under development in several high-tech sectors such as in aviation and the maritime counterbalancing prospects of full automation towards unmanned operated transport systems (Mohrmann et al., 2015).

In assessing their safety performance, we can not only deal with new systems and technological innovation. Existing systems in their full maturity have long histories of past performance and have gone through a series of decisions, assumptions and modifications that are hardly fully known, let alone documented. The notion of transition management in matured, complex systems with a high level of technological change potential is in its early phases of development. A distinction between disruptive and derivative technology is crucial to understand its dynamic behaviour. Due to the very high-performance levels, such catastrophic consequences can manifest themselves as very high consequence and very low probability events beyond the responsibility of individual actors and entities. Interferences may occur due to unknown interrelations between components that have been forgotten, neglected or unexplored. In practice, such dynamics are consigned to the category of 'unknown unknowns' but are actually discernible as design-induced consequences during operations. Foresight also includes knowledge and operational experience-based hindsight.

The role of accident and incident investigations can gain a new dimension if such aspects are incorporated in the investigation methodology. A common investigation methodology across industries and disciplines should lay the basis for such a new approach in order to create a level playing field. This would need legal recognition and procedural embedding into practice, such as achieved in aviation by the International Civil Aviation Organisation (ICAO) in its Annex 13.

1.4.2 Foresight in technological innovation: trends and opportunities

Traditionally, many industrial companies have concentrated on learning from past events and developed internal safety policies and industry norms after that. Past events include accidents, production problems, distribution, and usage problems. Many safety authorities, including regulatory agencies, have also followed this pattern. Feedback to the design of technology and organisations and managing safety during operations have greatly benefited from such learning. Social scientists designed, created, and proclaimed a category of Non-Plus-Ultra-Safe systems, such as aviation. There are, however, necessities and opportunities to combine feedback and feed-forward learning, integrating safety as a social value at all systems levels and lifecycle phases, equivalent to health, environment, wealth, sustainability, and prosperity.

Safety management based on a systematic combination of learning of past events and issues and analysis and methods for insight into the future challenges seems still not very widespread within several key high-risk areas. This ESReDA project group aims at reinforcing feedback and feed-forward loops between hindsight and foresight experiences and expertise.

Safety is to be revalued as a strategic societal value, instead of the presently preferred notion as a key performance indicator within organisations, to be assessed against other operational aspects such as economy and efficiency. Safety is a public value, not only a corporate value within an ETTO (Efficiency Thoroughness Trade Off) decision making context on an operator level. A shift back from control to comprehension is inevitable in dealing with modern, complex, and dynamic socio-technical and socio-organisational systems in their operating environment.

Only by re-addressing the context of such systems, can a credible foresight on their nature and safety performance be established.

A transition is taking place in safety thinking. It is moving from reactive, to proactive, and to predictive thinking. This transition is reflected in thinking about both technological change and developments in society:

- in technological developments with respect to technological innovation and disruptive applications;
- in socio-economic and social developments with respect to risk awareness, perception, risk acceptance and management.

A 'Zero Vision' paradigm is emerging: no risk is acceptable and lethal accidents are intolerable. At the same time, systems become more embedded, complex, and dynamic. In the transport sector, although systems safety performance has achieved a Non-Plus Ultra-Safe (NPUS) level, the scale of operations themselves is increasing with respect to volumes, numbers and sizes of transport technologies and the energies that can be released from them. The law of diminishing returns applies to conventional safety management solutions, and new directions are needed to achieve improvements. The present authors see a trend towards new notions that deal with foresight during operations such as early warning signs (EWS), or recovery from non-normal situations achieved through resilience engineering. Both developments erode the need to remain vigilant and proficient with respect to safety. Investments in road safety have been reduced in some European countries. Consequently, the death toll in Europe is increasing again. Safety in aviation is jeopardised by the limits to growth due to the capacity of the infrastructure, both airside and landside. Such system related developments can be foreseen by analysing their architecture and exploring higher order drivers for change and efficiency, such as business models, policy making and governance.

With respect to socio-technical systems with a non-plus ultra-safe performance level (especially those in aviation, railways, maritime, nuclear and the process industry) can be considered as belonging to a specific category of high energy density systems, capable of creating catastrophic consequences of a physical nature. Preventing accidents of an unprecedented magnitude remains a prime reason for existence for safety investigations.

There is no golden bullet that will serves as the single encompassing safety performance indicator. An analysis of the safety performance in aviation indicates a complex interaction between airworthiness requirements and passenger service performance indicators. Rather than aiming at a further decrease of the overall accident rate as performance indicators, safety enhancement efforts could be invested in a better understanding of the system principles and properties. Safety investigations are a pivotal approach to this purpose.

The need for a system change can be recognised in two ways:

- an incremental shift in the derivative solutions for known problems; or
- a substantial shift marked by disruptive solutions for new problems.

In the second case however, innovation processes and adaptations cannot be implemented by a single actor or from a single perspective or discipline. The concept of cyclic innovation (Berkhout, 2000). Unlike a traditional, linear design model, cyclic innovation emphasises the complex interaction of multiple actors and paradigms. This concept promises sustainable effects, which if not predictable are at least are descriptive or comprehensible.

The magnitude of energies that are to be controlled during normal operations and can be released during accidents is comparable between aviation, railway, and the nuclear sector. However, the variations within each mode of transport are great. To demonstrate the nature of classes of socio-technical systems, a specific class of high energy density dense systems with specific catastrophic potential is defined. High speed trains, large commercial aircraft and nuclear power plants indicates Many managers do not want to invest in innovations and push their current approaches to the limits with sometimes disastrous results. (see Minsky)

Such a class of systems requires specific approaches with respect to technological foresight. A sudden release of the energy content requires specific control over recovery and rescue capabilities. Resilience alone is not enough to take control after the release of such amounts of energy. The role of the operator as the ultimate manager of the total energy content is critical as the last line of defence in controlling the stability of such systems in both normal and non-normal situations. If such systems are not inherent stable, a delicate balance must be maintained in controlling the stability and flexibility during operations. Destabilising such systems by design - such as with the Boeing737MAX, or non-normal operating conditions - such as with repair and maintenance of NPPs, or de-qualification of operator skills - such as with deprived basic flying skills of pilots, puts high pressure on operating performance standards. Foresight of potential failure modes during design evaluation, and operational feedback by weak signals - such as whistle blowers, become indicators for timely adaptation, modification, and Good Operatorship requirements.

Weak signals are not weak by definition. Based on signal theory, there are several reasons for a weakness of signals:

- strong signals can be suppressed to weak signals;
- the signal can be misinterpreted because of distortion during transmission;
- a signal can be missed in the spectrum at the receiving end;
- a signal can be overruled by a signal of another nature;

- the frequency of transmission can fall beneath a perception threshold level.

In practice, weak signal debates deal with either the technical, behavioural or social nature of signals, with primary production processes or secondary processes, while the diversity across actors and stakeholders may create confusion and disagreement of their validity as service providers for user's safety or for technical reliability.

Table 1: High energy dense systems (1 MW = 1000 kW)

	Weight tons	Speed Km/h	Altitude meters	Energy Mega Watt
High Speed Train	430 tons	250 km/h	ground level	1053 MW
		320 km/h	ground level	1740 MW
A380 Jumbo jet	MTOW 575	900 km/h	10,000 m	75,000 MW
	at take-off MTOW575 tons	260 km/h	ground level	1500 MW
	at landing MLW 386 tons	260 km/h	200m above ground level	1252 MW
Nuclear power plant	Average size			800 MW
	Borsele (Neth)		Sea level	450 MW
	Chernobyl		Sea level	600 MW
	Fukushima		Sea level	784 MW

1.5 Foresight in a World at Risk

Many countries have developed contingency plans for tackling worst cases and wicked problems. These are most often based on foresight methods and, notably those that use scenarios (van der Heijden, 2005) in which the most dramatic outcomes are described. Scenarios usually relate to outcomes that can be described as probable, plausible, or possible (Voros, 2003). One often finds knowledge summaries of experiences that are assumed to have similar outcomes. Scenarios are conceivable e; not "wild cards".

In scenarios, it is assumed that the development of events is seldom unambiguous or predetermined. When conditions are complex, uncertain, and ambiguous—or

when one needs to take a long-time perspective—it is difficult and often risky to make precise predictions. Renn (2008) proposes a classification for risk management where methods and procedures are linked to the concepts of complexity, uncertainty, and ambiguity. The 'Achilles' heel' of such risk and crisis management models lies in the 'translation' from challenges, goals and tools to action and active preparedness.

The Covid-19 pandemic can be used as an illustration of the treacherous nature of foresight management, i.e. foresight as a management capability (Amsteus, 2008); words and plans are not translated into resilient action. Thus, the scenarios preparing for pandemics like this become just stories of a foretold illness and death for most countries.

1.5.1 The Coronavirus pandemic

On March 11th, 2020, the World Health Organisation officially changed its designation of SARS-CoV-2, the illness caused by a new coronavirus, from an epidemic to a pandemic. Earlier, on 31 December 2019, China informed WHO about several cases of a new pneumonia, possibly originating from a fish market in Wuhan, China. On 7 January 2020, the virus emerged in Europe. Within few weeks, it spread globally to become a pandemic that affects an exceptionally high proportion of the population. Covid-19 is the unofficial name of the disease caused by the SARS-CoV-2 virus.

What do we (the authors as lay people) know about the coronavirus? At the outbreak, not much. We know that coronaviruses are rather common human and animal viruses. Four such viruses cause symptoms of the common cold, and three (SARS-CoV, MERS-Cov and Covid-19) cause more serious lung infections (pneumonia). Like SARS and MERS, the novel Covid-19 is a disease that starts in animals and is initially transmitted from animals to people. We also know that the virus is not a living organism, but a protein molecule. It is covered by a protective layer of fat, which, when absorbed by the cells of the eyes, nose and the inner lining of the cheeks, changes its genetic code (mutates) and converts them into aggressor and multiplier cells. Since the virus is not a living organism, it does not die as such but disintegrates over time. The disintegration time depends on the temperature, humidity and type of material where the virus particle lies.

Besides these facts, we have learned that the virus is highly contagious and when resulting in the Covid-19 disease, may cause severe acute respiratory syndrome,

and even death, particularly in elderly people. However, the virus is fragile, only protected by the thin outer layer of fat. That is why we have learned that washing our hands with soap is the best remedy to protect ourselves. The foam dissolves the fat layer, the protein molecule disperses and breaks down on its own. However, the recommendations from the authorities to wash our hands with soap for 30-60 seconds, practice social distancing, avoid travels by public transportation, self-isolate when needed, etc., is far easier to say than to practice. To convert the habits of the population in such manner is not easy, and the success is rather patchy in the global arena.

1.5.2 Any early warnings?

The Covid-19 is novel and not identified previously, and it is different from those that cause cold or even SARS and MERS. At least in Europe and the US, seemingly most public recommendations and measures were reactive and imposed late in the transmission cycle when proactive trend spotting, compulsory and collective preventive actions, and a genuine global emergency preparedness were needed.

To identify and analyse weak signals and early warnings is an important brand of foresight studies. The core challenge is to recognise phenomena that have crisis potential and to assess appurtenant risks and emergency options early enough to handle these strategically (Karlsen & Øverland, p. 145-146; Rossel, 2011; Kaivo-oja, 2012). So, could this corona virus pandemic and the global breakdown resulting from it have been foreseen, given the experience with previous pandemics; e.g. the Spanish Flu in 1918-20, the Asian flu in 1957, the Hong Kong flu in 1968, SARS in 2002, the Swine flu in 2009, the Ebola epidemic in 2014 and the MERS coronavirus epidemic in 2015? Arguably, should all nations have been on alert to make-sense of the initial indications sent from China?

Apparently, many medical and scientific authorities and groups projected very negative scenarios for the pandemic globally and nationally and these projections were intended as warnings. In 2019, the US Department of Health and Human Services carried out a pandemic exercise named 'Crimson Contagion'. This imagined a flu pandemic starting in China and spreading around the world. The simulation predicted that 586,000 people would die in the US alone. The core scenario message was, according to a group of New York Times journalists, rather scary (Sanger et al. 2020):

WASHINGTON — The outbreak of the respiratory virus began in China and was quickly spread around the world by air travellers, who ran high fevers. In the United States, it was first detected in Chicago, and 47 days later, the World Health Organization declared a pandemic. By then it was too late: 110 million Americans were expected to become ill, leading to 7.7 million hospitalized and 586,000 dead.

However, did this scenario trigger an alert to the US authorities? Hardly. Such projections may have generated increased anxiety, but arguably, that is much better than complacency. The journalists claimed it only resulted in a (not to be disclosed) report, emphasising, 'how under-funded, under-prepared and uncoordinated the federal government would be for a life-or-death battle with a virus for which no treatment existed'. Moreover, the US president for weeks stubbornly preached that the pandemic was negligible and controllable. By 3 April 2020, the Johns Hopkins University (JHU) reported that the number of people confirmed infected by the virus exceeded 1 million (thereof 250,000 in the US) and the death toll over 500,000 globally. Five months later Worldometer reported 27 million infected and nearly 900,000 deaths .

MacKay and McKiernan (2004 point out the role that hindsight plays in foresight studies. They argue that the past is not an isolated static state, but one that is clearly linked with the future. However, biases may influence our perceptions and conceptions of the past. These biases act as constraints on our ability to understand the driving forces that emerge from the past, play-out through the present and become the critical uncertainties in the future. A foresight bias results from a shallow perception of history and is characterised by a combination of hindsight biases, creeping determinism, and searching for information that corresponds to people's views about both the past and the future. The authors propose counterfactual analysis as an antidote to the foresight bias, linking counterfactuals with scenarios, thus translating the experiences of the past to future challenges. Unfortunately, such techniques seem to have been largely neglected in recent scenario studies like the one performed by the US government.

On the other hand, are better, more useful ideas to be found in the literature on 'risk society', on decision-making in situations of (extreme) uncertainty, or on anticipation, resilience, and sense-making?

1.5.3   From micro-cosmos to macro-chaos

Seemingly, most governments have not had in place an early warning system and supporting decision mechanisms that could have prevented the outbreak or at least lessened the spread of the virus to a tolerable extent and at a more controllable speed. A capacity for early warning could have made it possible to mount a proportionate response at the initial breeding ground in China, and to instantly disseminate the information to the rest of the world. In January 2020, Chinese researchers had published the genetic code of the virus, a requirement necessary to develop test equipment and start developing a vaccine. Some countries like China, South Korea, Singapore and Germany effected comprehensive testing and other measures and managed to restrict the spread and the number of deaths. Iceland is the only country that did a massive testing of citizens having no symptoms, which may be a key to understand the real spread of the disease.

However, many countries hesitated to act until the disease exploded in Europe and the US in March 2020. This forced governments to adapt their policies ad hoc and to express strong opinions in areas they knew little about. Thus, we witnessed quite different measures and opinions in various countries, varying from initial neglect and laissez-faire to extensive shutdowns of society. In some countries, e.g. Sweden, UK and the US, leaders proclaimed that millions had to be infected and many thousands of elderly people had to die before the pandemic would burn out. Others, e.g. Denmark and Norway, declared the opposite view: vulnerable groups (e.g. elderly, disadvantaged, physically impaired people, people living in highly exposed housing, etc.) should receive extra attention. Therefore, the Danish and Norwegian governments closed schools, bars and restaurants, shops, dentists, hairdressers, exercise studios, physiotherapists, etc. Every public place where people usually met and mingled posed a threat to these groups and the favourite measure was to lock them down.

The domino effects of the corona crises are widespread and total. In many countries the health care system is loaded to maximum capacity. Some countries like Italy, Spain and the US, most probably face a collapse or must shift to seemingly harsh triage decisions unknown in peacetime. The number of businesses that instantly locked down and the number of unemployed skyrocketed after the pandemic was declared. The economic pain has spread at a velocity equal to the spread of the virus itself. In many countries, the economy falters, as business owners and employees wonder if any stimulus package will reach them. The longer

the economic meltdown lasts, the more unlikely it becomes that the community will recover its former vitality and, moreover, the greater the risk of unsettling the social fabric that holds the economy together. Predictions made in popular and social media is that the pandemic will change the world forever (Allen et al. 2020).

1.5.4 Revisiting the Risk Society in a world without a leader
In 1986, Ulrich Beck published the now classic Risk Society. The book called attention to the dangers of environmental and industrial catastrophes and changed the way in which we think about contemporary societies. Ever since, the global dangers highlighted by Beck have taken on new forms and assumed greater importance. Financial crises have produced worldwide consequences that were completely out of control, terrorism has shifted from the regional to the global arena, waves of pandemics have swept the planet, and climate has been the most significant change-maker and defining marker in politics.

The term risk society describes contemporary social communities that seek to organise themselves in response to a future marked by global disasters, e.g.: technological vulnerability, climate change, pandemics, terror, military conflicts, political, economic, and social unrest. Global structures decouple many of the risk factors from defined localities and territories. The impact of today's and future risks can be universal. They will cross boundaries between states, geographical regions, gender, class, and cultures. Communities that pay close attention to a future under uncertain, ambiguous, and complex conditions will launch measures to prevent and reduce the impact of both current and future risk factors. In this way, they will be reflexive, expanding the capacity to sense and make-sense of novel emergence (e.g. a crisis), as well as setting the stage to reconsider the conception of a safe and robust society. However, there is a huge gap between the present reality and the ideal future solution.

Arguably, modern society is lacking the proper capability to understand the role of the future, given our perception of the past and present. This hampers the capacity to intercept emerging global shocks and anticipate novel trends, as well as setting the stage to reconsider our conceptions of present human agency. While amid the coronavirus pandemic in 2020, the point made by Beck (1999, p. 78) reminds us on the unfeasibility of being informed and rational in managing 'unknown unknowns':

"The ultimate deadlock of risk society … resides in the gap between knowledge and decision: there is no one who really knows the global outcome – at the level of positive knowledge, the situation is radically 'undividable' – but we nonetheless have to decide … so risk society is provoking an obscene gamble, a kind of ironic reversal of predestination: I am accountable for decisions which I was forced to make without proper knowledge of the situation".

The global nature of this pandemic, and the unknown features of the disease Covid-19, is changing world politics, in which risks are handled individually by various nation states for political gain. It demonstrates a global inequality and a local vulnerability and states a position far from what Beck (2009) calls a 'cosmopolitan material politics'. Rather, it is what the historian Yuval Harare characterises as (2020, p. 42-43), a 'Disease in a world without a leader'. The acute crisis facing humanity is not only due to the corona virus, but also because of a lack of trust and solidarity between humans:

"To defeat an epidemic, people need to trust scientific experts, citizens need to trust public authorities, and countries need to trust one another".

The effects of the coronavirus pandemic are evident everywhere: empty streets, shuttered shops, overflowing hospitals; closed kindergartens, schools and universities. Millions of people are laid idle by a State-ordered work ban, by the shared lack of business dealings, or by being forced to work from their bedrooms. However, despite the similar global effects, the world lacks a leadership with a common and unified strategy to cope with the coronavirus pandemic.

1.5.5 Future global chocks and the need for resilience
With no vaccine yet available, the pandemic is a drag on the global economy and a blight on social life. Arguably, there is a lack of societal resilience, meaning a capability to tackle and recover from such global shocks. However, there is abundant literature and studies reminding us about the most important lessons from past crises. One such is the OECD study on 'Future Global Chocks: Pandemics' which states (Rubin, 2011, p. 80):

"The key to any progress against infectious diseases is a structure that brings together these diverse interests in a lasting fashion. Without such a structure, the commitment to reducing the impact of infectious diseases on our national, economic and personal security will be subject to the political vagaries of the moment, leaving us unprepared for the next global health crisis."

The OECD study argues of international research centres on for the setting-up infectious disease, to serve as a nucleus for safe applications of interdisciplinary sciences globally to the benefit of all.

Another strand comes from the academic world, examining the questions 'what do we talk about when we talk about managing crises', and 'what are the threats, dilemmas and opportunities'? A point of departure may be the book 'Coping with Crises' (Rosenthal et al., 1989). This addresses major crises during the 1970s and 1980s and was followed-up ten years later by 'Managing Crises' (Rosenthal, Boin, and Comfort; 2001). The latter stated that, on its own, learning from the past has limited value to improve preparedness or the management of future crises.

Rather than accept this fatalist position, futures and foresight researchers point instead to the benefits of 'futures literacy' (Miller 2015, 2018). This is the capacity, make sense of contemporary trends shaping the future and involves informed hindsight of past events.

In the case of the coronavirus pandemic, the role of social media disrupts the supply of objective and valid information. Fake news, speculations and unsubstantiated opinions interfere with the control that governments seek to achieve through their information channels. In opinion panels, experts are selected based on their political usefulness, or play a role as whistleblower—there to criticise official theories on what the pandemic is about, its challenges and the responses thereto. These contemporary behaviours in the media add considerably to the chaos phase in disaster management.

Informed hindsight must be related to making sense of the trends shaping the future. In his chapter on viral epidemics, Alkan (2001; p.267-280) points out that communities exist in a fragile equilibrium with their ecological environment. A disturbance of this balance can cause epidemics. Alkan argues that preventing future outbreaks of deadly epidemics is nearly impossible. What society can do, however, is organise for a resilient response that best copes with cruel decisions under conditions of uncertainty. Alkan states (2001; p.277):

'Crisis management during epidemics is not simply a function of adequate models and smart scientists. In the end, crisis managers have to make decisions that encompass more than just scientific information. They have to deal with typical crisis dilemmas. Making decisions based upon an incomplete data base is the hallmark of response to crisis. As viral epidemics emerge and re-emerge, it is

preparedness, a high degree of suspicion, and rapid appropriate response that will limit the spread of these diseases in the future.'

Viral pandemics are here to stay, and they are examples of unforeseen, complex, transboundary crises with a series of domino effects on social organisation, health, and welfare. Alkan (2001, p.278) argues that while modern society is becoming more risk averse, viruses continue to modernise themselves.

If we diagnose the nature and architecture of complex systems, our ability to cope with with pandemics will no longer be restricted to responding to the consequences of disaster. The coronavirus pandemic serves as an example of a wider pattern in which viruses and diseases transfer from animals to humans. Wildlife markets, bush meat and other indigenous food chains are a primary source of contamination and spreading of new diseases. Changing the food chain and, indeed the wider system of food and nutrition is at the root of preventing pandemic events like Covid-19. Progress in the agricultural industries and virus resistant food chains will stop pandemics at their origin by coping with virus mutation, transmission to humans and across population groups, and uncontrolled spread to other world regions. Design principles regarding distributed production, unravelling chains and disconnecting networks to prevent knock-on effects, are already very well established in other industrial sectors. We can learn a lot across industry at the level of functional relations and design concepts. Analysing and understanding dynamic system behaviour and the architecture of complex systems are a prime challenge to robust systems design (Klir, 1987).

1.5.6   Chance favours the prepared mind
A transnational response structure is urgently needed. In his inaugural address as a newly appointed professor and dean at the opening of the Faculté des Sciences at Lille (7 December 1854), Louis Pasteur claimed that," In the field of observation, chance favours only the prepared mind". Strategic planning of emergency preparedness and management calls for building societal resilience capacity to sense and respond to emerging, and often what Beck (1986) named, 'invisible' risks. The coronavirus is an example of such hidden enemies. People may be infected and contagious without knowing, since many of the symptoms are mild and resemble more common diseases like colds and flu. Partly, Covid-19 is covert, invisible and not identified in large parts of the population, but partly evident, contagious and deadly in other population segments.

It is urgent to ensure that critical systems are robust, diversified and hold adequate reserve capacity. That has apparently not been the case for most countries prior to the outbreak of the coronavirus pandemic in 2020. The early warning systems were not in place, and even the reactive capacities are seemingly inadequate. Furthermore, the global partnerships are too weak and not coordinated to receive, share and integrate sources of information conducive to handle the pandemic and the societal risks resulting from it.

We know from research on previous crises that it is more important to understand the phenomenon than just to mitigate the consequences. Although we do not know all aspects of the coronavirus and the pandemic it has caused, complexity can be beaten by transparency, not by simplicity. The focus has been on the emergent consequences of exposure rather than on the transmission mechanisms themselves. It is the size of the consequences, rather than the nature of the pandemic, that has driven governmental responses. This is not an act of resilience and anticipation, preparing to prevent a next pandemic; just a firefighting effort to save what might be left after the current crisis. Consequently, it pays to be prepared.

1.6 Foresight towards a full information paradigm

Cacciabue (2004) discriminates two types of risk analysis: retrospective and prospective studies. These are complimentary and contribute equally to the development of assessment and measures. This approach rests on both empirical and theoretical platforms for evaluating socio-technical context and models. In practice, retrospection aims to identify data and parameters associated with specific occurrences, operational experiences, and context. Prospection, in contrast, aims to evaluate consequences of scenarios using a spectrum of methods, models and techniques. Taken together, this framework identifies the knowledge base needed to foresee future developments, their boundary conditions, initiating events, systemic process, and failure modes (Cacciabue 2004). Applying this approach may provide an encompassing set of safety performance indicators for foresight of the safety states of a system and to identifying areas of concern. Such areas are based on information collection and processing as described by Klir and Godet.

1.6.1 A full information paradigm

The simultaneous use of feedback and feed-forward mechanisms can be theoretically underpinned by the 'full information paradigm' of Klir (1987)—see fig 1. According to this paradigm, the body of knowledge and experience acquired in a system over decades, provides a basis for considering safety and risk (Stoop, 1990). Such a body of knowledge dominates legacy systems such as energy, process industry and transport; it makes the Non-Plus Ultra-Safe (NPUS) safe, but also reluctant to change. Their ability to adapt is hampered by vested mental constructs, assumptions and simplifications, expertise and consensus on scientific paradigms, methods, notions and techniques, both theoretical and practical.

'Old views' have to be discarded and abolished in case of a paradigm shift in safety thinking, similar to Schumpeter's 'creative destruction' on economic theory. Otherwise an opaque blending is created by mixing old and new views into a hybrid concept. In the past, we have seen a stall of such a dialectic process by proclaiming A versus B concept of safety, to be replaced by another version of C versus D. Such a debate does not restrict itself to an academic discourse but may hamper progress by creating confusion during application of these versions in legacy systems. A fall back on old views and repetition of debates across domains and disciplines frequently occurs, allocating public, corporate and personal responsibilities for safety, emphasising the roles of whistle blowers and regulators.
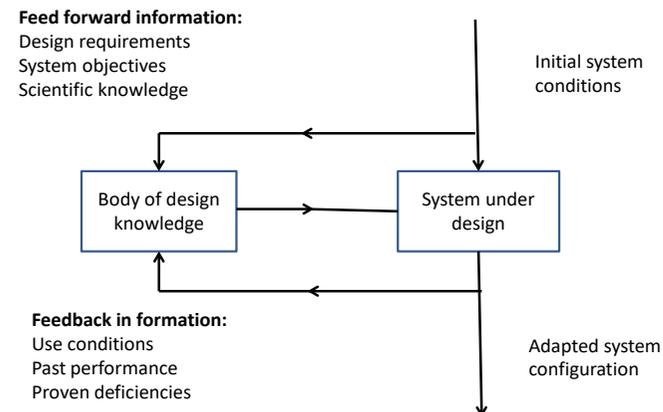
Figure 1: Full information paradigm (Hierarchical ordered control loops) according to Klir (1987)

Solving complex problems would be done better if there was greater scope for individuals to innovate solutions. To create space for individual competence, creativity, flexibility and innovation, we advocate the abolition of three obsolete notions: simplified accident models, human-error schemes, and judicial concepts of cause and blame. Taking each of these in turn:

- predefined, simplified accident models should not be used to reconstruct the course of an event. This is 'model-forcing' rather than 'model-fitting'. Instead, groups of actors should develop shared understanding of an event by using the scenario concept;
- 'human error' schemes prejudice problem-solving. Instead, a new view on human behaviour should be adopted as this invite, rather than precludes, deeper understanding of human behaviour in context;
- judicial concepts of blame and cause are fitted to the legal context of deciding liability in the courtroom. Their application should be challenged as a means for actors to understand multilinear interactions. Instead, these interactions are better understood using systems concepts, especially as the operation of feedback and feed-forward.

Abolition of the use of accident models, the notion of cause and human error as proposed by social scientists is likely to meet resistance to change due to:

- a lack of understanding of system engineering theory by non-technical scientists and practitioners;
- mono-disciplinary paradigmatic perspectives in psychology on human performance and cognition;
- disciplinary demarcation lines between technical and social sciences, and;
- cognitive stubbornness and resistance to change at both an individual, corporate and governance level.

1.6.2   The Greek Triangle according to Godet

The Greek Triangle, as formulated by Godet in 1994 and later developed into the networking action scheme (Godet 2010), sees prospective strategy as a management tool that links anticipation to action through appropriation.
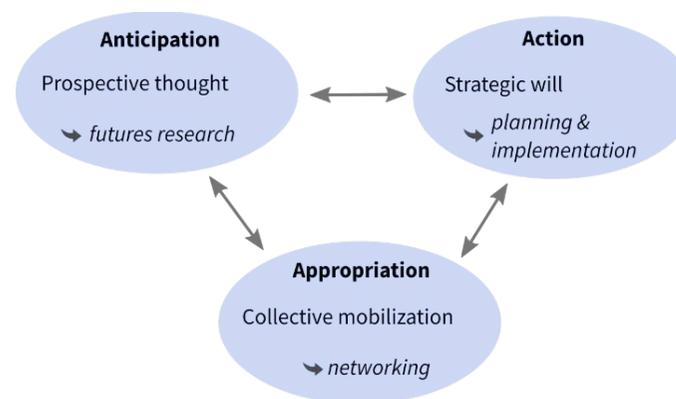
Figure 2: The relationship between Anticipation, Action and Appropriation (Source: Adapted from Godet 1994, p. 4)

Godet defines these terms:

- anticipation is the awareness of the future, and prospective thought;
- appropriation is joint commitment, collective mobilisation and sharing of values; and
- action is strategic resolve, and planning.

The triangle helps to discern the plausible future, and to develop strategy accordingly.

The three points of the triangle represent the pull, or image of the future (visual); the push, or drivers, of the present (quantitative); and the weight, or barriers, of the past (deep structure).

1.7   Foresight in safety – taking actions for a change

1.7.1   Corporate foresight

Corporate foresight is often seen as the capability of an organisation or firm to ensure its long-term survival and competitiveness by envisaging trends and detecting changes and consequences.

Corporate foresight (Rohrbeck, 2010) has been defined as: *'…an ability that includes any structural or cultural element that enables the company to detect discontinuous change early, interpret the consequences for the company, and formulate effective responses to ensure the long-term survival and success of the company.'*

However, due to shortening of product lifecycles, increased technological change, increased speed of innovation, and increased speed of the diffusion of innovations, the long-term perspective has become hard to defend. Rather, there is a constant pressure to explore and develop new business ideas, penetrate new markets and compete with aggressive competitors.

Rohrbeck et al. (2015, p. 8) argue that three novel areas of research within the field of corporate foresight should be pursued:

- Managerial cognition, which emphasises the role of the individual and group cognition in shaping perception and influencing decision-making;
- Forward-looking search, which is based on the behavioural theory of the firm. It emphasises that, as individuals are subject to bounded-rationality, firm decision-making cannot be conceptualised as purely rational or produced by analytical reasoning;
- Prospective sense-making, which considers organising as a process in which individuals build on their past experiences, and collectively reflect on these episodes to converge behind common objectives and lines of action.

Hopefully, such research endeavours would also be conducive to forming a research stream on strategic safety foresight in organisations.

1.7.2 Tools and techniques
Foresight studies must be integrated into the total safety administration of high-risk companies, industrial factories, transport enterprises, etc. This means adjustment of all the major elements of the current approach as developed over decades. These elements include risk analysis, accident investigations, mapping of unwanted events, dynamic learning, legal requirements, internal safety standards and procedures, competence development, continuous safety education, training, and change management. Hindsight lessons, insight competence and foresight studies must be part of a holistic safety management system. Debatably, the implementation of such a holistic model seems today rather rare in most private companies and public enterprises.

The foresight discipline has developed and enlarged its methodology over the years. In particular, the use of scenarios, the Delphi technique, panels, and games have become widely used, often in combination with other methods. At the same time, the content of the methods has partly changed. Whereas these methods were once the province of experts, now they are increasingly participatory; with employers, consumers, and citizens as actors.

Scientists use future techniques in their research (futurists) as do think-tanks and similar institutions. They draw on a wide range of foresight methods, including those listed in Table 2. The list is merely illustrative, and other methods exist. Note that these methods can be used for a wide variety of purposes e.g. diagnosis, prognosis, prescription or being normative, predicative, etc.

Table 2: List of foresight methods. (Popper 2008a&b; Karlsen & Øverland 2010)

- Anticipatory thinking protocols - Causal layered analysis (CLA) - Environmental scanning - Scenario method - Delphi method - Future history - Monitoring	- Back-casting (eco-history) - Cross-impact analysis - Futures workshops - Failure mode and effects analysis - Futures wheel - Technology road mapping	- Social network analysis - Systems engineering - Trend analysis - Morphological analysis - Technology forecasting - Visions

The list is not at all complete. Several authors include other methods in their foresight research. The point of this list is that foresight methods may be used for different purposes, e.g. diagnosis, prognosis, prescription or being normative, predicative, etc.

Amongst those methods listed in Table 2, trend analysis is particularly widely used by e.g. 'public planners, think tanks, foresight departments in companies and multinational enterprises. In fact, trend analysis, which is widespread within several commercial sectors, research institutes, and universities, have become a necessary tool in strategic planning, including policy making and decision making. In a systematic approach to safety, trend analysis should be linked to two

important fields: the connection to knowledge, to risk and change management and to risk analysis and learning from accidents and incidents.

Three recent different examples of trend analysis with different perspectives are shown below.

Table 3: Examples of trend analysis topics from three institutions

Institution	Institution	Institution
Simon M Atkinson/IPSOS[5]	**WATCH INSTITUTE[6]**	**FORBES[7]**
1. Dynamic Population Growing Opportunity and Growing Inequality 3. Megacities 4. Increasing connectedness and decreasing privacy 5. Healthier and sicker 6. Rise of individual choice and fracturing of the mass market 7. Rise of the individual and decline of social cohesion 8. Cultural convergence and increasing extremes 9. Always on versus off the grid 10. Emergence of public opinion's revolutionary force	1. Demographic shifts 2. Economic outlook 3. Geopolitical Issues 4. Technological Advances 5. Environmental challenges Watch Institute has identified 5 megatrends in each of the sector mentioned above – and list, describe and analyse altogether 25 megatrends.	1. The increasing datafication (sic) of our lives 2. The Internet of Things and how everyday devices are becoming more 'smart' 3. Exponential growth in computing power is fuelling massive tech advances 4. The incredible rise of artificial intelligence 5. The unstoppable freight train is automation 6. 3D printing opens up amazing opportunities for manufacturers (and others) 7. We're interacting with technology in diverse ways 8. Blockchains: An invention that could change our world 9. Platforms are the way forward for businesses

1.7.3   Foresight: from safety, via anticipation towards resilience

Park et al. (2013, p. 358-359) claim that in complex systems, risk analysis alone is inadequate to fully protect system functions and components.

This is because: "The classic risk analytic paradigm begins with hazards identification – an exercise that is problematic in the context of complex systems and emergent threats because hazards may be largely unknown".

Instead, they propose to combine risk analysis with what they call resilience analysis when working out catastrophe management plans. They claim that resilience in a complex systems context is a 'dynamic, emergent property in the context of a specific failure scenario'. Both risk management and resilience are vital to every organization. While risk analysis is well known, especially in private sector enterprises, prominent in the resilience analysis are four recursive processes, which may be modelled as a cycle (Park et al. 2013, p. 360):

1.   Sensing, by which new system stresses are efficiently and rapidly incorporated into current understanding;

2.   Anticipation, by which newly incorporated knowledge gained by sensing is used to foresee possible crises and disasters;

3.   Adaptation—the response to the information produced by sensing and anticipation;

4.   Learning, by which new knowledge is created and maintained by observation of past actions.

Resilience and anticipation deal with risks in different, but compatible, ways. Anticipation is the process of becoming aware of previously unanticipated events. According to Wildavsky (1991), anticipation is a mode of control by a central mind or actor; efforts are made to predict and prevent potential dangers before damage is done. For example, accident prevention is based on anticipating potential accidents and is enhanced by three processes of mindfulness: (1) preoccupation with failure, (2) reluctance to simplify interpretations, and (3) sensitivity to operations. (Weick and Sutcliffe, 2001; p.54). It is a strategy which aims to cope

---

[5]   https://www.ipsos.com/sites/default/files/10-Mega-Trends-That-are-Reshaping-The-World.pdf, Accessed 06 January 2020

[6] https://issuu.com/megatrendswatch/docs/global-megatrends-preview?ff=true, Accessed 06 January 2020

[7]   https://www.forbes.com/sites/bernardmarr/2017/12/04/9-technology-mega-trends-that-will-change-the-world-in-2018/#23c8f0805eed, Accessed 06 January 2020

with known threats that an organization is aware of. Anticipation means to direct your resources at one or a few specific threats, so you are best capable of dealing with that specific scenario. The point is that anticipation as a safety strategy was insufficient in today's uncertain and complex world. Under circumstances of great uncertainty and complexity, resilience is a better strategy than anticipation for managing risks. On the other hand, resilience is the capacity to cope with unanticipated dangers after they have become manifest, learning to bounce back. Dealing with unknown hazards 'as they declare themselves' is another expression for resilience (Wildavsky 1991). This perspective from Wildavsky has been further developed by several scholars within organization and risk management theory. One perspective is 'high reliability organisation' theory. Another perspective that builds on Wildavsky and the literature of high reliability organisations is 'Resilience engineering'. Debatably, Wildavsky`s dissection between anticipation and resilience has been blurred in this literature since the resilience concept in the Resilience engineering literature includes both what Wildavsky would have denoted anticipation and resilience.

Let us take a closer look at the resilience concept and the theoretical puzzle. As stated above, resilience is the idea that an individual, a technological or social system has the capacity to handle events that challenge boundary conditions. It encompasses the ability to prevent something dysfunctional from happening, or the ability to prevent something dysfunctional from worsening, or the ability to recover from something dysfunctional once it has happened (Westrum, 2006). Dysfunctional challenging events occur because plans and procedures have fundamental limits, or because the environment changes, or because the object itself adapts, given the changing pressures and expectations for performance (Woods 2009). The capacity to respond to such events, (i.e. dysfunctionalities) resides in the expertise, strategies, tools, and plans that people in various roles can deploy to prepare for and response to specific classes of change. Hence, we expect resilience to demonstrate an ability to avoid problems, to handle problems when they must be faced and to recover from damage once the dysfunctionality has happened.

Resilience is also the process of being mindful of errors that have already occurred and correcting them before they worsen or cause more serious harm. Resilience is related to accident mitigation and enhanced by two processes of mindfulness: (1) commitment to resilience, and (2) deference to expertise. Organizations committed to resilience develop knowledge and skills to cope with and respond to

errors, capability for swift feedback and swift learning, speed and accuracy in communications, flexible role structures, quick size-ups, experiential variety, skills at re-combining existing response repertoires, and comfort with improvisation. Such organizations move decision-making rapidly to those with the necessary expertise (Weick & Sutcliffe 2001).

Besides, resilience may be seen as a 'dynamic non-event' (Weick 2001), it is both dynamic and invisible. It is dynamic because it is an ongoing condition in which problems are instantly controlled due to compensating changes in components. It is invisible in the sense that it does not reveal the worst case scenarios, i.e. how many mistakes and breakdowns could possibly happen and in the sense that reliable outcomes are constant, i.e. there is nothing to pay attention to since nothing seemingly is happening inside the intended performance envelope. Visibility should be enhanced by identifying operational scenarios other than incidents and accidents.

Hindsight is the ability to understand, after something has happened, what should have been done or what caused the event. It is another way of describing retrospection. Hindsight is a useful skill that can be cultivated. Hindsight often refers to a lesson learned from something that went wrong. In hindsight, you'd know you should've paid attention to the giant 'danger' sign. In the context of foresight studies, hindsight is a form of organisational sense-making, and resilience is seen as the capacity to bounce back to normal operations after a catastrophe or some other major mishap.

Arguably, such concepts should contribute to the ambition of linking the theoretical world of foresight and the practical world of safety closer together, by explicating key concepts and implicit assumptions in both fields. However, the concept of resilience seems almost ineffable: it resists definition and description. if resilience is meant to encompass both the capability to respond, to monitor and to anticipate and by the end of day also learn both from successes and failures, resilient engineering research should illustrate the necessity to link these aspects when building resilience in organisations. An open question is - what concrete things and conditions could an observer use to make sense of resilience in airline operations, railways, NPPs and other sectors mentioned earlier in the chapter?

Besides, the theoretical puzzle prevails: How do we recognise resilience in ontological terms as long as we do not expect a person or system having a total breakdown? Subsequently, how do we perceive the ontological and

epistemological aspects of resilience to be visualised and presented? Resilience is not only a technological device, but also covering an organisational or an individual capacity meant to prevent dysfunctions to materialise or to appear if, and only if something totally unexpected happens. We may say that resilience is a systemic phenomenon that is not expected to be activated, i.e. it is not foreseen to have a future. However, if a breakdown happens, resilience is expected to serve as a safety net recovering the capacity of the system or the individual. Can this puzzle ever be solved?

1.7.4   What next?

More research beyond the level of short term, specific impact assessment studies, is needed at both national and EU levels to identify adequate and appropriate methods; and to investigate the utilitarian value of applying corporate (management) foresight perspectives to safety in a medium and long-term perspective.

1.8   Foresight in safety: the new approach

1.8.1   Five major elements in the new approach

Innovation and pioneering work is needed to apply foresight theories and methods in the field of Safety. Apart from national security, food and nutrition safety and a few other fields, safety seems to be rather absent as research object in the foresight tradition.

The new approach of the ESReDA Project Group 'Foresight in Safety', largely informed by the situation in European high-risk industries and public safety institutions. It can be characterised by five factors:

1.   A broad perception of the concept safety which may benefit from the scientific foresight tradition. So far, it seems that safety in general has had a low priority in the development of the theories and the methods as well as in the practical application of foresight insights. PG's work may therefore be looked at as a kind of pioneer work trying to combine a basic area in the modern society (safety) with a very promising and innovative scientific research discipline (foresight).

2.   The time horizon assumed in this study is essentially the near future (0 – 10/15 years?). This is not aligned with the traditional foresight approach which emphasises the value of a middle, or long-term, time perspective.

3.   The safety setting is pragmatic. The goal of foresight work in safety should help to promote and increase safety. This study emphasises the value of hindsight experiences and learning from past events, but at the same time including proactive methods and measures: as data from early warning signs, lessons from whistle-blower-cases, the challenges with loss of memory in companies and public institutions etc.

4.   We will promote a holistic programme to enhance safety in industry, transportation, public serves etc.: combining hindsight and foresight, combining lessons learned from past experiences with future trends and studies, combing systematic safety approaches from own sector and own company with experiences from other similar companies – also abroad.

5.   Lastly, we will propose to explore the possibilities to meet the safety challenges (defined as total safety) within your sector with the positive effects of a synergy approach which includes a wide perspective: the potential of enhancing safety by cooperation both within and across sectors, across national borders, across scientific disciplines and traditions, combining hindsight and foresight etc.

This background is reflected in the various topics which are covered in the present work.

1.8.2   Implementing the foresight approach

Foresight as an academic, scientific discipline is, above all, characterised by three fundamental, complementary dimensions: uncertainty, complexity, and dynamic interactions.

The first dimension, uncertainty, is a consequence of choosing the future as the subject. Uncertainty increases with the choice of time frame: with near, medium and above all long-term horizons (30 – 50 years) the uncertainty factor is extremely large. In addition, a number of other choices contribute to increase the degree of uncertainty: as organisational level (group, municipality, region, nation, continent), choice of approach (such as political, social-economic, cultural etc.), choice of sector (such as business and multinational enterprises, government

institutions or enterprises, ideal companies or various types of organisations, including NGOs and supranational institutions and organisations).

The second dimension, complexity, highlights the everyday phenomenon that we seldom predict the results of an innovative development. Technological innovation leads to brand new products, and new patterns of technical and social interaction. The development of products such as colour TV, personal computing, tablet devices, mobile phones, electrical and driverless cars or buses, are just a few examples from the past decades. The emergence of social media and the widespread use of digital tools such as Facebook, Instagram, Twitter, Google, Wikipedia, have gained within a few years, are others. The very complex interaction between the climate and the environment, changes in the settlement pattern (from rural to urban domination) are other examples. An important feature of the complexity of today is the tempo at which these changes occur. The pivotal function that knowledge production has gained in social developments and the enormous resources that multinational corporations can allocate to innovative operations further accelerate the tempo. The combination of complexity and accelerating tempo underline the importance of foresight methodology.

The third dimension, dynamic interactions, deals with the dimension of time. This dimension covers various time scales, both short term—during operations—and long term—throughout the system lifecycle. In foresight, time also covers transition phases and system states that emerge during the transition from one phase or state to another. In such transition periods, hybrid situations and conditions may create temporary disruptions and deviations from optimal performance which could be foreseen and addressed. Such hybrid periods may both create a better or worse performance that anticipated (Vincenti, 1990). They can be submitted to system erosion and deliberate interventions by extrapolating performance beyond design parameters (Minsky, 1986). In a foresight approach, resistance to change, system stability and system oscillation should be considered in advance as inherent/intrinsic properties to prevent emergent behaviour (Stoop, 2019).

The implementation of foresight theories and methods in the future safety work – with these three dimensions integrated – needs further research and studies, a willingness to share insight and experiences across frontiers, being between enterprises, authorities, research institutions, think-tanks, organisations – also across national borders.

1.9   Conclusions and recommendations

1.9.1   Objectives in an uncertain and complex future
'The future is complex and uncertain, and so are its threats to safety and security. These threats are in a different league to our existing approaches to safety, which operate on timescales that are too short, and with scopes that are too narrow. The fact that our approach to safety is outclassed by the threats we face [survival of mankind, climate and environmental problems, new artificial products …] seems to be either fatalistically accepted or simply not faced at all. Our contention is that these threats are tractable, but that it requires rethinking what we think we know about safety, and a readiness—urgency, even—to explore new ways. Foresight, we think, symbolises this new frontier.

1.9.2   Foresight and safety
Foresight can benefit safety. Some of the foresight methods and concepts reviewed in this chapter can be adapted to this end.

However, success is likely to be greater if the foresight community and the safety community communicate with each other.

- The foresight approach seems to have high potential utilitarian value for finding safety enhancements in the short term.
- The use of foresight notions and methods has so far only to a small degree been incorporated in systematic safety management at a governance and corporate level.
- The impact of residual risks and side-effects should be part of a foresight approach considering the long-term dynamics and uncertainty of innovative developments.
- In the foresight approach, the full information paradigm should be applied, benefiting from a feedback and feed-forward learning process
- In the foresight approach, higher order driving forces should be considered, as they represent socio-economical innovations, political philosophy, and social values.

The answer to complexity is transparency: de-risking of disruptive architecture facilitates foresight.

- The legacy of systems, their technological nature and temporal dynamics should be considered as inherent constraints. Before introducing them, the long-term effects of innovations in complex systems need to be better predicted and discussed by all affected, incl. across life cycle borders.
- More research is needed at both national and EU levels to identify adequate and appropriate methods and to investigate the utilitarian value of applying foresight in safety in a medium and long-term perspective.
- Explore the value of importing experiences and knowledge about the use of foresight methods to the safety arena.

The present authors see foresight located alongside safety insight and oversight:

- First, gain insight by safety investigations in critical events and occurrences as described in the ESReDA approach.
- Then, gain oversight by putting these events in the architecture of a systemic context, discriminating structure, culture, content and operating context
- Finally, gain foresight by understanding and predicting future behaviour of the system

Safety is an indispensable strategic value in the transition process from derivative to disruptive solutions in developing innovative as well as legacy systems. The main challenge for safety professionals is to develop new notions, methods, tools and techniques to cope with the challenges that accompany such a transition. These efforts could benefit from unexplored and so far uncharted domains and disciplines. Foresight is a promising prospect when addressing safety. But it will need global leadership. Will the UN, OECD, EU and WHO jointly support such an endeavour?

To paraphrase Richard Booth (1979) in his inaugural lecture in 1979:

"Safety is too important a matter to be left to futurologists".

1.10 References

Alkan, M.L. (2001). Viral Epidemics: Past and Future, in U. Rosenthal et al. (eds) Managing Crises. Threats, Dilemmas, Opportunities. Springfield, Ill.: Charles C Thomas Publisher, Ltd.

Allen, J. et al. (2020). How the World will Look After the Coronavirus Pandemic. Foreign Policy, March 20, 2020. Retrieved 4 April 2020 from: https://foreignpolicy.com/2020/03/20/world-order-after-coroanvirus-pandemic/

Amalberti R., 2001. The paradoxes of almost totally safe transportation systems. Safety Science 37, p 109-126

Amsteus, M. (2008). Managerial foresight: Concept and measurement. Foresight 10(1):53-66. DOI: 10.1108/14636680810856026

Beck, U. (1986). Risikogesellschaft – Auf dem weg in eine andere Moderne. Frankfurt am Main: Suhrkamp Verlag.

Beck, U. (1999). World Risk Society. Cambridge: Polity Press.

Beck, U. (2009). World at Risk. Cambridge: Polity Press.

Berkhout G., 2000. The Dynamic Role of Knowledge in Innovation. The Netherlands Research School for Transport, Infrastructure and Logistics TRAIL. June 2000

Booth R., 1979. Safety: too important a matter to be left to the engineers? Inaugural lecture Aston University, 22 February 1976

Cacciabue P.C. 2004. Human error risk management for engineering systems: a methodology for design, safety assessment, accident investigation and training. Reliability Engineering & System Safety 83, (2004) 229-240.

Farrier T.A., 2017. Investigations, recommendations and safety management systems. MO3763 Senior Safety Analyst JMA Solutions, LCC, ISASI, June 29, 2017

Georghiou, L., 2001. Third Generation Foresight: Integrating the Socio-economic Dimension, in: Technology Foresight – the approach to and potential for New Technology Foresight, Conference proceedings, NISTEP Research Material 77

Godet, M. (1994). From anticipation to action. A handbook of strategic prospective. Paris: Unesco Publishing.

Godet M., 2010. Future memories, Technological Forecasting & Social Change. 77 (2010) 1457-1463.

Harari, Y. (2020). Disease in a World without a Leader. Time Magazine, March 30, p. 42-43.

Hollnagel E., Nemeth C and Dekker S., 2009. Resilience Engineering Perspectives. Aldershot, Ashgate

Hollnagel, E., Paries J. Woods DD, Wreathall J. 2011. Resilience Engineering in Practice: A Guidebook. Burlington, VT: Ashgate Publishers Co.

Jackson, M., 2013. Practical Foresight Guide. Chapter 3 – Methods. Shaping Tomorrow. Creative Commons License

JHU (2020). Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Retrieved 3 April 2020 https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6

Kaivo-oja, J. (2012). Weak signals analysis, knowledge management theory and systemic socio-cultural transitions. Futures 44, 206-217

Karlsen J., Øverland E. F. (2010). Carpe Futurum. Oslo: Cappelen Akademisk forlag

Karlsen J., Overland E. and Karlsen H., 2010. Sociological contributions to futures' theory building. Foresight, 12(3), 59-72

Klir G., 1987. The role of methodological paradigms in system design. Dept. of Science. Thomas J. Watson School of Engineering. State University of New York at Binghamton. New York

Koivisto, R., 2009, Integrating future-oriented technology analysis and risk assessment methodologies in Technology Forecasting & Social Change 76 (2009) 1163-1176.

MacKay, R.B, McKiernan, P. (2004). The role of hindsight in foresight: refining strategic reasoning. Futures 36 (2004) 161–179.

Martin B. 2010. The origins of the concept of 'foresight' in science and technology: An insider's perspective. Technological Forecasting & Social Change 77 (2010) 1438-1447

Minsky H., 1986. Stabilizing an Unstable Economy. McGraw-Hill Companies

Mohrmann F., Lemmers A. and Stoop J., 2015. Investigating Flight crew Recovery capabilities from System Failures in Highly Automated Fourth Generation Aircraft. Journal for Aviation Psychology and Applied Human factors.

Nordlund G., 2012. Time-scales in futures research and forecasting. Futures, 44, 408-414.

Park, J., Seager, T.P., Rao, P.S.C., Convertino, M. and Linkov, I. 2013. Integrating Risk and Resilience. Approaches to Catastrophe Management in Engineering Systems. Risk Analysis, Vol 33, No. 3: 356-367.

Perrow C., 1999. Normal Accidents. Living with High Risk Technologies. Princeton University Press.

Popper, R., 2008a. ''Foresight methodology'', in Georghiou, L., Cassingena, J., Keenan, M., Miles, I. and Popper, R. (Eds), The Handbook of Technology Foresight, Edward Elgar, Aldershot.

Popper, R., 2008b. How are foresight methods selected? Foresight, 10(6), 62-89

Safety Science, 2014. Special Issue on the foundations of safety science. Vol 67, August 2014, pp 1-70

Pupulidy I., 2019. Can Believing In SMS Make Us More Vulnerable? https://safetydifferently.com/can-believing-in-sms-make-us-more-vulnerable/

Renn. O., 2008. Risk Governance. Coping with Uncertainty in a Complex World. London: Earthscan.

Rohrbeck, R., 2010. Corporate Foresight: Towards a Maturity Model for the Future Orientation of a Firm. Springer Series: Contribution to Management Science, Heidelberg and New York

Rohrbeck, R., Battistella, C., Huizingh, E. (2015). Corporate Foresight: An Emerging Field with a Rich Tradition. Technological Forecasting and Social Change. 101,1-9.

Rosenthal, U., Boin, R. A., Comfort, L.K. (2001). Managing Crises. Threats, Dilemmas, Opportunities. Springfield, Ill.: Charles C Thomas Publisher, Ltd.

Rossel, P. (2011). Early Detection, Warnings, Weak Signals and Seeds of Change: A Turbulent Domain of Futures Studies. Futures, 44(3), 229-239.

Rubin, H. (2011). Future Global Chocks: Pandemics. Paris: OECD. Report, IFP/WKP/PGS(2011)2.

Safety Science, 2014. Special Issue on the foundations of safety science. Vol 67, August 2014, pp 1-70

Sager, T., 2017. Foresight methods. Trondheim: NTNU, Concept report no 53

Sanger, D.E., Lipton, E., Sullivan, E., Crowley, M. (2020). Before Virus Outbreak, a Cascade of Warnings Went Unheeded. New York Times, 22 March 2020. ret.3/4/20 https://www.nytimes.com/2020/03/19/us/politics/trump-coronavirus-outbreak.html

Snowden D., 2007. The origins of Cynefin. Cognitive Edge Pte Ltd. www.cognitive-edge.com

Selin, C., 2008. The Sociology of the Future: Tracing Stories of Technology and Time. Sociology Compass 23/6: 1878-1895

Steed G. and Tiffin A., 1986. A National Consultation on Emerging Technology, Science Council of Canada, Ottawa, 1986

Stoop J.A., 1990. Safety and the Design Process. Doctoral Thesis Delft University of Technology, April 1990

Stoop J., De Kroes J. and Hale R., 2017. Safety Science, a founding fathers' retrospection. Safety Science 94 (2017) 103-115

Stoop J., 2020. Foresight between whistle blowers and resilience, this book

Taleb N., 2008, The Black Swan. The impact of the Highly Improbable. Penquin Books Ltd

TCI 2004. Veiligheidsborging van grote infrastructuurprojecten J. Stoop, p 128-137. In: Grote infrastructuurprojecten: inzichten en aandachtspunten (achtergrondstudies) van de Tijdelijke Commissie Infrastructuurprojecten (Commissie Duivesteijn) - Onderzoek naar infrastructuurprojecten. Tweede Kamer der Staten General, Vergaderjaar 2004-2005, Dossier 29283

van der Heijden, Kees (2005). Scenarios. The Art of Strategic Conversation. Chichester, England: John Wiley & Sons, Ltd. ISBN 978-0-470-02368-6

Van Kleef E. and Stoop J., 2016. A life cycle analysis of an infrastructural project. ESReDA 51st Seminar on Maintenance and Life Cycle Assessment of Structures and Industrial Systems. 20th-21st October, Clermont-Ferrand France

Van Vollenhoven P., 2001. Independent accident investigation: every citizen's right, society's duty'. European Transport Safety Lecture, ETSC, Brussels, Jan 2001.

Vincenti W., 1990. What Engineers Know and How They Know It. Analytical Studies from aeronautical History. The John Hopkins University Press

Voros, J. (2003). A Generic Foresight Process Framework. Foresight, 5(3), 10-21.

Weick, K. and Sutcliffe, K. (2001) Managing the Unexpected – Assuring High Performance in an Age of Complexity, Jossey-Bass Publishers.

Weick K., 2001. Making Sense of the Organisation. Oxford: Blackwell Business

Westrum, R. 2006. A Typology of Resilience Situations. In: Hollnagel et al., Resilience Engineering: Concepts and Precepts, Ashgate, Aldershot, 55-66.

Wildavsky A., 1991. Searching for Safety. New Brunswick N.J., Transection books.

Winters B., 2017. How to manage safety based on operational excellence principles? A case study at KLM Royal Dutch Airlines. MSc Thesis Delft University of Technology

Woods D., 2009. Escaping failures of foresight. Safety Science, 47, (4), 498-501

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

Zimmermann K., Paries J., Amalberti R. and Hummerdal D., 2011. Is the Aviation Industry ready for resilience? Mapping Human factors Assumptions across the Aviation Sector. In: Hollnagel et.al. 2011, Resilience Engineering in Practice

2 Foresight between whistle blowers and resilience

John Stoop, Kindunos, The Netherlands

2.1 Executive summary

In the safety debate emphasis is laid on the need for a paradigm shift:

- From reactive to proactive
- From prescriptive to responsive

and:

- coping with the unanticipated.

In short: the relation between safety and foresight has become a focus of attention in both theory and practice.

Several successive schools of thought can be identified in the socio-organisational and -psychological scientific domain. The role of operational feedback in each of these schools -in particular from the hot seat- is quite different. After a period of exclusion of operator feedback and compulsory compliance with a single actor - managerial- control, a re-integration of operational feedback and multi-actor involvement is emerging. A recognition of 'weak signals' in their systemic context is emerging as a means to cope with system complexity and dynamic interactions. Such recognition acknowledges the value of human variability in task performance, irrespective of blaming, shaming or framing. Such recognition should facilitate foresight on acceptable safe operator performance.

Due to the very nature of socio-technical systems however, foresight historically has been assessed in a wider context than operations variability to reduce uncertainty on unanticipated and unacceptable future performance. New technologies, disruptive designs and innovative concepts demand foresight on a safe performance without the benefits of future operational experience and feedback. From our analysis it is concluded that the old school of Reason and Rasmussen is deficient while the new school of resilience is not yet fully capable of coping with foresight in legacy systems such as aviation.

Economic theories, disruptive designs and innovative technologies, professional airmanship and subject matter expertise are identified as prime change drivers in socio-technical systems of a legacy nature. Such drivers determine the acceptance of schools of thought beyond their own internal rationales or scientific paradigm. Reflecting on the role of foresight it is concluded that new approaches such as resilience engineering have potential but can only applied successfully if they take into account the inherent properties of the legacy of the systems in which they are applied. In resilience engineering, the outsiders role of whistle blowers becomes obsolete as subject matter expertise is acknowledged as input from within the system. Such input is not restricted to individual operational experts, but is also covering independent and qualified safety investigations and inspections during both design and operations on the organisational and institutional levels.

2.2 Introduction

This paper explores the role of resilience engineering contributing to foresight in general. It focuses on feedback from reality and dealing with complexity with respect to reducing uncertainty and predicting future behaviour. It delves into several rationales that have come up in the debate about foresight and resilience engineering and puts these rationales in the context of managing risk and safety. This chapter discusses the role of whistleblowing and resilience in assessing weak signals as indicators for mishaps in matured and established socio-technical domains, referred to as legacy systems. Several competing schools of thought in the socio-psychological domain about human behaviour are explored, contrasting the 'old' school of Reason and Rasmussen with the 'new' school of resilience engineering. Questions are asked about the validity of assumptions and the role of operational feedback in such concepts, in particular regarding whistle blowers.

Moreover, these socio-psychological schools are confronted with technological thinking about foresight, in particular in the aviation domain. Aviation as a legacy sector is based on technological flexibility through the variation-selection process and knowledge management as a driver for innovation. How aviation as a unique socio-technical system has been dealing with uncertainty and foresight is explored in view of acceptance of resilience as a new notion in legacy systems. A revision of resilience engineering, with additional essentials 'initiative' and 'reciprocity' opens up opportunities to accept resilience as an engineering approach in aviation. Such acceptance complies with Good Airmanship principles in this legacy system integrating foresight in system feedback processes on both the individual, organisational and governance level.

2.3 Foresight in context

Over the past few years, a crisis in safety science and risk assessment is proclaimed (Safety Science 2014; Stoop, De Kroes and Hale 2017). In exploring new perspectives, a literature analysis indicated a reconsideration of fundamentals of safety management, risk analysis and risk management (Aven 2016, Pasman and Reniers 2016, Lannoy 2016) and a generic applicability of independent safety investigations (Vuorio et.al. 2017). Simultaneously, changes in a socio-economic context from New Economy to Circular Economy, raise questions about the validity of existing safety notions and paradigms. There is an increasing interest in resilience engineering, recovery from non-normal situations and feedback of operational experience from practitioners. Such interest is aiming to bridge the gap between Work as Done and Work as Imagined (Hollnagel 2011). Evidence Based Interventions in the medical sector are discussed as a prospective approach in processing empirical data, based on best available evidence to justify a remedy, given the state of the art in the disciplines involved. In high tech transport sectors such as aviation, railways and the maritime, forensic engineering is acknowledged as a powerful approach in providing an evidence-based intervention (Strauch 2002, Stoop 2015). Field operators and engineers are concerned with 'weak signals' as indicators for immanent failure. 'Weak signals' are considered a symptom of degradation of a system in its operational phase, exposing their assumptions, simplifications, linearizations and knowledge limitations (Dekker 2011, Dekker and Pruchnicki 2013). Safety theories and notions as developed in the 1960's and 1970's are criticized, based on experience and expertise of practitioners in various domains. Rather than hindsight, foresight should be favoured to predict, analyse and control imminent danger (Roed-Larsen and Stoop 2017).

In essence, the safety foresight debate is about uncertainty and predictability. Foresight is required because disasters are unpredictable and unacceptable, in particular in complex socio-technical systems. Foresight is concerned with questions such as: where to find data, how to interpret the information and how to adapt and change? In this quest, a specific role for whistleblowers is proclaimed. Whistleblowers fulfil a role of interpreters of scarce and uncertain information, based on their professional, domain specific knowledge and experience. How such a role can be conducted however, seems to be dependent on the specific reporting culture and type of feedback in their sector. In the nuclear sector, operational experience feedback is advocated, in the transport sector, independent safety investigations are institutionalized, in the medical sector, resilience engineering is preferred. Although not fully similar to whistle blowing in a strict sense, in the ICT sector hacking and sabotage are predominant as failure coping mechanisms, requiring a specific form of ethical engineering (Van den Hoven 2013).

The debate on feedback from operations has been dealt with from two perspectives:

- Feedback during recovery from major disruptions after an event
- Feedback during normal operations exploring the gap between theory and practice.

In dealing with uncertainty, flexibility, variety, divergence and adaptive potential, allocation of these systemic, dynamic properties can result in two equivalent, primary system configuration options:

- Keep organisations and institutions constant and vary technology. This system configuration is referred to in the Cynefin model of Snowden (2007) as 'complicated'
- Keep technology constant and vary organisational and institutional arrangements. Such a system configuration is referred to in the Cynefin model as 'complex'.

Keeping both technology and organisations constant creates closed, rigid systems without the ability to respond and adapt, while keeping both technology and organisations variable will create chaotic systems lacking effective control options on disruptive technologies.

These options require reflection on two main issues in such a configuration allocation to either man or machine:

- Human performance and the debates between the 'old school' of human factor thinking and the new 'Resilience school' of organisational thinking
- Engineering design in high tech systems and sectors with respect to variability and selecting either derivative or disruptive designs, discriminating adaptation from innovation.

These perspectives, options and configuration allocations will be dealt with in the next paragraphs by analyzing the aviation sector as a case study.

2.4 Unravelling complexity

2.4.1 Competing paradigms

In establishing a new way of thinking in safety, an artificial contrast is frequently created between an 'old' and 'new' view. Over the past decades, debates in safety have been initiated in distinguishing between occupational versus process safety, internal versus external safety, deterministic versus probabilistic thinking, technological versus social safety, safety versus security and Safety I versus Safety II. Such dialectic controversies have not been fruitful due to a seemingly endless variation on the same theme of contrasting and mutually exclusive notions and competition between scientific disciplines and industrial domains.

In their battle for recognition of humanities as a scientific discipline in safety issues, a disdain for technology and engineering design as a scientific activity has been expressed. Over 40 years, phrases were launched such as: 'Safety, too an important matter to be left to engineers' (Booth 1979) or expressed by Edwards in his presentation to the British Airline Pilots Association Technical Symposium advocating a dominant role for human factors in aviation safety (Edwards 1972). This plea coincided with the roll out of the first of a new generation of wide body aircraft, the Boeing 747, representing a leap in technical reliability and safety. Putting safety first as an objective in the Vison Zero philosophy is criticized as a 'shining example of altruism' from the perspective of trading- off safety against other system goals. Claiming zero accidents as a goal should be 'equivalent to the cries of fundamental religious groups on the right path to salvation or paradise' (Hale 2006). In 2017 however, this 'hard and shining ideal' of zero fatal accidents was actually achieved (sic!) by the international community of commercial aviation (CASV 2017). More recently, the right to exist of safety science as an academic discipline as superfluous to psychology and organisational theory was brought up in a Special Issue of Safety Science of August 2014 (Safety Science 2014).

Without achieving consensus and a synthesis that is both theoretically consistent and generically applicable in a new socio-economic and technological context, such debates frustrate progress. Rather than dialectically designing a new variation of safety notions within the same scientific paradigm from a theoretical supply perspective, a demand driven approach could be favoured with a general, basic understanding of complex socio-technical systems and the context in which they operate. Woods suggests to overcome this dialectic stall in the safety debate by defining a new unit of analysis: the man-machine-interface unit, replacing the either man or machine perspective (Woods 2016).

In the academic safety debate two competing paradigms exist: a technological systems engineering perspective and a resilience engineering perspective (Stoop 2015):

- A systems engineering approach provides a new perspective by shifting from a disciplinary to a problem solving oriented approach (Stoop 1990, Stoop 2017/1)
- A resilience engineering approach provides a new paradigm by shifting from a technical, causal approach to a socio-organisational approach with a focus on consequences and recovery from mishap and disaster (Hollnagel et.al. 2011).

As postulated by resilience engineering professionals, the latter approach conflicts with some of the fundamental assumptions which define human factors, ergonomics and socio-organisational theories as applied in industry (Zimmermann et.al. 2011).

They state that Resilience is 'the antithesis of the traditional and still prevailing, human factors and safety paradigm', referred to by Hollnagel as the 'Traditional Safety Perspective' (Zimmerman et.al. 2011). Adhering to this traditional perspective should not meet the needs of ultra-safe, complex modern industries such as aviation and may prevent further progress. Traditional ideas seem to 'remain entrenched in the perspectives and approaches of industry practitioners'. According to Amalberti, matured systems such as commercial aviation may no longer have the flexibility for dramatic or profound change (Amalberti 2001). Adaptations are supposed to remain restricted to the same underlying scientific paradigm. Their adaptation in safety thinking applies an epidemiological model as an extension of the usual sequential models. Although commercial aviation is highly standardised and regulated at an international level, there should be room for interpretation and variation of how people perform, understand and manage their work (Zimmermann et.al. 2011). Zimmermann et al. claim that it was their aim to advocate Resilience Engineering attitudes by the rejection/acceptance of the Traditional Safety perspective. They pose the question whether aviation is ready to make the paradigm shift to Resilience in view of -to their opinion- an apparent much-needed paradigm shift. They intend to 'dispel the myth that aviation is a purely technical domain in which standardisation has eliminated all

variations in how people do their work'. To their opinion, 'flying, controlling and maintaining aircraft involves more than just checklists, radio frequencies and torque settings' (Zimmermann et.al. 2011). Cultural differences between world regions should justify striking a balance between rule following and creativity, in particular in a context of diminishing resources and skills. Coping with adverse situations and conditions should not only advocate resilience on the micro level, because the macro level has stretched assets and resources too far. The system as a whole should favour resilience as a property. Although not yet formulated in terms of resilience, this is exactly what the aviation sector has achieved since the foundation of the International Civil Aviation Organisation (ICAO). Aviation has adopted a system life cycle perspective with continuous adapting, multiple feedback loops across life phases, actors and system levels (Stoop and Kahan 2005).

Zimmermann et al. (2011) notice a paradox in the relationship between Resilience and Safety: 'an unsafe system may be more flexible, more cautious, and may inadvertently foster resilience at the micro level. Similarly, a stable, safe system would have difficulty maintaining flexibility'. They observe a 'natural' tendency to increase production levels when things go right. Increasing production could increase the inherent –volume driven- risks, reduce flexibility and tighten coupling. They state: 'As aviation keeps evolving towards higher levels of standardisation, automation, procedures and stability, we must recognise that this comes at the expense of Resilience' (Holling 1973). Such a strive for operational excellence in order to increase production is driven by New Economy arguments of optimizing production algorithms (Winters 2017).

In proclaiming the myth that aviation is a 'purely technical domain in which standardization has eliminated all variation in how people do their work', social scientists are ignoring the technological and design assumptions and restrictions that are inherent to high tech safety critical systems in which open, global network configurations dominate. Since its conception, ICAO has dedicated its attention and efforts to all aspects of the aviation system performance regarding fees and fares, tariffs and trades (Freer 1986). ICAO has created an encompassing and coherent framework of Annexes to the ICAO Convention since 1951. Eventually, the civil aviation community has achieved a Non-Plus Ultra-Safe state (Amalberti 2001). In aviation, a distributed and delegated responsibility was allocated to the operators under the notion of Good Airmanship to avoid rigidity in their task performance and to enable them to deal with unanticipated situations. To avoid a chaotic system with too many degrees of freedom and disruptions, ICAO chose a strategy with technology as the flywheel for progress, keeping organisational and institutional standardization and harmonization as the prerequisite for access to a high level playing field (Freer 1986).

The desire of Zimmermann et.al. to introduce Resilience in aviation as a paradigm shift raises fundamental questions (Zimmermann 2011):

- Is there a need to make a paradigm shift in safety thinking in aviation?
- Does aviation need resilience to make such a shift?
- How did aviation become so safe in the first place as a Non-Plus Ultra-Safe system?
- What have been the safety achievements in this legacy system?
- Can we identify 'natural' tendencies as change agents for adaptation?
- How can aviation deal with foresight in view of major changes in its socio-economic, geo-political and technological context?
- Which scientific paradigms, theories and notions obstruct a transition to a Next Generation aviation industrial concept and system architecture?

In answering these questions, we elaborate on:

- Feedback loops such as whistle blowers and establishing institutional arrangements
- Change drivers such as economic business models
- Forensic engineering as a knowledge development and diagnostic potential
- System architecture regarding choices about stability, uncertainty, flexibility and trade-offs
- Creative destruction of obsolete constructs such as human error, drift into failure and complexity by replacing them with new constructs such as resilience engineering.

Developments towards resilience as a new concept for safety enhancement have their origin in criticisms on the human performance and organisational management as developed by Reason and Rasmussen. These concepts have allocated a specific role for whistle blowers and their foresight capabilities.

2.4.2   Reason: the traditional approach revisited
In his early work, Reason (2015) focused on the systemic factors underlying what he defined as 'organizational accidents'. Such accidents should differ in sharp

contrast from 'individual accidents' where damaging consequences have limited impact, restricted to their direct environment. In addition, they are supposed to have 'quite different causal pathways' compared to organisational accidents, resulting merely in loss-time injuries. Individual accidents should not have potential for predicting the likelihood of organisational accidents. In his revisited perspective on organisational accidents, Reason shifts the focus of intervention and control potential from management to those who are in the first line of defence: the operators on the spot. They are supposed to have an improved error wisdom and the power to halt the accident trajectory before harm or damage can be done. In his approach, awareness is a pivotal notion. A system safety approach should require the integration of systemic factors –labelled as collective mindfulness- and individual skills – labelled as personal mindfulness-. Political and commercial pressure are considered underlying factors for senior management to underplay in hindsight emergent, obvious threats. Because incompatible goals and organisational shortcomings may lead to disregarding clear warning signals, there should be no unambiguous responsibility for responding to weak signals by senior management. Reason advocates a shared responsibility with line management, and newly defined Safety Duty Holders, as the subject matter experts in assessing risks. All employees should be made aware of their individual safety responsibilities, supported by standards, procedures and job descriptions. A state of chronic unease should be maintained in the safety war (Reason 2015).

Reason allocates a specific responsibility to designers in their "frequent lack of awareness of the capabilities and limitations of the end user" (Reason 2015). According to Reason, many design-induced errors arise because "designers underestimate the extent to which necessary knowledge should be allocated in reality rather than in theory". In his opinion, organisational accidents are assumed to be the result of a mismatch between theory and practice. Training the mental skills of operators on underlying risk awareness are considered hallmarks for High Reliability Organisations. In order to make front-line workers more vigilant, organisational support from management is required. Individual mindfulness of danger needs to be informed, sustained and supported by a collective mindfulness of the operational risks (Reason 2015). This should enhance system resilience, converted to a lasting mental skill of foresight and maintaining situational awareness. By applying mindfulness, as Reason states, it is possible to foresee and recover from an accident. Predefined knowledge, theories and models, generated by safety scientists may even displace or marginalize existing local or system-specific safety knowledge embedded in operational practices. Hiring external safety professionals and experts with well-intended efforts, might even have a detrimental effect (Almklov, Rosness and Storkersen 2014) because their subject matter expertise might dominate managerial expertise. Reason emphasises an indispensable role of error for front-line workers: 'an incident story without mention of error or individual wrong actions is a story without a beginning. Accidents and incidents are inevitable in complex and tightly coupled systems and –hence- they are normal'. Due to hindsight biases and distorting influences in dealing with unexpected events, a narrowing of focus on the systemic factors may induce a 'premature closure on the actions of those at the sharp end', disregarding the balance between individual and collective mindfulness. Local factors distinguish systems that suffer from accidents from those that do not, because local circumstances are necessary and sufficient. Organisational factors are only conditions, not causes and insufficient to bring about the disaster (Reason 2015).

Towards a shared responsibility

Changes in the initial conditions of –complex- systems of systems create difficulties in understanding their behaviour and adaptation to the changes. These changes may incrementally decline a system into disaster by environmental pressure, social processes and unruly technology that normalize increasing risk (Harvey and Stanton 2014). Adapting to such changes throughout the lifetime of systems of systems, may be too short to enable the development of sufficient knowledge and experience to cope with the consequences. While responsibilities for systemic risks remain at an organisational level, regulations are to be developed to shift the official ownership of risk from organisation to the individual. Placing responsibilities at an individual level, is based on the assumption that each individual will do everything within their power to mitigate the risk. This assumption ensures a more rigorous safety management than the old approach of assigning risk at an organisational level, where accountability was more difficult to ascribe (Harvey and Stanton 2014). These insights in assessing risk should explicitly take into account recent incidents, changes to policies, predicted changes in government, predicted lifespan of technical components and national/international economic climates. Assessing a 'Risk-to-Life' comes down to trust in the skills and experiences of the subject matter expert involved in the risk assessment. Such a moral and ethical burden puts high demands on foresight capabilities and their potential role as 'early warning' signalling expert. Such an individual responsibility institutionalizes a role as potential whistle blower for a

subject matter expert and Safety Duty Holder. They are faced with the responsibility to communicate with stakeholders across disciplinary and paradigmatic borders of a technical, social and organisational nature.

2.4.3 Rasmussen's' role on systems modelling

In the domain of human behavior a shift of focus occurred from inferred and uncertain states of mind towards characteristics of human factors that can be framed in generic performance models. Rasmussen takes this shift one step further by proclaiming a distinction between stable conditions of the past, versus a present dynamic society (Rasmussen 1997). The present society is allegedly different by a very fast change of technology, a steadily increasing scale of industrial installations, a rapid development of information and communication technology and an aggressive and competitive environment which influence the incentives of decision makers to use short term financial and survival criteria.

Rasmussen states that modeling can be done by generalizing across systems and their particular hazard sources. Risk management should be modeled by cross-disciplinary studies, considering risk management to be a control problem and serving to represent the control structure involving all levels of society for each particular hazard category. This, he argues, requires a system-oriented approach based on 'functional abstraction rather than structural decomposition'. Therefore, task analysis focused on action sequences and occasional deviation in terms of human errors, should be replaced by a model of behavior shaping mechanisms in terms of work system constraints, boundaries of acceptable performance and subjective criteria guiding adaptation to change (Italics added). System models should be built not by a bottom-up aggregation of models derived from research in the individual disciplines, but top-down, by a systems oriented approach based on control theoretic concepts.

According to Rasmussen, rather than striving to control behavior by fighting deviations, the focus should be on making the boundaries explicit and known. Risk management should provide opportunities to develop coping skills at boundaries. For a particular hazard source, the control structure must be identified, including controllers, their objectives and performance criteria control capability. Information should be available about the actual state of the system. Control over the pace of technology at a societal level created a specific role for the regulator in protecting workers. By stating safety performance objectives, safety becomes just another criterion in multi-criteria decision making and becomes an integrated

part of normal operational decision making in a corporate setting. In this way, the safety organization is merged with the line organization. This requires an explicit formulation of value criteria and effective means of communication of values down through society and organizations. The impact of decisions on the objectives and values of all relevant stakeholders are to be adequately and formally considered by a newly introduced notion of 'ethical accounting' (Reason 2015).

A full scale accident then involves simultaneous violations of all the designed defenses. The assumption is that the probability of failure of the defenses individually can and will be verified empirically during operations even if the probability of a stochastic coincidence is extremely low. Monitoring the performance of the staff during work is derived from the system design assumptions, not from empirical evidence from past performance. It therefore should be useful to develop more focused analytical risk management strategies and a classification of hazard sources in order to select a proper management policy and information system. When the anatomy is well bounded by the functional structure of a stable system, then the protection against major accidents can be based on termination of the flow of events after release of the hazard. When particular circumstances are at stake, the basis for protection should be on elimination of the causes of release of the hazard. Design of barriers is only accepted on the basis of a predictive risk analysis demonstrating an acceptable overall risk to society. When the predicted risk has been accepted, the process model, the preconditions, and assumptions of the prediction then become specifications of the parameters of risk management. Preconditions and assumptions must be explicitly stated in a Probabilistic Risk Assessment. In this view, fortunately, Rasmussen states, it is not necessary for this purpose to predict performance of operators and management. Data on human performance in operation, maintenance, and management can be collected during operations and used for a 'live' risk analysis. Thus, predictive risk analysis for operational management should be much simpler than the analysis for a priori acceptance of the design. This also should require far less subject matter expertise. Such performance data should be collected through other sources than accident investigations; incident analysis and expert opinion extraction may compensate for the lack of abundant accident data. According to Rasmussen, the models required to plan effective risk management strategies cannot be developed by integrating the results of horizontally oriented research into different features of hazard sources and systems configurations. Instead, vertical studies of the control

structure are required for well bounded categories of hazard sources, although uniform control strategies would suffice (Rasmussen and Svedung 2000).

In conclusion, in their advocacy for managerial control, Reason and Rasmussen initially have positioned the feedback from design and operators in an outsiders role of whistleblowing. This has only partly been compensated in their revision by introducing a Safety Duty holder and ethical accounting for shop floor workers. The assumptions, limitations and simplifications of Reasons' and Rasmussens' concepts have initiated a debate among sociopsychological and -sociological researchers on a successive concept for operational control and managerial oversight in safety critical systems: the resilience engineering concept.

2.4.4   The fallacy of lack of foresight and management control

Claiming a role for resilience engineering

In his theory, James Reason shifts stability of systems from the individual operator level to the organisational level. Such a stability is shifting from individual control to organisational and hence, managerial control. As stated by Hollnagel (2011), individuals have a natural and uncontrollable variance in behaviour, restricting the ability of higher management order to control individual behaviour as compliant to their desired/imagined pattern.

Resilience is discriminating between organisational control and predetermination of planned tasks and procedures. While organisational control deals with variety in performance (As Done), predetermination is controlled by the specifics of task and mission characteristics (As Imagined). The nature and imagined behaviour of the systems is determined by both complexity/coupling and legacy/change rate of its technology.

Discrepancies and anomalies between performance and the potential role as 'early warning' signalling expert as Imagined and as Done are either intentional deviations -stigmatized as 'violations' from rules and regulations- or unintentional -triggered by internal patterns of slips, lashes or mistakes-. Reason developed a generic and normative categorization of human error, based on individual characteristics (Generic Error Modelling System, GEMS).

In aviation, anomaly management occurs on an organisational level: compliance with predefined performance is organised by compliance to drafting a flight plan, pre-flight preparation and in-flight responses based on scenarios and Standard Operating Procedures. Various modes of operations are foreseen, based on the specifics of flight phases, as the ability to switch between operational modes and balancing stability and manoeuvrability, while maintaining flexibility and adaptivity to variety in cultural and conditional aspects. Operational excellence can be achieved by organisational robustness and managerial control (Winters 2017).

Table 1 Organisational and technological control

	Low technological control	High technological control
High organisational control	Fire fighting Medicine	Aviation Nuclear power plants Process industry
Low organisational control	Fishing industry	ICT

There is an increasing role for resilience moving from high organisational control and high predetermination to low organisational control and predetermination, with a shift from proactive to reactive interventions.

Such characterizing of systems by their legacy, high tech nature, change rate and complex/coupled properties identify strategic choices that have to be made in controlling modes of operations of systems: do we select organisational resilience instead of technological resilience (Zimmermann et.al. 2011)? Can we rely on collecting precursor data of what went right as 'proactive' -and consequently superior-instead of investigating what went wrong as a 'reactive' reduction of uncertainty. Or do we need both to comply with the full information paradigm (Klir 1987, 1994)?

Resilience engineering revisited

With the distinction between organisational control and technological control, Zimmermann et.al. (2011) suggest a dilemma in choosing either one of them as the exclusive approach. Such a dilemma however, does not comply with the evolution of a socio-technical nature, such as aviation. Reluctance to accept resilience engineering as the new way forward did pose the question: is the aviation industry ready for resilience (Zimmermann et.al. 2011)? The other question: is resilience engineering ready for the aviation industry as a legacy systems of a Non-Plus Ultra-Safe nature, is as appropriate.

Woods identified several initial fundamentals for resilience engineering from a sociological perspective (Woods 1996). In his inquiry to make progress in resilience engineering thinking, he identifies two additional fundamentals: initiative and reciprocity (Woods 2016, 2019).

In overcoming brittleness in complex systems, he heavily leans on engineering design principles that are basic knowledge in aerospace engineering, in particular the principles of operating envelope and graceful degradation. He coins the fundamentals of a new notion of 'graceful extensibility' to cope with inherent variability and surprise events in a continuous changing world (Woods 2019). In coping with immanent failure, he turns to exploring the design of governance mechanisms and system architecture in order to control long term performance of complex systems, facing multiple cycles of change. In his exploration of initiative and reciprocity, a specific role for communication and interaction across system life phases and system states emerges. Feedback from operational experience to planning and design could be re-integrated in such systems design. This could provide a timely interference with actual system performance, based on foresight and proactiveness. Implicitly, Woods introduces the principle of Good Airmanship for all industrial sectors. Explicitly he acknowledges the value of complementarity across engineering, biological, social and cognitive sciences. This creates opportunities for new thinking of systems design and operations, combining socio-psychological notions with engineering design methodologies. Optimization and control strategies could be developed from an integral systems perspective. In such a perspective, there is ample room for disruptive and innovative thinking, necessitated by changes in environment, economy and risk perception.

2.5   Selecting strategic options

Creating a disruptive change should comply with both economic and technical developments in complex systems as the new context for developing intellectual constructs on dynamic systems behaviour, mobilizing new domains and disciplines. Selecting either organisational or technological change is dictated by the sector and its inherent technology to avoid a drift into chaotic systems.

2.5.1   Economic developments

It is doubtful whether there is a 'natural' tendency to increase production when things are going right. Trade-offs between efficiency and thoroughness in a traditional economy are frequently conducted at the expense of safety. Such trade-offs are realised by increasing flexibility and organisational resilience. Eventually, new opportunities are being created in a New Economy market model for aviation. Subject matter expert(ise) frequently plays the role of whistleblowing in such situations. They are labelled also frequently as 'resistance to change' or 'unconscious cognitive stubbornness' in objecting such change (De Boer 2012). Resistance to change and unconscious cognitive stubbornness may have a positive or negative effect on performance. On one hand they may block sharing of mental models in a team, hindering a shared understanding of the situation. They may create a cognitive lockup in supervisory control tasks, change blindness, cognitive mismatch, fixation and eventually may create accidents in dealing with contradicting signals. On the other hand, they may stimulate vigilance, danger avoidance, stimuli detection and rapid reflection on immanent situations. They may induce less automatic, intuitive behaviour and enhance analytic competences. These notions are considered instrumental attributes of Good Airmanship and Good Seamanship.

New Economy models focus on lean efficient production, eliminating superfluous costs and waste. They do not take into account the consequences of reductions in training costs and subsequent, decay of proficiency and basic flying skills of pilots, as demonstrated by the AF447 disaster.

With respect to economic developments and models, Minsky identified four different phases of driving forces for business models at a macroscopic level of economy (Minsky 1986). :

- Optimizing expectations on a short term with operational trade-offs at a corporate level
- Speculative extrapolations of these expectations in a seemingly stable situation
- Profit expectations on a long term despite stalling investments and erosion of precautionary measures
- Innovative powers of disruptive solutions and creative destruction of old concepts, disclosure of new markets, substantiated by research and development investments to achieve value preservation.

Such disruptive innovations are supported by disclosure of unchartered scientific domains and a new interdisciplinary cooperation (Woods et.al. 2016, Woods 2019, Stoop 2017/3).

In particular in the domain of human decision making, the work of Slovic (2004) on emotions and empathy, Kahneman (2013) on cognition, intuition and perception and Taleb (2008) on rare events and after the fact explanations have drawn attention in the safety science community.

Over the past decade, circular economy principles have been developed. In the environment, zero emission, recycling and closing circular chains are advocated. Sustainability requirements lead to disruptive technologies, new business models and entrepreneurial competences (Berkhout 2000). Systems should be intrinsically safe, while safety is considered a strategic asset in the value chain. Such changes in the socio-economic environment also require disruptive changes in safety thinking and scientific interests. Traditional scientific constructs may run short in explanatory potential (Stoop 2017/2).

According to Troadec, the chairman of the French safety investigation authority BEA, based on the experiences of the Air France AF447 accident, only flight recorder retrieval clarified operating circumstances. Combination of ergonomics of warning designs, training conditions and recurrent training processes DID NOT generate expected behaviour, showing limits of current safety models of human behaviour (Troadec 2013). The AF447 case triggered new and unchartered scientific interests in the man-machine interaction domain, focusing on non-normal situations, intuition, habituation and exploration of the 'startle' effect (Mohrmann et.al. 2015).

2.5.2 Technological developments

With respect to developments in aviation in 1949 at the foundation of ICAO, technology was chosen as the flywheel for progress (Freer 1986). Technological flexibility, variability and technical adaptation were chosen as the prime system change agents (Vincenti 1990) under conditions of tight coupling to an international institutional framework of ICAO standards and operational practices. This choice was evident: during the negotiations at Yalta in 1945 between Roosevelt, Stalin and Churchill on the progress of aviation after the ending of the Second World War, none of the participants was willing to grant primacy to another economic system than their own, being either a Capitalist, Communist or Commonwealth model. The only alternative was to agree on a technological harmonisation and standardization for reasons of interoperability and accessibility of the international aviation network (Freer 1986). Such a flywheel function was

readily available for technology after the world war due to the huge R&D and production potential in the aviation industry in the USA, UK and Soviet Union.

This selection of technology as the flywheel for progress demands a very high organisational continuity and stability at the corporate level to introduce a high level performance (safety) playing field. Harmonization was achieved by introducing certification and supranational standardization such as at the sectoral level ICAO Annexes structure for all primary systems functionalities. The role of the State as the prime mover for change was selected as the natural entity for imposing legislation and enforcement on their State owned carriers.

Simultaneously, a very high technological flexibility to adapt to new developments and operational conditions, constraints and specificity was required, introducing a rapid technological development in the context of private corporations, stimulating competition and innovative exploration.

As a consequence, a combination of high technological flexibility –nowadays indicated as 'unruly technology' and low organisational or individual flexibility – nowadays labelled as 'resistance to change' and 'cognitive stubbornness' – are two complementary notions that in conjunction enable both flexibility and reduction of uncertainty in acceptance of technological innovations. The fading role of the State as the leading entity in this development process and the merging of a multitude of aircraft manufacturers into a limited number of leading global companies has called for reflection on the future of aviation. Tensions have arisen with respect to the pace and rate of innovation and organisational adaptation in adapting to new global economic, market and environmental developments. Programmes like Horizon 2050 have been created, facilitating innovative research and development programmes on a sectoral level.

Advocating resilience engineering has not been embraced by the aviation community (Zimmermann 2011). Resistance to organisational change, cognitive stubbornness and underspecification of technological development have been noticed as obstacles for such an acceptance. These phenomena however are functional and complementary in making progress under conditions of minimizing uncertainty. Unruliness is a precondition for technological adaptation and innovation. This property of technology has been recognized already in 1949 with the foundation of ICAO. It has been described by Vincenti as a basic property of aerospace engineering design (Vincenti 1990). To reduce uncertainties in this technological developments and to guarantee a safe operational performance, an

elaborated system has been developed, linking the various phases of the system life cycle. Exchange of knowledge and experience is established by an international agreed system of certification and licensing by modelling, simulation, testing, training and investigations. In due course, the scope expanded from aircraft airworthiness criteria to flight envelope and system viability criteria (Stoop 2017.2).

2.6   Vincenti: the variation selection model

Specifications and regulations are considered properties of control mechanisms at a sectoral high performance level that enable progress. Simultaneously, they create resistance to organisational change and facilitate underspecification of technological development to enable deviation and adaptation. At a sectoral level, harmonization and standardization and sharing design and operational experiences and knowledge are prerequisites to implement this philosophy of technological progress. Foresight on operational behaviour of innovative and disruptive solutions is established by a sophisticated framework of Annexes to the ICAO Agreement by certification, testing and training.

In his analytical study on aerospace engineering methodology, Vincenti indicates the transition from craftsman thinking in experimental progression towards knowledge based design of artefacts and evidence based learning (Vincenti 1990). In the 1930's the empirical and experimental design of aerofoils was gradually replaced by analytical and mathematical understanding of the mechanisms that ruled aerofoil design. Such transition from scientific theory and aerodynamic models as developed by Bernouilli, Navier Stokes, Mach, Schlichting and others towards a knowledge-based design was supported by wind tunnel testing of scale models and flight tests. Scientific research focused on the role of viscosity, transition between laminar and turbulent flow, laminar flow aerofoils and elliptic lift distribution. This application of scientific research in order to reduce uncertainty in the attempts to achieve increased performance created a growth in knowledge. This knowledge was applied directly in the design of new combat aircraft. The British Supermarine Spitfire was designed based on elliptical lift distribution on its wings. The US North American Mustang was designed based on the laminar flow characteristic of its aerofoils. Both aircraft represent a leap in aerodynamic performance. The German Messerschmitt Me 262 marked the transition from piston engine powered to the fighter jet age.

Many technological innovations became available for civil aviation applications in the desire to expand civil aviation to a global network after the war and beyond. In the fourth generation of fighters, the application of IT controlled thrust vectoring enabled the Russian Sukhoi SU-35 to perform the Puchachev Cobra manoeuvre.

2.6.1   Presumptive anomalies

Increased knowledge in turn acts as a driving force to further increase knowledge. As defined by Constant (quote by Vincenti 1990) the phenomenon of 'presumptive anomaly' may stimulate better understanding of the behaviour of an artefact:

"Presumptive anomaly occurs in technology, not when the conventional system fails in any absolute or objective sense, but when assumptions derived from science indicate either that under some future conditions the conventional system will fail (or function badly) or that a radically different system will do a much better job."

Vincenti concludes that presumptive anomaly, functional failure and the need to reduce uncertainty in design act as driving forces to a growth of engineering design knowledge.

Challenging design assumptions, model simplifications and operational restrictions in examining the validity of this knowledge store have contributed to the growth of design knowledge. Through safety investigations, systemic and knowledge deficiencies were identified, leading to novel safety principles in engineering design. Eventually, this has led to Knowledge Based Engineering and Multidisciplinary Design Optimization as a specific school of aeronautical design thinking (Landman 2010, Torenbeek 2013, Van Tooren 2003).

The search for performance optimization and reduction of uncertainties has created a continuous exploration of design variations and selection of better performing design solutions. This has created generations of commercial and military aircraft designs with similar morphology, configurations and properties. Such solutions can either have a derivative or disruptive nature. Vincenti elaborates on the role of this variation-selection process in the innovation of aerospace design (Vincenti 1990). Developing 'anomalies' should be considered in a historical context of design requirements, gradual changes in the operating context and consequences of design trade-offs.

Although 'anomalies' may temporarily deviate from prevailing engineering judgement, specific concerns may force to deviate from this mainstream in exploring innovations.

Foresight on performance has been both tested at the component and subsystem level prospectively by modelling and simulation and retrospectively by flight testing and operational feedback. Such 'unforesightedness' comes with balancing gains as well as costs. The outcomes of such a balancing may favour specific design trade-offs, but should be considered in their historical context and operational demands. As speed increased, drag became dominant in the design trade-offs in designing retractable gears. The generalized knowledge that retractable gears were favourable, was the product of an unforesighted variation-selection process and was valid for a specific class of aircraft designs (Vincenti 1990). Similar trade-offs in context can be observed in the design of modern commercial aircraft in balancing weight and fuel consumption versus structural integrity and dynamic stability (Torenbeek 2013).

Flight envelope protection was introduced to refrain the pilot from entering the margins of the operational envelope (De Kroes and Stoop 2012). The application of automation in cockpits has a proven track record of substantial gains in safety, efficiency and accuracy, but comes at a cost of loss of pilot situation awareness in critical situations, increased cognitive task loads and loss of basic flying skills. In aviation, the notion of 'unforesightedness' due to trade-offs has been acknowledged on both the component and the systems level.

Warnings against costs in trade-offs in design and operations requires subject matter expertise: an understanding of the relations between technological and socio-organisational aspects is indispensable. Otherwise, an undefined and compiled notion of 'complexity' is generated to disguise the ignorance of understanding 'emergent' properties –as defined by Rasmussen- and dynamics of 'complex systems with tight couplings' –as defined by Perrow-, which might - according to Turner- 'drift into failure' due to 'human error' –as defined by Reason- . In such a combination of undefined notions, the ability of 'foresight' is easily lost, in particular when analysing design trade-offs and feedback from reality by safety investigations have been dismissed from the diagnostic toolkit. Losing specific and context dependent knowledge in safety critical situations resulted in loss of understanding why in a specific case an accident could occur. By losing oversight over the nature of a triggering event, remedial control options are lost as well.

All this occurs in Non-Plus Ultra-Safe systems where Vision Zero has been achieved for the first time ever in large commercial aviation due to the fact that in 2017 no fatalities occurred. Such an achievement has crossed the – according to Amalberti- 'mythical barrier' of the 10-7 , falsifying the assumed asymptotic nature of safety performance and all their derivative assumptions in such systems (Amalberti 2001). It also questions the ambitions of human behavioural sciences to serve as a promising and needed 'antithesis' for a 'conventional' technological perspective. According to Troadec and Arslanian of the French BEA on the AF 447 case, factual evidence in air safety investigation experiences have demonstrated limitations of present human performance scientific thinking.

Rather than challenging the interpretation of various scientific schools of thinking, a descriptive diagnosis of the nature and dynamics of complex systems should provide insight in their architecture and developments towards a next safety integrity level. Unravelling rather than accepting their complexity becomes of prime importance for achieving Vision Zero and First Time Right principles in a safety for design approach (Stoop 1990).

2.6.2   Complexity, a social construct
In order to control the complexity of this development process, a distinction is necessary between structural complexity (single functional structures) for the benefit of flexibility and functional complexity (multifunctional structures) to enable adaptation. Since combining both types of complexity creates uncontrollable uncertainty in performance variations, limiting any trustworthy foresight of intended behaviour, such a combination of complexities can only be combined in one design at a high cost of increased uncertainty and reduced controllability. A choice should be made for either structural complexity or functional complexity. Additional design properties such as robustness, redundancy, reliability and resilience of technical artefacts to reduce the uncertainty in the design, should be guaranteed throughout the design process and operational life. To the purpose of foresight in aviation safety, various safety design principles were derived from theoretical notions, experimental design evaluations and safety investigations: fail safe, safe life, damage tolerance, crash worthiness, graceful degradation, self-relianceness, situation and mode awareness. In this process, the role of accident investigations and forensic engineering to disclose failure cannot be underestimated (Petroski 1992, Noon 1992, Carper 2001, Barnett 2001, Arslanian 2011).

This design philosophy of preferring technology as the flywheel for progress has been specific for the aviation industry. In the process industry the choices have been different: technology was chosen as a constant, while organisational variety and change was selected as the engine for change for multinational companies without State interference. The context of multinational corporations in a different socio-economic competitive and political climate with the state of technology and its inherent maturity level, differs from the aviation industry. In aviation international cooperation, interoperability, accessibility of global networks and a harmonized and standardized high performance level playing field prevail. This sector has seen the development of several generations of aircraft of similar configuration, performance and operating envelopes.

In the process industry technological development has been different across multinational companies, each with their specific organisational constitution and structures (De Rademaeker et.al. 2014, Pasman and Reniers 2014, Lannoy 2016). Safety is embedded in the organisation rather than in its technology, differentiating between line or staff responsibilities, creating tensions between subject matter expertise resources, foresight capabilities and operational feedback responsibilities. Such differences raise questions about the role of technology, its variability, unruliness and physical boundaries of its production principles and operational processes. But above all, such a choice for organisational change raises questions about the control over organisational change and technological change and vice versa, by either subject matter experts, corporate management, national public governance or supranational institutions. This dilemma between technological and organisational has created a specific role for whistle blowers in an organisation and a choice between top-down or bottom-up initiation of change, including the power relations in an organisation. In the engineering design community, the role of designers and technical experts is quite different, where the role of change agent is fulfilled by inventions and disruptive changes according to the theory of Vincenti's on presumptive anomalies and the variation-selection model. In aviation, the role of pilots as the delegated and distributed responsible expert operators is established by the notion of Good Airmanship, providing feedback from reality while safety investigations provide feedback form anomalies, deficiencies and failure.

2.7 Foresight and whistle blowers, an analysis

2.7.1 Some observations

In describing the development of safety in the aviation sector, technology has been chosen as the flywheel for progress, keeping organisational and institutional arrangements constant. The way technological design alternatives were developed and selected has gone through a process of variation-selection, testing 'anomalies' on their trade-offs by feedback from reality. In such a validation process, a distinct role has been allocated to safety investigations to provide evidence of the system's functioning under normal and non-normal conditions. Reducing uncertainty and variance in operator behaviour has been covered by the notion of Good Airmanship, covering delegated and distributed responsibilities between corporate and individual performance in the global network. The role of design in reducing uncertainty has evolved towards Knowledge Based Design and Value Engineering paradigms.

In order to decide on the consequences, feasibility and acceptability regarding the safety properties of derivative or disruptive solutions, new safety notions have to be developed.

Resilience engineering has presented itself as a serious prospect candidate.

However, 'old school' safety notions, such as human error, drift into failure and normal accidents have dominated the debate over the past decades, accompanied by mathematical and quantitative assessment of risk. Such old school notions have been challenged by sociological theories about 'complexity' and 'unforesightedness', popularized by notions such as 'unknown unknowns' and 'unpredictability'. Such a reductionist approach from a socio-organisational perspective does not pay credit to technological and systemic analytic potential that is practically available in the engineering design community. The link between technology and organisation as two primary and mutually independent characteristics is yet to be reinstalled by adhering to a socio-technical systems approach, acknowledging both hierarchical and network characteristics (Woods 2016, Boosten 2017). In such an approach, both operational performance, organisational arrangements and institutional conditions have to be taken into account for the sake of innovation (Berkhout 2000).

2.7.2   Analysis of assumptions

In proclaiming fundamental shifts in dealing with human behavior, Rasmussen disconnects design from operations, eliminating the feedback and feed forward relations between these two life cycle phases of systems. He replaces a design orientation with an operational orientation, controlled exclusively by corporate management. Systems performance is only to be discovered by deviations from normal and intended performance during operations through 'emergent' behavior. The role of the State is reduced to providing performance standards, criteria and limits. In his construct, there is no room for accident investigations. Minor accidents are considered statistical aberrations from normal, while major accidents are unique events, beyond control and learning. Eliminating safety investigations -providing operational transparency and knowledge on a system's life time behavior- reduces a control perspective from dealing with cause to only dealing with consequences after the release of a hazard. As stated by Perrow (1999), consequences are consequently assumed to be 'normal' to any complex system behavior. By taking this perspective, Rasmussen reduces safety from a sectoral strategic value to a corporate operational constraint. Rasmussen also applies a different definition of 'systems'. In this construct, systems are considered open horizontally organized networks, while in the engineering perspective, a hierarchical dimension prevails at the sectoral control mechanism with a distributed allocation of responsibilities and control mechanisms. Such an engineering perspective does not notice a paradox in almost perfectly safe systems as proclaimed by Amalberti (Amalberti 2001). The reductionist perspective of Rasmussen on systems as horizontal networks and his restriction to a corporate level opens up debates on discrepancies between Work As Imagined (by management) and Work as Done (by operators). This perspective leaves out the assumptions as formulated during the design and the development of the system itself.

The shift in perspective as proclaimed by Reason and Rasmussen also rejected the tools and techniques that were readily available in the engineering domain. Rasmussen suggests to replace the engineering toolkit by tools and techniques from the mathematical domain –QRA in particular-. In validating the applicability of QRA to this managerial construct, frequent criticisms on their assumptions and limitations have been formulated by the QRA and resilience engineering community (Aven 2016). Over time, Reason's and Rasmussen's assumptions proved to be inadequate: later versions of human behavior reinstalled an interest in operational feedback from incidents, Just Culture and High Reliability Organization behavior. A shared responsibility between management and operators is proclaimed, introducing the notion of 'mindfulness' with allocation of a prime responsibility to the shop floor level of performance. The 'ethical accounting' as defined by Rasmussen introduces the phenomenon of Whistleblower. Any impact of decisions on the objectives and values is inevitably normative: they are either undesirable and non-compliant with established ethics in an organization –defining a negative connotation for a whistleblower- or are the ethical responsibility of a corporate employee, -defining individual mindfulness- and contribution to a 'Risk-of Life' assessment of risk.

Accepting any of the newly proposed paradigms as successor of 'old school' – including their obsolete- notions should be accompanied by an assessment of residual risks and side effects.

Such acceptance should not be restricted to the individual level of 'whistle blower' functionality. At the institutional level, safety investigations by independent agencies have seen a global development in the aviation sector. It is a part of the legacy of aviation, supported by forensic engineering and governance as distinct scientific disciplines (Stoop and Dekker 2010).

In its efforts to enhance safety in aviation further, ICAO has drawn up a set of management processes based on the theories of Reason and Rasmussen, that could be adopted by corporate management (SMS) and state safety programs (SSP). This initiative was not to suggest to exclude, discount or downplay other preventive activities. An empirical analysis of Farrier (2017) showed an unintended outcome of the role of safety investigations in aviation. It has become clear that ICAO's various moves to consolidate guidance on SSPs has been to downplay the role of accident investigations in the SMS environment, or even to disconnect them entirely from other preventive processes.

The trend seem to be to discount investigations as a part of the larger preventive process. He concludes that for a variety of reasons, investigations -including their recommendations that result from them- are not always a good fit with each other. Farrier notices inherent tensions between the two philosophies: a focus on what might happen versus what has happened: a desire to consider hazards in the abstract instead of focusing on concrete experiences of actual loss. A focus on Hazards as Imagined versus Hazards as Experienced shift the attention from perceived issues to 'precursors', expecting a higher added value of the latter. A

downplaying of 'reactive' investigations takes place against support for 'proactive' management efforts. In practice, the underlying philosophy of Reason and Rasmussen supports the notion that 'safety culture' has preventive value and costs a lot less than investigations and design based safety impact assessments. Farrier concludes: accident investigations and their recommendations need to be properly baked into the fabric of current and future safety management systems. 'Proactive' outcomes need not be pursued exclusively through 'proactive' sources of data. Safety investigations should form the basis for follow-up inquiries and analysis, while their recommendations should be scrupulously tracked and managed. This is in accordance with the principle of the Full Information Paradigm (Klir 1987) that feedback and feed forward should be combined to achieve full information on systems behaviour.

Farrier states (2017): Introducing new concepts such as Safety Management Systems and State Safety Programs puts two principles of safety management and safety investigations in opposition instead of leveraging their respective advantages. Such opposition pits the active against the passive, the hard work of investigation and analysis against the easy tasks of collecting and recording. Both have their place in the aviation safety professionals' toolkit, and neither should be disregarded or discounted (Farrier 2017). In discarding safety investigations from the analytical toolkit, such investigations are expelled from foresight from within a system and forced into a role of adversary whistle blowers (Vuorio et.al. 2017, Wilson and Straker 2018).

Such an exorcizing also has consequences on the investigative functionality of the capability to change systems. This introduces two problems (Karanikas, Roelen and Piric 2018).

First, the interpretation of investigative findings is submitted to differences in perspectives between investigators and safety managers. Investigation reports are consensus documents on the investigative reconstruction of an event. A transition from what happened to how to deal with the consequences has to take place by analytic interpretation and adaptive intervention on those investigative findings. Such a transition depends on the capabilities, responsibilities, resources, response capabilities and intervention strategies of each of the stakeholders in the safety enhancement process.

Second, such an intervention strategy lacks procedures for transforming drafting investigative recommendations into incorporating these findings in a safety management system in a specific corporate, stakeholders and governance context.

2.8   Discussion

In making an inventory and analysis of the role of whistle blowers in foresight various scientific opinions about uncertainty, variety, flexibility and controllability emerge. There are different perspective with respect to how to maintain control over complexity of socio-technical systems.

2.8.1   Old school of thinking

'Old school' human factor thinking has become deficient: it contains a normative opinion about human performance due to 'human error', has little predictive potential due to an unnoticed 'drift into failure' and has no control over consequences due to 'normal accidents'. There is a wilful decline of cause in favour of consequences. The rejection of accident investigations as a source of information deprives the concept from operational feedback. Foresight should be provided by incidents, early warnings and whistle blowers. While a first version claims managerial control responsibility over the system performance, a revisited version shifts responsibilities back to front line operators and designers, demanding a permanent awareness and mindfulness to predict, to cope and to anticipate disaster. Managerial responsibilities are reduced to only 'conditional' factors, replacing safety as a sectoral, strategic value.

Warnings against costs in trade-offs in design and operations requires subject matter expertise: an understanding of the relations between technological and socio-organisational aspects is indispensable. Otherwise, an undefined and compiled notion of 'complexity' is generated to disguise the ignorance of understanding 'emergent' properties –as defined by Rasmussen- and dynamics of 'complex systems with tight couplings' –as defined by Perrow-, which might - according to Turner- 'drift into failure' due to 'human error' –as defined by Reason- . On the instrumental level, safety oversight was replaced by a Safety Case approach as the coping mechanism for management over emerging risks.

In their claim for exclusive control over organisational safety performance, oversight was replaced by a shared but undefined responsibility to evade liability and accountability (Koivisto et.al. 2009)

In such a combination of undefined notions, the ability of 'foresight' is easily lost, in particular when analysing design trade-offs and feedback from reality by safety investigations have been dismissed from the diagnostic toolkit. Such reframing of the theoretical concept of foresight and uncertainty in socio-organisational terms fitted in quite well with the New Economy principles that were favoured in the 1990's by the UK government (Martin 201, Stoop 2017.2). Consequently, operational safety feedback warning systems -such as Good Airmanship and Seamanship in aviation and maritime- could not be reconciled with such exclusive corporate management responsibilities. For subject matter experts, an antagonistic role emerged as a Whistle Blower for early warnings of immanent systemic mishaps.

Losing specific and context dependent knowledge in safety critical situations on the operational level resulted in loss of understanding why in a specific case an accident could occur. By losing oversight over the nature of a triggering event, remedial control options are lost as well.

2.8.2   New school of thinking

The 'new school' thinking in human factors claims a necessary paradigm shift but still focuses primarily on consequences instead of causes, on operations instead of design and prefers to analyse the positive rather than the negative. They apply a multi-actor approach and take a non-normative perspective in an operational environment. The emphasis is on organisational flexibility, irresponsive of technology, legacy and socio-economic context of the systems under scrutiny. This school does not (yet) reinstall a highly necessary relation with technology, engineering design and system theory as fundamental characteristics of socio-technical systems.

Consequently, their plea for adhering to resilience may favour recognition of their discipline and perspective, but may not fulfil the needs of highly elaborated and matured industrial legacy sectors such as aviation.

The needs of such sectors and systems are dictated by their specifics and operational context. Rules of a higher hierarchical order, economic market mechanisms and control strategies at the level of system architecture and configuration, public governance, economic and business models and social culture prevail.

Advocating resilience without taking into account such a context and hierarchy may even jeopardize the goals of such legacy systems in disregarding strategic decisions and choices made in the past. These strategic decisions have been successfully applied in aviation by avoiding slipping complex systems into chaotic states, achieving an unprecedented non-plus ultra-safe performance level. In aviation, a high organisational and institutional stability combined with technological change has accommodated permanent economic growth, adaptation to new business requirements and societal constraints. Institutional arrangements were made at the sectoral level such as establishing ICAO. Selecting technology as the flywheel of progress and independent investigations at a State level proved feedback from reality, combined with delegated responsibilities to the cockpit crew by Good Airmanship principles (McCall 2017). Such Good Airmanship principles are derived from the maritime history, where a very open operating environment forced to comply with Good Seamanship and standardized operating procedures to survive unexpected situations. The very high rate of diversity and open operating environment did not allow room for variation and interpretation, but adaptation to the margins of a physical operating envelope in non-normal situations.

2.9   Conclusion

The debate about applicability of resilience has brought about the need to have foresight on future performance and stability in a dynamic operational environment, relative to the 'old school' of safety thinking (Zimmermann et.al. 2011). For aviation however, there is no need per se for proselytizing to a new belief of Resilience. A more pragmatic approach of incorporating useful notions in the needs of the sector prevail. Taking both technological and organisational variety on board may jeopardize the overall stability of the sector and threaten the present non-plus ultra-safety performance of the sector. Creative destruction of old paradigms is a necessary step towards innovation but is a serious risk to the sector in a phase of expanding capacity and growth combined with developing into a next generation of aircraft, airports and traffic management systems. There is no 'natural tendency' towards increased productivity, but as Minsky has shown, socio-economical market mechanisms and societal developments of a higher order dictate change. A chaotic system may emerge from such uncontrolled series of changes if technological and organisational configurations are made flexible

simultaneously. Such a transition from complex to chaotic also changes the role of subject matter experts as professional safety assessors into whistle blowers and restricts the ability to incorporate their foresight during such changes. Although Reason and Rasmussen have revised their concepts, fundamental deficiencies have not been addressed (Reason 2015). Advocating new concepts as such to accommodate the enhancement of safety during changes is a valuable and necessary plea. There are some valid hesitations in the aviation community to embrace resilience engineering. Aviation may 'drift into failure' by disrupting the architecture of the sector too much, not only on the safety aspects. Releasing organisational variety may cause stagnation of technological innovation by emphasizing legal liability and accountability to failure as unforeseen consequences. Recent British legislation on Corporate Manslaughter and Corporate Homicide has aggravated the legal liability situation after the Concorde crash in July 2000 by introducing massive repercussions for manufacturers after failure of their products. There is an unexplored relation between Resilience and Safety. In such a context, the role and position of whistle blowers is undefined.

Foresight is about reducing uncertainty and predicting future performance. New approaches, theories and notions are still open and their desirability, feasibility and applicability is still undetermined. The future role of the State, increase in automation, security, sustainability and circular economy principles are not yet fully explored, let alone validated regarding their consequences. There are no Golden Bullets in enhancing safety in such developments.

Revising resilience engineering by adding two fundamentals -initiative and reciprocity- may create a basis for cross-disciplinary participation, communication and commitment. This could make the outsiders role of whistle blowers obsolete and could reinstall their role as subject matter experts from within the system. Such a transition poses challenges on creating a shared repository of expertise, experiences and knowledge management, combining feedback and feed forward loops to design and operations of complex systems. As such, it may benefit foresight in safety by identifying early warnings of system degradation.

2.10   References

Almklov P., Rosness R. and Storkersen K., 2014. When safety science meets the practitioners: Does safety science contribute to marginalization of practical knowledge? Safety Science 67 (2014) 25-36

Amalberti R., 2001. The paradoxes of almost totally safe transportation systems. Safety Science 37, p 109-126

Arslanian P.L., 2011. Acknowledgement speech ISASI's Jerome Lederer Award 2011. ISASI Forum, October-December 2011 p 12-13

Aven T., 2016. Risk assessment and risk management: Review of recent advances on their foundation. European Journal of Operational research 253 (2016) 1-13

Barnett P., 2001. Ethics in Forensic Science. Professional Standards for the Practice of Criminalistics. Protocols in Forensic Science Series. CRC Press

Berkhout G., 2000. The Dynamic Role of Knowledge in Innovation. The Netherlands Research School for Transport, Infrastructure and Logistics TRAIL. June 2000

Boosten G., 2017. The (Congested) City in the Sky. The capacity game: finding ways to unlock aviation capacity. Inaugural Lecture October 2017. Amsterdam University of Applied Sciences. Aviation Academy.

Booth R., 1979. Safety: too important a matter to be left to the engineers? Inaugural lecture Aston University, 22 February 1979.

Çarper K., 2001. Forensic Engineering. Second Edition. CRC Press LLC

CASV, 2017. Civil Aviation Safety Review 2017.

De Boer R.J., 2012. Seneca's error: An Affective Model of Cognitive Resistance. Doctoral Thesis, Delft University of Technology, 7th May 2012.

De Kroes J. and Stoop J., 2012. Stall shield devices, an innovative approach to stall prevention? Air Transport and Operations Symposium 18-20 June 2012, Delft University of Technology

De Rademaeker E., Pasman H. and Fabiano B., 2014. A review from the past, present and future of the European Loss Prevention and Safety Promotion in the Process Industry. Process Safety and Environmental protection, July 2014

Dekker S., 2011. Drift into Failure. From Hunting Broken Components to Understanding Complex Systems. Ashgate Publishers

Dekker S. and Pruchnicki S., 2013. Drifting into failure: theorizing the dynamics of disaster incubation. Theoretical Issues in Ergonomic Sciences, 2013. http;//dx.doi.org/10.1080/1463922X.2013.856495

Edwards E., 1972. Man and Machine: Systems for Safety. Proc. British Airline Pilots Association Technical Symposium, British Airline Pilot Association, London, 21-36

Farrier T.A., 2017. Investigations, recommendations and safety management systems. MO3763 Senior Safety Analyst JMA Solutions, LLC, ISASI Forum June 29, 2017

Freer D., 1986. Chicago Conference 1944. Special Series ICAO Bulletin 41. http://www.icao. Int/publications/Pages/ICAO-Journal aspx

Hale A., 2006. Method in your madness: System in your safety. Valedictory lecture Andrew Hale, Delft University of technology 15 September 2006.

Harvey C. and Stanton N., 2014. Safety in System-of Systems: ten key challenges. Safety Science 70 (2014) 358-366

Holling C., 1973. Resilience and stability of ecological systems. Annual Review of Ecology and Systematics. Vol 4 (1973) p 1-23

Hollnagel E., 2011. Resilience Engineering in Practice. A Guidebook, edited by Erik Hollnagel, Jean Paries, David Woods and John Wreathall. Ashgate Studies in resilience Engineering

Kahneman D., 2013. Thinking, Fast and Slow. Farrar, Straus & Giroux Inc

Karanikas N., Roelen A. and Piric S., 2018. Design, scope and focus of safety recommendations: results from aviation safety investigations. Policy and Practice in Health and Safety. ISSN 1477-3996 (Print) 1477-4003 http://www.tandfonline.com/loi/tphs20

Klir G., 1987. The role of methodological paradigms in system design. Dept. of Science. Thomas J. Watson School of Engineering. State University of New York at Binghamton. New York

Klir G., 1994. On the Alleged Superiority of Probabilistic Representation of Uncertainty. IEEE transactions of fuzzy systems. Vol 2, no 1 Febr. 1994

Koivisto R., Wesberg N., Eerola A., Ahlqvist T., Kivisaari S., Myllyoja J. and Halonen M., 2009. Integrating future-oriented technology analysis and risk assessment methodologies. Tchnological Forecasting & Social Change 76 (2009) 1163-1176

Landman Q., 2010. Risk assessment in preliminary Aircraft Design Using Bayesian Networks and Knowledge Based Engineering. Literature report. Delft University of Technology

Lannoy A., 2016. Risk Management, safety and dependability: looking back from 1990 to 2015, which future? 51st ESReDA Seminar on Maintenance and Life Cycle Assessment of Structures and Industrial Systems. October 20-21, 2016, Clermont-Ferrand, France

Martin B. 2010. The origins of the concept of 'foresight' in science and technology: An insider's perspective. Technological Forecasting & Social Change 77 (2010) 1438-1447

McCall J., 2017. Modern day heroes: a multiple case study of how successful flight crew and air traffic control coordination helped prevent disaster. International Journal of Current Research Vol 9, Issue 11, pp. 61268-61275, November 2017

Minsky H., 1986. Stabilizing an Unstable Economy. McGraw-Hill Companies

Mohrmann F., Lemmers A. and Stoop J., 2015. Investigating Flight crew Recovery capabilities from System Failures in Highly Automated Fourth Generation Aircraft. Journal for Aviation Psychology and Applied Human factors.

Noon R., 1992. Introduction to Forensic Engineering. The forensic library. CRC Press inc.

Pasman H. and Reniers G., 2014. Past, present and future of Quantitative Risk Assessment (QRA) and the incentive it obtained from Land-Use Planning (LUP). Journal of Loss prevention in the Process Industry 28 (2014) 209.

Perrow C., 1999. Normal Accidents. Living with High Risk Technologies. Princeton University Press.

Petroski H., 1992. To Engineer is Human. The Role of Failure in Successful Design. Vintage Books

Rasmussen J., 1997. Risk management in a dynamic society: a modelling problem Safety science Vol 27, No2/3, pp 183-213, 1997

Reason J., 2015. Organizational Accidents Revisited. CRC Press Book

Roed-Larsen S. and Stoop J., 2017. Uncertain future. Unsafe future? 53th ESReDA Seminar. Enhancing Safety: the Challenge of Foresight. November 14 - 15, European Joint Research Centre, Ispra Italy

Safety Science, 2014. Special Issue on the foundations of safety science. Vol 67, August 2014, pp 1-70

Slovic P., Finucane M., Peters E. and MacGregor D., 2004. Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk and rationality. Risk Analysis, Vol 24, no 2, 2004

Snowden D., 2007. The origins of Cynefin. Cognitive Edge Pte Ltd. www.cognitive-edge.com

Stoop J., 1990. Safety and the Design Process. Doctoral Thesis Delft University of Technology, April 1990

Stoop J. and Kahan J.P., 2005. Flying is the safest way to travel: How aviation was a pioneer in independent accident investigation. European Journal of Transport and Infrastructure Research, 5, no. 2 (2005), pp. 115-128

Stoop J., 2015. Challenges to the Investigation of occurrences. Concepts and Confusion, Metaphors, Models and Methods. Side Document of the ESReDA Project group Dynamic Learning from Accident Investigation. January 2015

Stoop J., De Kroes J. and Hale R., 2017. Safety Science, a founding fathers' retrospection. Safety Science 94 (2017) 103-115

Stoop J., 2017/1. Resilience for Engineers. 7th Resilience Engineering Association Symposium, Poised to adapt. Enacting resilience potential through design, governance and organisations. 26-29 June Liege, Belgium

Stoop J., 2017/2. Drift into failure, an obsolete construct. Second International Cross-industry Safety Conference. Work as Imagined – Work as Done, Balancing Rule and Reality. 1-3 November 2017, University of Applied Science, Aviation Academy Amsterdam

Stoop J., 2017/3. How did aviation become so safe, and beyond? 53th ESRedA and EU/JRC Seminar. Enhancing Safety, the Challenge of Foresight. 14-15 November 2017, Ispra Italy

Strauch B., 2002. Investigating Human Error. Incidents, Accidents and Complex Systems. Ashgate

Taleb N., 2008. The Black Swan. The Impact of the Highly Improbable. Penquin Books Ltd

Torenbeek E., 2013. Advanced Aircraft design. Conceptual design, Analysis and Optimization of Subsonic Civil Airplanes. Wiley Aerospace series

Troadec J.P., 2013. AF447 presentation by Director of BEA. Second International Accident Investigation Forum, Singapore 23-25 April 2013

Van den Hoven J., 2013. Value Sensitive Design and Responsible Innovation. Wiley Online Library. https://doi.org/10.1002/9781118551424.ch4

Van Tooren M., 2003, Sustainable Knowledge Growth. Inaugural Lecture Delft University of Technology March 5th 2003, the Netherlands

Vincenti W., 1990. What Engineers Know and How They Know It. Analytical Studies from aeronautical History. The John Hopkins University Press

Vuorio A., Stoop J. and Johnson C., 2017. The need to Establish Consistent International Safety Investigation Guidelines for the Chemical Industries. Safety Science 95, pp. 62-74. (doi:10.1016/j.ssci.2017.02.003)

Wilson K. and Straker D., 2018. Fiction versus reality: The Impact of Hollywood on Accident Investigation. ISASI Forum, Oct-Dec 2018. P 24-26

Winters B., 2017. How to manage safety based on operational excellence principles? A case study at KLM Royal Dutch Airlines. MSc Thesis Delft University of Technology

Woods D., Patterson E., Corban J. and Watts J., 1996. Bridging the gap between user-centered intentions and actual design practice. Human Factors and Ergonomics Society Conference 40th Annual Meeting 1996.

Woods D., 2019. Essentials of resilience, revisited. https://www.researchgate.net/publication/330116587

Zimmermann K., Paries J., Amalberti R. and Hummerdal D., 2011. Chapter 18: Is the Aviation Industry ready for resilience? Mapping Human factors Assumptions across the Aviation Sector. In: Hollnagel et.al. 2011, Resilience Engineering in Practice

3 Failures of Foresight in Safety: Fantasy Risk Analysis and Blindness

Nicolas Dechy, Institut de radioprotection et de sûreté nucléaire (IRSN), France,
Yves Dien, Collectif heuristique pour l'analyse organisationnelle de sécurité (CHAOS), France,
Jan Hayes, RMIT University, Australia,
Nicola Paltrinieri, Norges teknisk naturvitenskapelige universitet (NTNU), Norway.

3.1 Executive summary

In order to foster the continuing debate about the best strategies to enhance foresight in safety, the aim of this chapter is to characterise some of the failures to foresee negative outcomes from a safety point of view. The approach followed is to review the lessons that should be learned from negative event, especially the accidents, across industrial sectors. It will enable some typical patterns to be identified that explain why companies and their regulators have recurring difficulties to anticipate risk related scenarios and accidents.

One such recurring theme is the inability to make the right assumptions when risk analyses are performed. It shows that some fantasy planning may occur especially when addressing major risk assessments. Seeking to identify a worst-case scenario is a useful concept, working in principle but not always in practice. Another recurring difficulty is mainly in recognising an accident waiting to happen. Indeed, early warning signs are usually available before an accident, but they may be too weak to trigger a learning loop or a risk analysis process. Some signals are strong, but they are not treated accordingly. The pathologies are some form of blindness (failure to see warning signs) and deafness (failure of those in charge to act on concerns raised). Those patterns of difficulties show some features of the foresight pitfalls thus giving directions for implementing measures for better anticipation.

3.2 Key messages

Failures of foresight in safety recall how difficult the challenge of risk anticipation is, especially for low probability and high consequence events including black swans. All actors in high-risk industries should remain humble.

Many provisions for foresight in safety have been implemented, but accidents highlight some of the flaws of the processes to anticipate risks. The exhaustiveness and efficiency achieved with traditional risk analysis systematic approaches remains a myth.

The implications are to remain sceptical and critical, to permanently update models and to challenge assumptions. Others are to seek out early warning signs, to prioritize risks in order to focus the available resources on critical risks.

Risk identification is a social construct. It is performed by analysts, designers, operators, and it involves decision-makers within resource (time, budget, expertise) constraints and should remain under scrutiny to avoid fantasy planning.

Analysis of risks, events and early warning signs can be assisted by tools. However, those tools integrate the designer's worldviews and purposes and may not be relevant to address some sociotechnical dimensions. Analysts use artefacts (tools, documents) to formalise their analysis, which may excessively constrain their questioning and attention, leaving them blind in important areas. To better capture risks, more open risk analysis approaches including different worldviews, opinions, transparent and flexible approaches to anticipate the unthinkable are required.

Most accidents are not inevitable but are preventable. Disasters are hard to obtain and not created overnight. They develop during an incubation period and during this time actors have an opportunity to identify latent flaws or early warning signs. Such signs and alerts provide opportunities to challenge safety beliefs and act but they are not always seized upon. It can lead to actions that are too little, too late.

One challenge is to develop high quality intelligence which requires fragmented data and disjointed information to be connected, in order to identify patterns like in a puzzle. This requires data and information structures, and processes but also people to make the expert link to the risk. Interpretation relies on worldviews, lenses and paradigms that should be debated. Assumptions and old patterns should be challenged, while new interpretations should be welcome.

Organisations are defined by what they chose to ignore and forget. Many deviations are normalised for too long. Memories of lessons from accidents are not kept and revived. Foresight in safety shifts organisations from fantasy risk planning, blindness, deafness, denial, apathy and inaction to the need to sustain proactive action on safety and thereby robust and resilient performance.

3.3 Introduction: defining challenges in foresight in safety

In order to foster the continuing debate about the best strategies to enhance foresight in safety, the aim of this chapter is to characterise some failures to foresee adverse events (serious incidents, accidents and disasters). By taking this approach, we aim to complement some of the literature in foresight in safety regarding conditions that favour failures of foresight. We have chosen case studies that we estimate are important to share and help to highlight some key vulnerabilities, rather than attempting a broad review of disaster cases.

3.3.1 Foresight and management

In everyday life, it seems that our abilities to foresee adverse outcomes from our daily activities are challenged by the limits of our planning. We tend to rely on overly optimistic assumptions that fail to integrate the surprises and unexpected events we seldom face. The same is true of organisations.

A century ago, formalisation of management as a new discipline recognised foresight as a key capability especially for engineers, managers and leaders (Stark, 1961): "'*Managing means looking ahead', gives some idea of the importance attached to planning in the business world, and it is true that if foresight is not the whole of management, it is an essential part of it. To foresee, in this context, means both to assess the future and make provision for it* (Fayol, 1916)". In summary (Kingston and Dien, 2017), foresight is about imagining the future possibilities based on knowledge of the past and present.

Stark (1961) considers the future in terms of extensions of the present which are potentialities or temporal possibilities. He defines foresight as a "*productive thinking with the elementary rules of logic its only constraint*"'. Part of foresight is 'reproductive' based on past experiences while another part is more 'creative'. He

considers that prediction can be conducted after foresight: prediction is rather a judgmental thinking with the establishment of subjective probabilities to a relatively narrower set of scenarios or period of time. Prediction is measured at a given time and verified after the event while foresight can be continuously assessed.

Ansoff (1975) already remarked that anticipating "*strategic surprises*" in a military perspective and business perspective has historically been a key issue. Many companies were surprised by the petroleum crisis in the seventies, although advance forecast about potential actions of OPEC[8] were publicly available and on the desks of some surprised managers. Ansoff points out that the assumption that those organisations were unaware because they lacked a forecasting and planning system was falsified as many had such a capability. Corporations and industries who had those planning processes were also surprised by other discontinuities. Discontinuities and surprises differ significantly from extrapolation of experience. Depending on levels of information and uncertainty and associated states of knowledge and ignorance, Ansoff (1975) opposed strategic planning which is adequate for strong signals, prepared periodically and organisation-focused; while strategic issue analysis is promoted to respond to weak signals and discontinuities, and requires a continuous, problem focused process.

With this distinction in mind, Ansoff (1975) identified "*an apparent paradox: if the firms wait until information is adequate for strategic planning, it will be increasingly surprised by crises; if it accepts vague information, the content will not be specific enough for thorough strategic planning*". Ansoff invites development of a "*gradual response through amplification and response to weak signals*" […] "*which permits gradual commitment on the part of the management*".

3.3.2 Challenges in foresight in safety

Within a reliability and safety perspective, Lannoy (2015) recalled that foresight requires a forecast to be made in an uncertain, ambiguous, controversial context or to construct a likely future by using information from the past, present and some expected future trends.

Turner (1976, 1978[9]) considers that administrative organizations may be thought of as cultural mechanisms developed to set collective goals and make

---

arrangements to deploy available resources and attain those goals. To manage safety in high-risk industries, risk anticipation activities received a lot of attention and resources for several decades in order to engineer safer systems and to demonstrate control to the regulators. In other words, risk anticipation or foresight in safety is socially constructed (Short, 1984) by different actors and through different processes, provisions and procedures.

In our experience with high-risk industries, the work of 'risk anticipation' or 'foresight in safety' currently relies on four main strategies:

- planning especially through risk assessment when addressing safety threats;
- monitoring the system (e.g. indicators), detecting and treating the early warning signs, weak or strong, indicating a threat to safety;
- setting up an operational feedback process for learning the lessons from past, internal and external events, in order to improve the system;
- preparing for the unexpected and crisis management which implies development of adaptive capabilities such as resilience.

In this chapter, we will not address the fourth strategy, though some research in these directions also provides some concepts and case studies related to weak signals of a crisis waiting to happen (e.g. Lagadec, 1994; Roux-Dufort, 2003) and also some indicators of brittleness (Woods, 2009).

In a conference organised by the French Institut pour la Maîtrise des Risques in 2015 titled "Exploring the unpredictable: how and how far?", in the aftermath of several unexpected disasters (Fukushima, German Wings, Deepwater Horizon, Eyjafjöl,…), several terms were employed by the authors in relation to foresight in safety and its failures (in table 1), with some synonymous, and some addressing different categories (Dechy et al., 2016).

With no surprise, the time dimension is essential to distinguish categories of foresight. However, the terms of the first line (atypical, unimaginable, inconceivable, unthinkable) underline the capabilities required to identify and recognise some scenarios with some difficulties to establish causal links between fragmented elements and limits of knowledge to model the phenomenon. Terms in the second line highlight the temporal difficulties to anticipate, either ultimately (unforeseeable) or about the occurrence time (unpredictable). Terms in the third line refer to their occurrence frequency or likelihood on a given period of time. The fourth category integrates the time dimension but refers to its prevention.

Table 1: Categories of phenomenon linked to event foresight and its failures

		Events	
		Expected, foreseen, without surprise	*Unexpected, unforeseen, surprising*
Category of foresight phenomenon	*Scenario imagination without time*	Identified because typical, imaginable, conceivable, thinkable	Unidentified because atypical, unimaginable, inconceivable, unthinkable
	Temporal prediction	Foreseeable, predictable	Unforeseeable, unpredictable
	Probability estimate in a time period	Probable, likely	Unlikely, improbable, 'black swan'
	Prevention until the period end	Avoidable, preventable	Unavoidable, inevitable

Risk assessment is about imagining and foreseeing what could go wrong and estimating how bad it could be in order that controls can be put in place. Poor risk assessment can lead to ineffective risk controls – controls that are ineffective or inadequate in several ways. In this way, risk assessment is a form of planning. As Clarke said, organizations that 'don't plan are seen as ineffective, poorly managed, irresponsible or sometimes just plain dumb' (Clarke, 1999, p 1). In this way, high risk industries mobilize engineers to identify process and system risks and related accidents scenarios, to model their causation, likelihood and severity. To conduct their work, they use several tools and procedures which may involve workers/experts at all levels of sociotechnical systems.

When uncertainty is high, planning is no longer simply "*a straightforward instrumental activity (a means to an end)*"(Hayes and Hopkins, 2014, p 83) rather, it can become a symbolic undertaking. When planning takes on a primarily symbolic role, the purpose of the plan becomes "*asserting to others that the uncontrolled can be controlled*" (Clarke, 1999, p 16). In this situation, symbolic plans represent a **"fantasy"** (Clarke, 1999) – in the sense of a promise that will never be fulfilled – and are often couched in a special vocabulary which then shapes discussion. The danger is that the plan itself takes on a life of its own and organizational effort is focused on managing the plan, rather than taking care of the physical system itself.

On top of the strategy to anticipate risks and related critical scenarios, a complementary strategy to prevent an accident is to foresee an 'accident waiting to happen' by relying on the feedback of system performance, especially its deficiencies such as safety related events, weak signals, precursors, near-misses incidents but also trends and drifts in key and safety performance indicators (KPIs, SPIs). Within this ESReDA FiS project group, we choose to label them 'Early Warning Signs' (EWS) as it enables to cover several concepts recalled hereafter.

Ansoff (1975) provided some elements of a definition of weak signals related to strategic management of companies and strategic surprise as "*early in the life of a threat, when the information is vague and its future course unclear*". A later definition was provided (Ansoff and Mc Donnell, 1990, p. 490) "*[a] development about which only partial information is available at the moment when response must be launched, if it is to be completed before the development impacts on the firm.*"

Some EWS of potential hazards of a system can be captured while it is designed and operated. Indeed, in-depth investigations of some accidents showed that some weak signals, precursors of accidents, near-misses (Vaughan, 1996; Llory, 1996; Carroll, 2004; Dechy et al., 2011; Jouniaux et al., 2014; Dien et al., 2012, 2014) have been recognized, at least by some actors, during an 'incubation period' (Turner, 1978) but they were disregarded. Notice that Vaughan (1996) introduces the distinction between weak signals (that are ambiguous to their link to a risk), mixed signals (signs of potential danger that are followed by signs that all was well) and routine signals that frequently recur and even if they are serious, perception of them is altered as they recur without damage.

For some of accidents, the missed opportunities to recognize the threats relates to issues of **blindness** and deafness. However, this is not an easy task and one should remain humble, vigilant and proactive. Indeed, several researchers warn investigators regularly that some signals of danger become clear only with the benefit of hindsight (Reason, 1990; Vaughan, 1996, Woods, 2009). It may lead to the following limit (Woods, 2005) "the past seems incredible, the future implausible". Though retrospective bias is a risk of event analyst, empirical analysis of several accidents showed that some signals are recognised by several actors and processes (auditing, learning) prior to a major accident (Dechy et al., 2011) showing that identification is possible without the benefit of hindsight.

Therefore, Turner (1976, 1978) considers that the challenge in normal operations is to develop "*high-quality intelligence*" (in a military context) to connect "*disjunct information*" distributed in complex systems so as to recognize an "ill-structured problem" (Simon 1973, Turner 1976). We would add to that point that the information is rooted in the history of the system and other systems (e.g. lessons from incidents in other countries, from similar systems and on generic aspects such as organisational failures). This organizational capability goes beyond effective communication as it requires an organization, processes, people to connect different fragment of information, to interpret them, to establish a link between them, referring to the "*puzzle*" metaphor (Lesca, 2001), to resolve ambiguities and to establish a well-structured problem (Turner, 1976).

In a French research project of Institut pour la Maîtrise des Risques (Jouniaux et al., 2014), the weak signal recognition process was defined in three phases:

- link data and fragments of information by experts or by data analytics pre-treatment;
- link this information to a risk or scenario; this relationship' relevance has to be qualified by experts;
- the signal is amplified when risk is redefined by management; strong signals can be minimized.

The two first steps can be aided by data analytics pre-treatment (e.g. big data, natural language processing), but any suspected link, or correlation, or surprise has to be qualified by an expert (Jouniaux et al., 2014). For more development on this issue, see chapter 10 by Marsden et al. (2020).

Near-misses and surprises are therefore opportunities to re-assess assumptions and effectiveness of risk prevention measures but also to imagine what could happen in other circumstances (applying the 'what if' motto).

This is not new, as Weick (1991) (quoted by Reason, 1997) was approaching the issue with a few proposals: "*we know that single causes are rare, but we do not know how small events can become chained together so that they result in a disastrous outcome. In the absence of this understanding, people must wait until some crisis actually occurs before they can diagnose a problem, rather than be in a position to detect a potential problem before it emerges. To anticipate and forestall disasters is to understand regularities in the way small events can combine to have disproportionately large effects*".

3.3.3 Failures of foresight in safety literature

Accidents continue to happen despite risk anticipation, foresight in safety and the implementation of risk control measures. Moreover, accidents recur that demonstrate failures to learn. Several opportunities to identify the risk or the accident waiting to happen were missed. All the prevention and protection measures, including foresight, come under scrutiny after a serious event during the investigation process.

Woods (2009) recalls that:

- "*establishing foresight encompasses extremely difficult forms of cognitive work and is an unstable process, given pressures on or from an organization's management;*
- *the difficulties arise from basic dynamic and cyclic patterns in how adaptive systems behave;*
- *emerging measures of how and where a system is brittle or resilient provide a critical resource for developing and sustaining foresight when organizations need to achieve high performance (faster, better, cheaper) and high safety (Hollnagel et al., 2005)*".

Turner (1976, 1978) considers a disaster as a "*cultural collapse*", because of the inaccuracy or inadequacy in the accepted norms and beliefs. The end of an accident is defined not in technical terms but refers to the "*full cultural readjustment*" that occurs when risk representation and risk management measures are changed. These deep changes do not occur for every near-miss.

Derived from an empirical analysis of several accidents, Turner (1976) identifies a sequence that leads to failure of foresight (table 2).

Table 2: Sequence of events of a failure of foresight (Turner, 1976, p381)

The sequence of events associated with a failure of foresight	
Stage I	Notionally normal starting point: (a) Initial culturally accepted beliefs about the world and its hazards (b) Associated precautionary norms set out in laws, codes of practice, mores, and folkways.
Stage II	Incubation period: the accumulation of an unnoticed set of events which are at odds with the accepted beliefs about hazards and the norms for their avoidance.
Stage III	Precipitating event: forces itself to the attention and transforms general perceptions of Stage II
Stage IV	Onset: the immediate consequences of the collapse of cultural precautions become apparent.
Stage V	Rescue and salvage -first stage adjustment: the immediate post collapse situation is recognized in ad hoc adjustments which permit the work of rescue and salvage to be started.
Stage VI	Full cultural readjustment: an inquiry or assessment is carried out, and beliefs and precautionary norms are adjusted to fit the newly gained understanding of the world.

The key stage regarding the failure of foresight process occurs with missed opportunities during the "*incubation period*"' when events accumulates, either not known to anyone or not fully understood by all concerned as it will be the case after the disaster, either it did not lead to changes in the risk controls.

Turner further invites us to identify conditions that make it possible for unnoticed, misperceived and misunderstood events to accumulate in a manner that leads eventually to cultural disruption. Turner (1976) identified several of those conditions that occur in stage II of 'incubation period':

- "*failure to comply with existing regulations*": they failed to realize that regulations apply or to implement them (this belongs to stage I and II about inadequate initial beliefs and norms);
- "*Rigidities in perception and belief in organizational settings*": accurate perception of the possibility of disaster was inhibited by cultural and institutional factors; 'collective blindness';
- "*The decoy problem*": attention is focused on "well -structured problems" and is distracted from "ill structured problems" in the background;
- "*Organizational exclusivity*": disregard of non-members' point of view, outsider's information or alerts are dismissed, considering the better knowledge of insiders, which can lead to forms of arrogance;
- "*Information difficulties*": associated with ill-structured, vague and complex problems; on top of communication difficulties (necessary but not sufficient condition); disjointed information is common in large organizations and the organizational risk is that they are not intelligently treated which leads to unresolved ambiguities of warning signs, orders and procedures, and responsibilities and controls;

- "*Involvement of strangers*": some people involved in the system are uninformed and untrained, which creates difficulties on top of oversimplified stereotypes about their likely behaviour;
- "*Minimizing emergent danger*": failure to see or appreciate the magnitude that remains under-estimated; under-valuation of evidence by the more complacent group and fearing the worst outcome; when impossible to ignore, (surprisingly) strengthening the response is not systematic; it may even lead to shift the blame or to believe in the use of quasi-magical means.

When reviewing Turner's added value, Weick (1998) warns that all organizations appear more vulnerable than they admit, because all develop culturally accepted beliefs and associated norms, and then accumulate unnoticed events that contradict with these world views. "*'Assumptions* […] *carry an organization's learning as well as its blind-spots'.* […] *'Assumptions conceal warning signals, deflect attention to safe issues, leave signals unnoticed because they are undefined and set the stage for surprises that necessitate revision in administrative practices'*".

Westrum (1992) distinguished three organizational cultures[10] according to the way they deal with safety-related information:

Table 3: How different organizational cultures handle safety information (Westrum, 1992)

Pathological culture	Bureaucratic culture	Generative culture
Don't want to know. Messengers (whistle-blowers) are shot. Responsibility is shirked. Failure is punished or concealed. New ideas are actively discouraged.	May not find out. Messengers are listened to if they arrive. Responsibility is compartmentalized. Failures lead to local repairs. New ideas often present problems.	Actively seek it. Messengers are trained and rewarded. Responsibility is shared. Failures lead to far-reaching reforms. New ideas are welcome.

Beside the processual, organizational and cultural views of foresight in safety, Perrow (1982, 1984) insisted on the inherent cognitive limits of operators and engineers to anticipate all interactions and cascading effects in complex systems. To some extent, many accidents are 'impossible accidents' at least from the perspective of those involved (Perrow, 1984). 'Accidents appear to be the result of highly complex coincidences which could rarely be foreseen by the people involved. The unpredictability is caused by the large number of causes and by the spread of information over the participants… Accidents do not occur because people gamble or lose, they occur because people do not believe that the accident that is about to occur is at all possible (Wagenaar and Groeneweg, 1987). In Perrow's view, foresight is always limited and some technologies should not be used because accidents are inherent, indeed can be said to be 'normal'.

3.3.4   Approach, structure and content of this chapter

First, in the next section (§3.4), we draw on past major disasters across a range of sectors (chemical, oil and gas, space) to identify patterns of failures of foresight. This empirical approach, that relies on case studies of accidents to highlight patterns of accident causation and especially organizational patterns of failure of foresight, has been implemented by several researchers on accidents and safety (e.g. Turner, 1976, 1978; Perrow, 1984; Llory, 1996, 1999; Reason, 1997; Dien et al., 2004; Llory and Montmayeul, 2010, 2018).

Indeed, from a practical point of view, those accidents and disasters investigation reports are often public and have been produced by large expert teams in relation to presidential or parliamentary commissions or independent safety boards (ESReDA, 2005). Their reports of several hundred pages may provide "*thick descriptions*" (Geertz, 1998), meaning very detailed accounts about daily activities of people, interactions and organizational and institutional processes. Those reports are of various qualities. Some can be considered as school cases that every safety specialist should read and know [e.g. Ladbroke Grove trains collision in United Kingdom in 1999 (Cullen, 2000); loss of space shuttle Columbia in 2003 (CAIB, 2003), Texas City refinery explosion in 2005 (CSB, 2007)]. In other cases, some reports are criticised publicly and associated with controversies in relation or not with the judicial investigation. No reports should be considered as perfect: the investigation may have grey zones or uncovered scopes and so can be complemented by other published material.

---

[10] *Other cultures may be categorized (e.g. Hudson (2001) added reactive, calculative, and proactive (replacing bureaucratic, between pathological and generative) within a safety culture maturity model).*

Notice that several sub-cultures or maturities co-exist within the same organisation, department which questions their overlapping, integration and interactions.

More fundamentally, this approach aims at identifying organizational "vulnerabilities". Indeed, accidents offer the 'gift of failure' (Wilpert according to Carroll and Fahlbruch, 2011) - an opportunity to learn about safe and unsafe operations. Accidents are the "*royal road*" (Llory, 1996) to access to real (mal) functioning of organizations[11], as some hidden phenomena in the "*dark side*" of organizations (Vaughan, 1999) may become more visible in accidents. This strategy is opposite to the study of normal operations and banality of organizational life (Vaughan, 1996) that is often conducted to identify "best ways" to cope with variability, to enhance reliability and to adapt and recover from adverse events ('High Reliability Organizations' (Roberts, 1990, Laporte and Consolini 1991), 'Resilience engineering' (Hollnagel et al., 2006), and 'Safety II' (Hollnagel, 2014)).

Second, analytical developments start in section §3.5 by discussing how risk analysis can fail and continue in section §3.6 with blindness patterns. This analysis also relies on our investigator and risk analyst's practical experiences in the field, either within high-risk industries, expert's institutes that support the safety regulation and consultancy firms, or also as researches in risk assessment and management. We provide a few conclusions and perspectives in section §3.7.

3.4   Accidents that highlighted some failures of foresight

3.4.1   Toulouse disaster in 2001 in France

On 21st September 2001, a powerful explosion occurred at the AZF fertilizer and chemical plant in Toulouse suburbs which lead to significant damage and effects[12]. The direct causes are still under debate and controversy between prosecution, lawyers for Total and other stakeholders even after the third trial in 2017 (Dechy et al., 2018). In summary, the explosion of off-specification ammonium nitrate was not prevented and turned into disaster due to several failures in risk assessment, management, governance, control and regulation (Dechy et al., 2004).

This accident belongs to the category of '*atypical*' accidents (Paltrinieri et al., 2012). It means that this low probability-high consequence accident was not among the worst-case scenarios captured by traditional risk analysis and

formalised within the safety case submitted under Seveso I and II directives by the licensee to the regulator, nor one of the scenarios used in the eighties and nineties to establish land use planning and emergency planning. This striking lesson revealed flaws not only in the risk analysis process used to identify the relevant scenarios but also in the negotiations upon the scenario's basis on which to define safety measures.

An underlying reason is that the residual scientific uncertainty on ammonium nitrate chemical sensitivity were underestimated. There were ambiguities in the behaviour of ammonium nitrate that belong to the category of "*occasional explosives*" (Médard, 1979). In some conditions (e.g. contamination by chemical impurities, fuel, air pressure…), inherent and residual explosion risk could increase. In addition, lessons from accidents in the last century were assumed to be learned, and also gave decision-makers confidence to exclude the occurrence of those conditions and initiators if industries operate normally. The conservative approach was therefore limited. There were also deficiencies in knowledge management about accidents lessons and chemical properties. Also, the fertilizer industry lobby pushed to consider that the "worst-case scenario" for ammonium nitrate storage were fires with toxic fumes, because such consequences are more likely, rather than a massive explosion. Imagination was therefore limited though the explosion risk remained inherent, especially if conditions were gathered. In addition, the "envelope approach" of safety case studies lead the licensee and regulator to focus on other scenarios of the plant and of the neighbouring plants which were more severe (several toxic cloud release) than a potential ammonium nitrate explosion.

Once approved in 1989, the land use planning (LUP) process and plan enabled local authorities to freeze further nearby urban development, but it was too late as buildings and houses were already close by, and the plan had no retroactive force to expropriate people (Dechy et al., 2005). In addition, the effects' distance was under-estimated because scenarios were rather incidents than worst cases. This occurred as an outcome of negotiation between the regulator and operator after Seveso I regulation. A reason was that regulator and operators wanted to find an incentive such as a way to value the financial investment in prevention measures with an impact on a reduced safety distance that impacts the land-use.

---

[11] *In reference to Sigmund Freud's metaphor: "Dreams are the royal road to the unconscious. "*

[12] *31 fatalities, estimates of French national health institute (InVS) are about 10 000 injured, 14 000 post-traumatic acute stress, 27 000 houses/flats damaged, 1,5 to 2,5 billion euros of damages; Dechy et al., 2004a).*

3.4.2   Buncefield accident in 2005 in United Kingdom

The Buncefield oil depot fire and explosion the 11th of December 2005 destroyed a large part of the site and the surrounding area. The immediate trigger for the catastrophe was a large petrol storage tank that overflowed whilst being filled from a pipeline. About 300 tons of petrol escaped from the tank, 10% of which turned to a flammable vapour cloud. Once ignited, the magnitude of the resultant vapour cloud explosion (VCE) was much greater than anyone knew was possible. The effects were fortunately more limited (43 injuries) as it occurred at 6 am of a Sunday morning (MIIB, 2008; COMAH Competent Authority, 2011).

The Buncefield oil depot was subject to the so-called ''Seveso-II'' Directive, but the scenario that occurred was not taken into account in the mandatory safety report, as in the case of Toulouse accident (Paltrinieri et al., 2012). Formation of a vapour cloud due to tank overfilling and consequent VCE were not deemed possible, neither by the company nor by the competent authorities. The design of the tank itself may have contributed to the vapour/mist formation in a manner that was not foreseen by designers. The tank was fitted with a deflector plate that led to a cascade of petrol droplets through the air. Moreover, most of the remaining fuel running down the wall hit a structural stiffening ring and detached from the tank wall, creating a second cascade of droplets. These conditions promoted the evaporation of the lighter components of petrol, (e.g. butane), which were allowed in higher concentration in the winter season. The unexpected strength of the subsequent VCE was caused by presence of equipment and trees increasing the turbulence of the flow and/or providing a certain level of confinement and a substantial energy of the ignition source (MIIB, 2008).

The worst credible scenario included in the site safety report was a major liquid fuel pool fire (COMAH Competent Authority, 2011). Although, risk analyses of oil depot in France in the early 2000's considered potential leaks that could form an explosive cloud especially for volatile hydrocarbons, the risk considered was the one of *unconfined* vapour cloud explosion, implying that it did not lead to overpressures over 200 mbar. Therefore, pool fires were often the worst case for safety cases reports with envelope effects out of the plant site. For this reason, the actual accident scenario can be defined as "atypical" (Paltrinieri et al., 2012).

Hazard identification has important aims: it may highlight possible malfunctions of the systems, outline related losses of containment and describe potential consequences. An atypical accident may occur when hazard identification does not produce a complete overview of hazards due to a lack of specific knowledge and a low awareness of associated risks, because it deviates from normal expectations of unwanted events or worst-case reference scenarios (Paltrinieri et al., 2013). This qualitative pre-assessment is the foundation of risk management. For instance, Seveso safety reports are supposed to conservatively evaluate worst-case scenarios and safety measures that are used for the operation licensing, calculation of effects' distance used in emergency response planning and in the design of safety area in land use planning (LUP).

Further accident databases analysis and research (Paltrinieri et al. 2012) demonstrates that VCEs in oil depots were not unknown before. In fact, since the middle of 1960s, there is record of VCE accidents occurring approximately every five years in oil depots around the world. Effective knowledge management searching for and considering such historical lessons and strong warnings was missing in Buncefield. It can be speculated that this was due to the inaccuracy of the analysis process (availability of resources to seek for accident data?) while assessing related risks, whose results were (only apparently) validated because it was consistent with similar process safety studies (addressing similar plants or the former plant documents) and within the basic experience. In other words, the risk analysis process was affected by social conventions and by an implicit code of practice that failed to integrate some knowledge about accidents.

3.4.3   Texas City refinery accident in 2005 in USA

On March 23, 2005, an explosion and fire at the BP refinery in Texas City lead to 15 deaths and 180 injuries. The CSB (2007) noted that: "*The Texas City disaster was caused by organizational and safety deficiencies at all levels of the BP Corporation*". The board member and CEO of the US Chemical Safety Board (CSB), Carolyn Merritt (2007) underlined: "*cost cutting, production pressures, and a failure to invest left the BP Texas City refinery vulnerable to a catastrophe.*"

Failures in major risk assessment have been noticed. The risk of the blowdown drum releasing a potential explosion cloud was identified as it was known as an "*antiquity of the fifties*" by the operator and the industry standard had changed to require a flare for new designs. It was known by the regulator (OSHA) who requested the removal of antique flare, but the operator (Amoco in the nineties)

relied on 'grandfathering' [13]to avoid the need for modification and replacement by a safer equipment. Internal audits of the BP group in early 2000's already blamed the BP refinery '*check book mentality*' for maintenance of equipment, for safety policy and more generally for selection of risk controls.

The failures to learn (Hopkins, 2008) and the blindness to process safety deterioration and alerts (e.g. loss of containment incidents were increasing) have been eased by the confusion with worker safety metrics that were improving (further explanations in section §3.6.3). CSB (2007, p18) investigation found that "*warning signs of a possible disaster were present for several years, but company officials did not intervene effectively to prevent it*." Merritt added (2007) that "*adhering to and enforcing federal regulations already on the books would likely have prevented this accident and its tragic consequences*."

Indeed, the CSB investigation showed that some BP members had identified the rise of major risks already in 2002. The new director of BP's South Houston Integrated Site observed in 2002 that the Texas City refinery infrastructure and equipment were "*in complete decline*". (CSB, 2007). An internal follow-up analysis concluded that "*the current integrity and reliability issues at TCR* [Texas City Refinery] *are clearly linked to the reduction in maintenance spending over the last decade*" (CSB, 2007, p153). Several other internal studies, surveys, audits and also serious incidents alerted and signalled the severity of deficiencies but the response of BP managers was "*too little and too late*" (Merritt, 2007). For example, there was a poor (only 30% of corrective actions were implemented) and declining implementation of corrective actions. Furthermore, a safety culture assessment conducted by an external company (Telos Group) alerted the managers in January 2005 about the critical and degraded state of the refinery. The Telos report identified the organisational and process safety problems that were found by the CSB in retrospect.

3.4.4 San Bruno pipeline failure in 2010 in USA

In September 9, 2010, eight members of the public were killed when a gas transmission pipeline ruptured at San Bruno, California (NTSB, 2011). The rupture occurred when a longitudinal seam weld failed. The weld had been poorly made during construction of the pipeline in 1956. The line had not been inspected or

tested since that time. Failure of pressure control at the upstream terminal led to a pressure rise in the line to close to the maximum allowable operating pressure (MAOP). Control room operators chose to troubleshoot the pressure problems but not to isolate the downstream pipelines. After exposure to higher than normal pressure for approximately one hour, the line failed.

Integrity management of ageing buried pipeline networks is an exercise in managing risk. In this case, the involved operating company PG&E (Pacific Gas and Electric) had put significant effort into developing a risk model of the system but the primary focus of the system was not on fault identification and repair. Indeed, the database contained inaccurate data, inappropriate risk algorithms and lacked any real-world connection. In summary, shortcomings in the GIS (geographic information system) and associated procedures include (Hayes and Hopkins, 2015):

- The database used as the basis for risk ranking included physical data that was optimistic and / or incorrect and there was no system of data checking in place.
- Algorithms for establishing inspection priorities averaged risk scores for a given pipeline segment across all threats to pipeline integrity thereby hiding problems, rather than highlighting them.
- Regardless of the identified threat, higher risk segments were mostly subjected to external corrosion direct assessment, a type of inspection which finds pipeline integrity problems for external corrosion threats only.
- There was no system in place to review the performance of the integrity management system overall i.e. to compare high risk segments identified with inspections done and with actual leaks seen to determine if the system was effective and/or how it might be improved.
- The system produced only a prioritised list of pipeline segments based on threats to integrity. Whilst such a system could, in theory at least, be used to determine where funds should be spent to improve integrity, it makes no attempt to comment on overall risk acceptability and the total budget required.

---

[13] *'Grandfathering' is a legal process that gives the benefits of anteriority to existing processes over new legal requirements which have no retroactive force.*

Another significant factor in this accident was the MAOP for the pipeline segment that ruptured. It had been determined based on the highest operating pressure seen in the system in the previous five years, rather than by testing. This was specifically allowed for pipelines of this age under the relevant regulations. Newer pipelines are required to be hydrotested but pressure testing requirements of the relevant standard had been grandfathered in this case, in a similar way to the old design of the Texas City refinery blowdown system. Given the flawed weld, it is unlikely that the pipeline would have passed a hydrotest and yet such a test was not required by the regulator, nor seen as necessary by the operating company.

3.4.5   NASA space shuttle losses in 1986 and 2003

On the 28th of January 1986, after an unusually cold night for Florida (minus seven degrees Celsius) which required a launch delay due to ice on the shuttle launch pad, the space shuttle Challenger exploded 73 seconds after its launch with all seven astronauts killed. The technical explanation of the accident centred on the failure of the joint between two segments on the right Solid Rocket Booster (SRB). The O-rings that were intended to seal this joint from hot gases leaking through the joint failed to perform properly, due to the extremely low temperatures for the intended launch environment. This leak allowed a flame to emerge from the SRB and to impinge upon the external fuel tank (Rogers et al., 1986).

On 1st of February 2003 the space shuttle Columbia disintegrated during its re-entry phase into the Earth's atmosphere after a 16-day mission on orbit around the Earth. The seven astronauts died in the accident. The Columbia mission was the 113th space shuttle flight. The technical cause for the loss of Columbia is clearly identified. During the shuttle's ascent phase, a piece of insulating foam separated from the left bipod ramp located on the external fuel tank. It struck the leading edge of the orbiter's left wing at a relative speed of about 800 km/h. The impact caused a hole in the shuttle thermal protection system, a particularly vulnerable area during re-entry in the dense layer of Earth's atmosphere (CAIB, 2003).

Beyond direct technical causes of the accidents, there is a similarity in the organizational patterns of the two accidents, with "*echoes*" of Challenger's causes in Columbia's (CAIB, 2003). Both disasters can be seen as symptoms of foresight blindness. Indeed, for the Challenger case, organizations (the NASA and its contractor Morton Thiokol Inc.) were unable to fully acknowledge the design flaw

of the rocket joint: they fail to recognize it as a problem to be fixed and they perceive it as an acceptable flight risk. According to the Presidential Commission it was "*an accident rooted in history*". There were warnings from several lower-level engineers from the subcontractor and concerns within some NASA engineers, that the joints were poorly designed, including one report that said they could cause a catastrophe. Unfortunately, it had no significant impact on decision makers. Warnings were unheeded by top managers.

In addition, NASA did not retain memory of the lessons learned. As Diane Vaughan noted (1996, p. 422): "*Few of the people in the top NASA administrative positions exposed to the lessons of the Challenger tragedy are still here. The new leaders stress safety but they are fighting for dollars and making budget cuts. History repeats, as economy and production are again priorities*". One effect of this policy is the feeling that as long as no serious problem occurs, the situation is seen to be under control ("so far, so good!!"): "*success-engendered safety optimism*" (CAIB, 2003, p. 114) and "*it could lead to a tendency to accept risk solely because of prior success.*" (CAIB,2003, p. 114). Risk is measured by past successes: "*The acceptance and success of these [past Challenger] flights is taken as evidence of safety*" (Feynman, 1986). Example of the Columbia accident is typical of inability to take account of small failures "forgotten" during the risk analysis carried out during the design phase. The fatal flight was the 113th mission of a space shuttle. The various shuttle orbiters had been hit by debris for each of the previous 112 missions[14]. It was the 7th detachment from the left bipod, knowing that a detachment from the right bipod was never seen. Those events were identified as safety issues by designer that defined specifications to prevent them. The consequence of these multiple failures was not a risk (re)assessment but a progressive shift in evaluation of the severity of incidents, gradually, mission after mission:

- From "safety of flight issue" to a "turnaround issue" (simple maintenance);
- From "Out-of-Family" problem (operation or performance outside the expected performance range for a given parameter or which has not previously been experienced) to "In-Family" problem (a reportable problem that was previously experienced, analysed, and understood);
- In other words, from jeopardizing safety to acceptable risk.

---

[14] *Knowing that no data were available on the fatal Challenger mission.*

3.5 Failures of foresight due to inadequate risk assessment

Risk prevention through anticipation relies especially on the ability to identify safety threats and to model risk levels adequately in order to ensure proportionate risk control measures. It is mainly depending on what is imagined and considered in input, at the modelling phase of risk analysis and finally about what is done with the output to prevent accidents. These processes are collective and socially constructed and relate to engineering standards and regulation.

When addressing low frequency-high consequences events that are for some beyond experience, this becomes more challenging. Many accidents have taken companies apparently by surprise as a result of poor engineering risk assessments. Among the mechanisms highlighted by accidents, the next paragraphs discuss some of the failures, especially on complexity modelling, imagination, quantification, and point to some recurring flaws and biases that downgrade these activities.

3.5.1 Limits in capturing the complexity of reality

Engineers have developed several formal methods in safety and reliability to identify (e.g. What if? Systematic questioning) and assess risk (e.g. failure mode and defects analysis, preliminary risk analysis, HAZOP, fault trees, bow-ties…) and more complex tools (e.g. 2D, 3D) to model physical and chemical phenomenon (fire, explosion…). A first limit stands in the inability to capture the complexity of reality into risk assessment approaches and risk modelling. The use of scenarios is fairly common, providing benefits and showing limits as well (see chapter 5, Ferjencik et al., 2020).

A striking lesson of both the Toulouse disaster and the Buncefield accident was that their scenarios were not identified as the worst-cases scenarios that were integrated in the safety case studies reports, emergency and land use planning area. We therefore called them "*atypical*" (Paltrinieri et al., 2012) as they are not enough typical to serve for those purposes. This recurring finding hampers the legitimacy of such engineering plans to anticipate risk and prepare emergency procedures (with the risk of becoming a "*fantasy*" (Clarke, 1999)).

Everyone knows that models are a simplification of reality. Similarly, some researchers (Perrow, 1984) have criticized system designers' abilities to address the complexity of sociotechnical systems and even to prevent and protect from inevitability of accidents in such settings. Though this 'normal accident' theory has been challenged for different reasons (Hopkins, 2001; Dechy et al., 2011; Dien et al., 2013) mainly because in most accidents, warning signs (weak or strong) are available before the major accident, but often not treated accordingly, the warning from Perrow should lead to vigilance attitude when establishing such scenarios.

The more we study risk assessment practices and failures of foresight, the more we become cautious about the interpretation of results provided by the application of formal approaches of risk analysis (e.g. hazard identification techniques) as they oversimplify reality [e.g. Buncefield with bow-tie, HAZOP, Paltrinieri et al., 2012; with FMEA, Thellier, 2017, 2018]. Several incidents, events, near-misses reveal some unexpected scenarios with unanticipated interactions and combinations of systems, sub-system, and component failures that were not captured in the risk analysis format. Then the question can become to wonder if they are used as opportunities to learn about those missed scenarios applying the 'what-if' motto as a driver for risk imagination.

Finally, it is widely acknowledged that engineering approaches poorly address and even divert from addressing human and organisational factors and the sociotechnical interactions and complexity (Rasmussen, 1997, Wilpert and Fahlbruch, 1998; Thellier, 2017; Vautier et al., 2018, Llory et al., 2018). This remains a major blindness and favours the "*cultural collapse*" mentioned by Turner.

3.5.2 Failures of imagination in defining the worst case

Preventing accidents through foresight requires safety threats to be identified. Efficiency criteria are the 'imagination', 'exhaustiveness' and the 'filtering' or defining a "hierarchy of risks". For decades, deterministic approaches have been widely used in several industrial sectors in order to specify appropriate measures and barriers to deal with major risks and protect workers and neighbours of industrial plants. This approach requires conventional scenarios to be defined and studied by postulating some "worst-case scenarios" that can lead to the complete degradation of a storage and pipe in order to study "envelope effects" (Hourtoulou, 2002, Libmann, 1996). Notice that the nuclear reactor meltdown was not formally postulated in conventional scenarios (Libmann, 1996) before Three Mile Island accident in 1979, though it had been imagined by some nuclear engineers, even before the WASH1400 probabilistic risk assessment report (Rasmussen, 1975). Though this report brought some advances, it was also largely reviewed and criticized.

For some researchers (e.g. Clarke, 1999, 2006), the "worst-case approaches" become an exercise in "*fantasy planning*" when they convince their users and fool them. It can lead to fantasy land use planning around pipelines (Hayes and Hopkins, 2015) or petrochemical plants if we look at Toulouse and Buncefield accidents. Indeed, in our experience (Dien and Dechy, 2016), the imagination is often limited by a dose of realism.

First, some initiators are excluded: e.g. at Toulouse disaster in 2001, some explosion triggering factors were excluded such as the confinement or the contamination factors, considering that those conditions would not occur in "normal operations" thus preventing any explosion. The industry lobby focused on the more likely event of a fire as the worst-case scenario, (Dechy et al., 2004, 2005). In the nuclear sector, some phenomena were not taken into account until recently (e.g. tornado in France) and the severity of natural hazards was under-estimated in Japan as shown by Fukushima accident (Diet, 2012) and worldwide. Meteorites are excluded by all industries.

Then, major risk modelling is supposed to espouse the so-called 'conservative approach' but in practice there are limits. "Worst case-scenario" are supposed to display "envelope effects" but are not always the maximum physically possible. All the parameters that influence them are not integrated in the model with all at the highest intensity. In France, in the early 2000s, for oil storage tanks in an open environment, it was considered that the highest overpressures of *unconfined* vapour cloud explosion were below 200 mbar. During the Buncefield accident (2005), overpressures were over one bar in some locations, due to differences in the initiating energy, the oil mist, the nature of hydrocarbons and turbulence factors. The impact of those parameters was more or less known by some experts and researchers but often not modelled by practitioners. But, in addition specific adverse conditions occurred. Indeed, trees acted to increase turbulence and played a role of flame accelerator (MIIB, 2008; COMAH Competent Authority, 2011; Paltrinieri et al., 2012). The oil mist explosion was very energetic because a deflector increases droplets when the liquid was leaking along the tank; there was a higher concentration of relatively more volatile components in the gasoline (in winter it is allowed by law).

At Fukushima, a tsunami wave with a height level of 14 meters was imagined before the accident but excluded for probabilistic reasons. Therefore, the chairman of the independent commission concluded "*It was a profoundly*

manmade disaster – that could and should have been foreseen and prevented. And its effects could have been mitigated by a more effective human response" (Diet, 2012, p9).

In the end, the exhaustiveness and efficiency achieved through the use of these systematic approaches remained a myth for some time. These beliefs (Turner, 1978) are better challenged today as some stakeholders of these analysis would maintain some doubts that the residual risks have been achieved (e.g. for French nuclear safety after Fukushima, Couturier et al.; 2016).

To better capture risks, our main lesson to share is to support more open risk analysis approaches including different worldviews, opinions, transparent and flexible approaches to anticipate the unthinkable.

3.5.3 Traps of quantification

Several benefits but also several traps from quantification can be identified (Lannoy, 2016). Probabilities and frequencies of accidents are commonly underestimated in several industries by more than an order of magnitude, before accidents occur and for several reasons. We could insist here only on a few limits that appear to us as noteworthy.

A first limit that appears to us as quite important is expert dependence or sensitivity to the expert approach. A European research project, ASSURANCE (Hourtoulou, 2002) showed that for the same chemical plant, with seven different experts from Europe who selected their scenarios, their modelling tools and data, the results could differ by a factor of 6 in effects distances modelled for worst-case scenarios and by three orders of magnitude for probability estimates.

A second limit, relies on the beliefs in quantification that can lead to perverse effects that can be detrimental to safety. Before the Challenger launch decision, engineers were asked to "*quantify their doubts*" about the robustness of the O-rings, as the engineer perceptions were treated as "subjective", "uncertain", "qualitative", "affective", "emotional" and could not comply with a technical culture that required quantitative data (Vaughan, 1996). Of course, the engineers would have been equally stumped if they had been told to demonstrate numerically that the system was safe, showing that it is not quantification itself that is necessarily a problem but the way in which it is brought to bear in decision making.

Problems with pipeline integrity management at PG&E are similar in that the data used for quantification was significantly incorrect. Some critical information had been entered into the system as 'dummy' values when a new database was introduced some years before the accident. This "garbage in" resulted in graphical output showing that risk was declining whereas, in fact, pipeline integrity management was grounded in 'garbage out' results.

Third, there can be some inappropriate use of statistical laws which are often used in reliability of equipment due to great number laws. For example, Gaussian distribution towards the average that are applied inadequately to low probability events and high consequences, infrequent extreme events, some of them could be considered as black swans hidden in the "fat tail" of statistical distribution (Taleb, 2007). We see this in the San Bruno case specifically where the entire pipeline network was divided up into several hundred segments with a risk score produced for each possible threat to integrity (external corrosion, ground movement, design and materials, third party interference) and each segment. The problem, however was that all threats for a given segment were averaged, thus effectively hiding high scores. In some cases, the use of multi-criteria decision analysis procedures may prevent some of the quantification traps (Linkov and Moberg, 2017; Merad and Trump, 2020).

3.5.4   Cognitive biases and social conventions

In addition, one should remember that the map is not the territory; therefore analysts and managers should be very cautious about the limits of the approaches, especially the dependency on experts with regards to the limits in their background knowledge, their procedures to treat limited data, their tools (e.g. Dien et al., 2012, Maguire and Hardy, 2013; Power, 2016; Merad and Trump, 2020). There are several cognitive biases especially with perception of low probability-high consequences events (work of Tversky and Kahneman, 1974; Taleb, 2007; Merad et al., 2016). A famous example is also related to the NASA space shuttle Challenger explosion, with under-estimates by NASA managers of the likelihood of a failure of a launch (Feynman, 1986).

The constraints in which risk assessments are performed should be addressed, as risk assessment are projects conducted under constraints (Merad, 2010). The resources, the methods or level of guidance and aiding, the level of openness and flexible mindset, should be questioned.

Several "worst-case scenarios" are conventions that are a social construct (e.g. a vapour cloud explosion could not occur on an oil depot because of lack of confinement). They may inherently integrate residual risk (Couturier et al., 2016) that is not treated accordingly (e.g. occasional explosives such as ammonium nitrate fertilizer that are not inherently safe (Dechy et al., 2004; Marlair and Kordek, 2005)). These conventions are changed especially after disasters and are in retrospect better acknowledged to be some fantasy planning (Clarke, 1999, 2006) that fooled their users for a while.

3.6   Failures of foresight due to blindness

First, we should notice that "blindness" may refer to several phenomena. It is obvious that it puts an emphasis on the inability to see and recognise from a cognitive and cultural point of view the early warning signs. But it can also be related to some failures to learn the lessons from strong signals such as lessons from accidents, by lack of memorisation or poor knowledge management. And it could also be understood as the inability to react to weak signals and change the course of actions as planned. Those contradictory signs may offer opportunities to reassess assumptions, models, controls and barriers, but are they seized? At some point, from a pathology such as blindness, it can shift to the inability to listen to and ear the alerts (deafness) or even some denial, apathy and inaction.

3.6.1   Engineering failures to reassess models against warning signs

The previous sections describes possible pitfalls in model development. The focus here is on issues with models in use. As the map is not the territory, in principle, a key preoccupation for risk analysts is to benchmark their predictions against real observations and collect new information that could challenge or help them to update and increase the reliability of their models. However, evidence from accidents shows that this is not always performed adequately either by the analysts in charge or by the professional community. This first argument hereafter relates to quality assurance in risk modelling; while the second addresses the opportunities to revise assumptions based on EWS treatment.

The strength of the explosion of Buncefield accident was unexpectedly severe. However, history shows (see next table n°3) that it was not the first accident with important effects for unconfined vapour explosion. Moreover, several modelling approaches were available to take into account various parameters that influence

explosion strength (e.g. multi-energy methods by TNO[15] since the eighties aimed at better taking into account turbulence and confinement parameters). The explosion in Saint-Herblain in 1991 in France was also surprisingly severe and did help to some extent to reveal to some experts (e.g. in France at INERIS, Lechaudel and Mouilleau, 1995) some under-estimates in unconfined vapour explosion modelling. However, common practices of risk analysts for such modelling were not so aware of this kind of phenomenon.

Beyond a quality assurance approach to benchmark the quality and the reliability of risk modelling, a complementary strategy is to be reactive to EWS. Those EWS should be proactively seized as opportunities to check safety assumptions, risk modelling, and therefore relevance of designs, rules and decisions. EWS should be both captured and treated within an engineering loop (redesign a design flaw) or operational loop (monitor or change the organisation).

To some extent, this issue deals about the treatment of risk under uncertainty. Several accidents (e.g. Therac-25, (Leveson and Turner, 1993), DC-10 crash in Ermenonville, (Llory, 1996)) have revealed these difficulties to be reactive and proactive to different magnitude of EWS.

As described above, PG&E operated a risk-based model for pipeline integrity management to determine inspection priorities. The major problems with the data and algorithms on which the model was based might have been identified before the San Bruno failure if only a link had been made between field experience and model predictions. Two kinds of field data were available (inspection results and pipeline leaks) and neither of these were used to verify that the risk model was operating as intended.

NASA engineers observed that O-rings of space shuttle boosters were damaged. Specifications of designers required no damage. The engineers discussed redesigning the system but this would take two years and could introduce new risks. Engineering preferred to choose evils that were known rather than unknown (Vaughan, 1996). This position was supported by the success of ongoing launches with anomalies but no major failures, which was taken as proof of safety, despite the fact that it was rather a confusion with reliability (Dien and Llory, 2006).

The Challenger accident provided another lesson on this issue on the eve of the launch, when low temperatures in Florida generated concerns for some engineers from the subcontractor and manufacturer of the booster. Engineers lacked data to challenge prevailing assumptions about O-rings behaviour for low temperatures and were asked to quantify their doubts in order to convince NASA managers to set-up a new safety criterion for the decision to launch space shuttle.

For this reason, an extended strategy of dynamic risk management is suggested by Paltrinieri et al. (2015) to define an appropriate decision-making process based on comprehensive monitoring activity. It should also integrate a dynamic learning as a follow-up from events (ESReDA, 2015) as described hereafter.

3.6.2   Failures to learn, to memorize and to manage knowledge

After an event, investigations seek to identify lessons to avoid a similar accident in the future, here and elsewhere. Failures to learn were numerous in BP's Texas City refinery before their major accident (CSB, 2007; Hopkins, 2008) and are a common root cause of accidents, potentially an "ultimate root cause" (Dechy et al., 2009, 2011, 2018; Dien et al., 2012; ESReDA, 2015).

Dynamic learning should avoid blindness, forgetting and continuous improvement should be observed. Accidents, however, recur (Kletz, 1993) with similar organisational root causes, as for NASA space shuttles and BP accidents. Losses of memory of lessons from accidents do occur (Kletz, 1993; Ferjencik and Dechy, 2016, Dechy et al., 2016), with both individuals and organisations forgetting (see chapter 4 by Ferjencik and Dechy, 2020). One should humbly acknowledge that learning to prevent the next accident, remains a high challenge.

Some accidents show that lessons from accidents are missed which highlights a lack of awareness and poor knowledge management. Problems with the modelling of unconfined vapour cloud explosions at Buncefield have already been described. Surprisingly, the trend to underestimating the overpressure that could be achieved continued after the event, highlighting the difficulty to learn and change systems.

---

[15] https://www.tno.nl/en/

	Location	Date	Loss of containment
Before Buncefield	Houston, Texas, USA	April 1962	Leak from a gasoline tank
	Baytown, Texas, USA	27 January 1977	Overfilling of a ship with gasoline
	Newark, New Jersey, USA	7 January 1983	Overfilling of an unleaded gasoline tank
	Naples, Italy	21 December 1985	Overfilling of an unleaded gasoline tank
	St Herblain, France	7 October 1991	Leak of gasoline from a transfer line
	Jacksonville, Florida, USA	2 January 1993	Overfilling of an unleaded gasoline tank
	Laem Chabang, Thailand	2 December 1999	Overfilling of a gasoline tank
After B.	San Juan Bay, Puerto Rico	23 October 2009	Overfilling of a gasoline tank
	Jaipur, India	29 October 2009	Valve left open

The ammonium nitrate accidents in 1947 at Texas City (United States) and Brest (France) ports, occurred in specific configurations with fuel and confinement in cargo of ships. Those accidents were used as a proof of the need for the two necessary conditions to have explosions. It led the fertilizer industry to consider that these two conditions could not happen in open ground storage in normal operation therefore explosion risk could be excluded. Scientific knowledge management and information sharing about physical properties is sometimes considered as insufficient and not addressed by systematic risk analysis procedures. Even in the study of better known physical and chemical phenomenon, surprises can happen. The case of "occasional explosives" (Médard, 1979) is typical; specifically, for ammonium nitrate where some residual risks were not intrinsically excluded thus forgotten (Dechy et al., 2004, Gyenes and Dechy, 2016) as dramatically recalled by the 2020 Beyrouth disaster. Several properties at the limits are not discovered through quality tests but could be more likely if safety and research tests were conducted more often. In the nuclear industry, some phenomenon are still research subjects, fifty years after the first nuclear power plants started.

An explanation of the severity of the consequences of the Texas City refinery explosion in 2005 (15 fatalities) comes from the location of temporary buildings for maintenance workers that were too close to the hazardous processes highlighting an inadequate siting procedure. It showed a lack of vulnerability analysis and worst case approach, but also a failure to remember the logic of targets removal learned from explosions in refineries (e.g. after La Mède (France) accident in 1991, control rooms became "blast proof"') or in silos (e.g. with the Blaye (France) explosion in 1997 where the administrative quarters were below the silo causing the death of workers not directly necessary to the process.

It can be seen in these cases that lessons from past incidents were not always explored with the aim "what if", rather past accidents were reinterpreted as proof of reliability or resilience instead of warnings of danger. The December 1999 storm that devastated western Europe and created an emergency situation at the French nuclear power plant of Blaye due to loss of power after a flooding of equipment, was fortunately properly managed. The side-effect is that it did not trigger an international strong learning process as Fukushima did. It is considered in retrospect as one of the precursors of Fukushima – an EWS that was lost.

3.6.3 Failures in monitoring and in listening EWS, failures to change

Early warning signs are often missed but there are opportunities during the incubation period to recognise them, to listen to alerts from people, especially during windows of opportunity (Edmondson, 2005) to implement changes. Here again, the challenge is not easy especially in highly complex systems with many EWS, a lot of noise, some difficulties to filter issues and determine the priorities. This is additionally more difficult for major risk prevention with low frequency and probability events.

In the Buncefield accident, the Automatic Tank Gauging (ATG) system preventing tank overfilling had been stuck 14 times in the months before the accident. Sometimes this was logged as a fault by the supervisors and other times it was not. Moreover, the contractor company that installed the ATG system never considered that the gauge should be investigated, even if they had been frequently called to rectify the matter (COMAH Competent Authority, 2011). The problem with

measurement of the critical parameter of tank level was therefore known by many people and yet it was not fixed.

The Texas City refinery explosion in 2005 (CSB, 2007) highlighted a few design flaws with latency effects, such as the "antique" flare (build in the fifties) which was abandoned in petroleum standards available since the eighties. The opportunity to remove it was investigated by a former owner of the refinery in the nineties especially under regulatory pressure, but the cost was used as a factor to postpone the corrective actions, as well as the 'grandfathering' argument. These missed opportunities to comply to a new regulation in order to improve safety (Ferjencik and Dechy, 2016) were normalised by control authorities.

But one of the most striking lesson from this accident remains the inability to learn (Hopkins, 2008) implying a difficulty to change which was "too little, too late" (Merritt, 2007). The numerous latent flaws (Reason, 1990, 1997) caused by the lack of maintenance were severe "*cost cutting, production pressures, and a failure to invest left the BP Texas City refinery vulnerable to a catastrophe*" (Merritt, 2007). Their severity was visible before the accident by many actors at the refinery (managers, operators "closest to the valves", health and safety engineers, investigators, auditors) (Dechy et al., 2011). A 2003 internal BP audit warned that the reasons for "such a poor state" of the infrastructure "in complete decline" were known to be "culture and money". The check book mentality was under-fire but was not turned around. The hindsight bias excuse does not apply here (and not only here, see chapter 11 about whistle-blowers, Dien et al., 2020)! It was not a failure of detection of weak signals, nor a myopia or blindness, but rather some deafness and denial to strong signals and inaction. CSB (2007) found: "warning signs of a possible disaster were present for several years, but company officials did not intervene effectively to prevent it."

In addition, BP managers failed to manage major risks and process safety, as they over relied on the wrong metrics related to worker safety (CSB, 2007; Baker et al., 2007). Notice that BP is not the only company that made this error, it remains a preoccupation in health and safety management. This tragic confusion contributed to their own blindness and deafness. Indeed, process safety and major risks are recognised to be hard to measure in safety literature and practice. However, several efforts have been made by industries to define key performance indicators to benchmark, leading and lagging risk indicators, safety performance indicators. In process safety, a famous indicator is the "loss of containment" (LoC) that is a precursor of an accident (fire, explosion, toxic cloud), which defines the separation between prevention and protection. At Texas City, this indicator was measured and was degrading over time: "*the number of loss of containment incidents at the Texas City refinery increased each year from 2002 to 2004*" (Baker et al., 2007, p187)" "*with an increase of "52 percent from 399 to 607 per year*" (CSB, 2007, p168). These indicators were not in the main picture of SPI's monitored by BP management. They relied too much on worker safety performance indicators and were measuring an improvement in the lost time injuries statistical indicator. This indicator was among the key performance indicators of the management especially for attributing bonuses to managers (Hopkins and Maslen, 2015) and communicating to control authorities.

The temptation to use a measurable indicator is a common issue as recalled by (Kingston and Dien, 2017) which can lead to the 'McNamara Fallacy' which is attributed to Daniel Yankelovich (Smith, 1972) and is described in four steps:

- "The first step is to measure whatever can be easily measured. This is okay as far as it goes".
- "The second step is to disregard that which can't be measured or give it an arbitrary quantitative value. This is artificial or misleading".
- "The third step is to presume that what can't be measured easily really isn't very important. This is blindness".
- "The fourth step is to say that what can't be easily measured really doesn't exist. This is suicide".

More generally, the blindness process is more subtle. It can for instance come from over reliance of management tools and processes that put under the light some phenomenon leaving in the shadow some others, reinforcing "organisational blinkers" (Largier, 2008). They can produce "*an effect of blindness by producing an artefact of rationality. They participate to the setting on frontstage of a unique definition of the organisational situation, though other definitions are always present, but stay in the backstage*" (Boussard, 2003). Often "*the most used indicators give more consistency and resistance to some organisational representations*" (Boussard, 2001).

Listening to divergent opinions (as promoted by Navy Submarine in CAIB, 2003) and to 'bad news' (e.g. at Texas City, "*bad news was not welcome*", CSB, 2007) is not always easy especially for managers under pressure to achieve high

performance without adequate resources. Divergent opinions on the new NASA policy "*Faster, Better, Cheaper*" associated with cost-cutting, by new administrator Dan Goldin were dismissed: "*When critics would raise the possibility that such cuts were going to affect safety the CAIB notes "Goldin described himself as "sharp-edged" and could often be blunt. He rejected the criticism that he was sacrificing safety in the name of efficiency. In 1994 he told an audience at the Jet Propulsion Laboratory, "When I ask for the budget to be cut, I'm told it's going to impact safety on the Space Shuttle ... I think that's a bunch of crap.*" (CAIB, 2003). EWS come from alerts from staff, analysts and auditors who may provide another interpretation. While managers value coherence and coordination in action, for fostering foresight in safety they should value more the diversity of analysis. Indeed, in Cybernetics theories, researchers valued the diversity of views with Ashby's principle of requisite variety that implies a greater diversity of the controlling system to be able the control a complex system (Ashby, 1956; Vautier et al., 2018).

3.7   Discussion and conclusions

High-risk industries invest many resources every day in risk anticipation and many measures have already been in place for many years. However, the failures of foresight recalled here, highlight how industries, their experts and their regulators can fail. Every contributor to risk anticipation and foresight in safety should remain humble, cautious and sceptical and voice their doubts towards the challenge of accident prevention as it remains very difficult in practice. Anticipating and preventing accidents is a continuous struggle or never-ending war, bringing new changes, new risks and new threats, and also new safety degradation to discover before it is too late.

"Disasters *'are not created overnight'*" (Turner, 1976, 1978); accidents are therefore "*hard to obtain*" (Perrow, 1984); and require a "*rare conjunction of a set of holes in successive defences*"' (Reason, 1997). Accidents are not the result of one error but a combination of multiples causes, conditions and influence factors (Dien, 2006, ESReDA, 2009). Accidents develop (Guillaume, 2011) during an "*incubation period*' (Turner, 1976, 1978), which sometimes lasts for years (with "*latent defect*"(Reason,1997); in the example of San Bruno for more than 50 years). "*Latent conditions*"' and "*resident pathogens*" within the workplace and organizations are "*time-bombs*" that can be identified and removed before the event (Reason, 1997) but sometimes they are not. In contrast to Perrow's view

(1984), the majority of accidents are not inevitable (Dechy et al., 2012, Dien et al., 2013), because of the frequent occurrence of EWS prior to serious events, with some of them recognised by some actors. This empirical accident modelling makes clear the possibility of accident prevention. But will the opportunities to recognize an accident waiting to happen, be seized in the time window available? Indeed, some windows of opportunity and recovery (Edmondson et al., 2005) are recognised by some actors and require responses which are not always implemented in due time showing a form of apathy.

While high-risk industries devote time, money and analysts to identify hazards, assess risks, learn from early warning signs, near-misses and from others' hard lessons and best practices through benchmarking, many flaws in risk prevention are found in accident reports and sometimes in internal audits and event reports prior to the accident or in external regulatory inspections. These flaws are among the root causes of accidents. The few accident cases (Toulouse, Buncefield, Texas City, San Bruno, Challenger and Columbia) used as references for this analysis have highlighted some of the flaws in risk anticipation and prevention.

There are many techniques, tools and procedures to identify risk and assess related scenarios. Some of the methods have limits, domains of validity, and conditions for being adequately applied and used, but these are sometimes forgotten. For instance, are they adapted to address extreme events or black swans (Taleb, 2007)? Also, we find it necessary to fight the recent growing trend that defines high consequence/low probability as black swans, against which little or nothing can be done. Risk assessment may be incapable of thoroughly quantifying them, but this should not be taken as an excuse. Accidents are made up of a chain of events and focusing on what we already know and understand may help to break such chain and lower both the probability and severity of disasters.

Beyond the methods, implementation is dependent on the judgement of risk analysts. So, who are the analysts (Dien et al., 2012), what are their competencies, what are their collective resources to perform the job of conducting 'risk work' (Power, 2016) or risk expertise (Merad and Trump, 2020)?

Stark (1951) already claimed that foresight is partly reproductive and partly creative. Therefore, foresight in safety requires 'out-of-the-box' thinking and there are some processes to foster imagination better than with traditional risk analysis methods. In other words, although the use of techniques can bring a systematic approach useful to demonstration of safety management to a regulator,

identification of risks requires imagination and creativity. Identification of risks requires diversity in the ways of thinking when questioning 'what-if'. Diverse views can be shared in brainstorming including in debates over work practices, in 'speak-up' (Edmondson, 1999), listening to divergent opinions, listening to 'bad news' or to those who disagree even outside the industrial system with citizens, residents, consumers, NGOs (Dechy et al., 2016). In a systemic perspective, this diversity of views is a way to obtain requisite variety to control the system.

As risk is a social construct (Short, 1984) so is foresight in safety. Failures of foresight can be approached as a 'cultural collapse' because of the inadequacy of accepted norms and beliefs (Turner, 1976). A key implication is to remain critical on the processes of risk identification, risk assessment, performance monitoring to detect EWS. As identified by Clarke (1999), 'fantasy planning' may occur sometimes at the expense of actors' consciousness when stakeholders put too much confidence in their collective choices which rely on inadequate assumptions and are impacted by a multitude of biases and constraints. This can lead to the "*decoy problem*", focusing on well-defined problems rather than ill-defined problems which can lead to collective blindness (Turner, 1976).

"*Organizations are defined by what they ignore – ignorance that is embodied in assumptions – and by the extent to which people in them neglect the same kinds of consideration*" (Weick, 1998; p74). Engineering underlying assumptions are not often challenged during these processes. Expected scientific procedures are sometimes inadequate and may lead to inadequate beliefs from stakeholders. Risk management by companies and regulatory science are subject to criticism. Different values and goals may lead to controversies between regulators and high-risk industries but negotiations do also occur. Worst case scenarios are reduced to realistic scenarios and only reasonable changes after near-misses are made. Norms and standards which have been approved by expert groups, institutions are "normalised" (Vaughan, 1996) within the organizational culture and it becomes harder for those in the system to question and challenge. "Fresh eyes" or *Candide*, external auditors and investigators can help and this is known for decades. But, the challenge is for actors of the system to challenge themselves, their competencies, their tools, their assumptions which require some mindset shift, to become more than a sceptic (questioning or doubt attitude in safety culture concepts).

As remarked by Weick (1998, p72) about Turner's input on cultural failures of foresight, the issue is not only about world-views, lenses and paradigms. The

"*mastery of pattern generation with sufficient requisite variety to match and register the patterned variety in the complex events* […] *is best captured by the imagery of kaleidoscope': 'just as the image of switching lenses can represent the changing of patterns in the realist schema, the changing of turning a kaleidoscope can represent the changing of patterns in the subjectivist schema, since the patterns of kaleidoscope may be internally generated with minimal dependence on information from outside. Turning a kaleidoscope can: (1) dislodge old patterns, (2) generate new patterns and (3) foster awareness that numerous configurations are possible*" attributed by Weick to Nord and Connell (1993, p117).

This remark invites organisations to create spaces and times where diversity of views and thinking fosters 'requisite imagination' (Westrum, 1992), that can help to recognize patterns, share explicitly doubts and uncertainties about systems behaviour, to identify well-defined and ill-defined problems. The goal is also to understand the assumptions, the artefacts, the tools used, the constraints met by operators, engineers, experts, managers, regulators in conducting their 'riskwork' (Power, 2016) and even to put under questions and scrutiny the expert work and regulatory science (Vaughan, 1996; Llory 1996; Maguire and Hardy, 2013; Boudia and Demortain, 2014; Merad and Trump, 2020). Every study of risk can be considered as a project (Merad, 2010) that has inherent constraints in the resources. The goal or 'preoccupation with failure' (in HRO, Weick and Sutcliffe, 2007) is to wonder if 'safety imagination' highlights or hides risks (Pidgeon and O'Leary, 2000). Spaces may be self-organized informally by groups of engineers such as observed with the 'Debris assessment team' at NASA to characterise the foam strike consequences or institutionalized such as 'tiger teams' after Apollo 13 crisis. During the Columbia mission in space, the informal team was not given the status of a tiger team, and its conclusions were dismissed by mission' managers.

The challenge of foresight in safety is to identify all risks and recognize all EWS. But this is impossible in practice in general and at a given time. A reduced scope is to focus on major risks which implies the critical scenarios, those which escalate and damage system, assets and stakeholders. Some filtering of important signals is necessary otherwise channels are flooded with more and more data to treat. Making sense, prioritising are key processes to develop the relevant focus of resources with issues at risk, and proportionate the risk controls and implement them in due time. Decisions and trade-offs must be made, aided and revised. Safety margins and the burden of proof should be challenged. Especially, time is providing new opportunities to capture new signs and new knowledge to revise

assumptions and judgments. Dynamic learning and risk management approaches should be developed and promoted (ESReDA, 2015; Paltrinieri, 2015).

Often, people within the system recognised early warning signs before the accident (Turner, 1976, 1978). These recurring empirical findings reject the hindsight bias excuse (Reason, 1990; Vaughan, 1996; Woods, 2009). This does not mean to reject the risk of the hindsight bias. In Texas City accident, several actors and several processes (auditing, learning) recognised EWS of safety degradation (Dechy et al., 2011). It is clear that some signals are blurred, contradictory or are "mixed signals" (Vaughan, 1996). But organisations are not monolithic (Dien, 2014), and some workers, engineers and managers may know that safety is deteriorating. Many employees, whistle-blowers and citizens have warned before disasters (in chapter 11, Dien et al., 2020). In some case, beyond blindness and myopia pathologies, deafness, denial and apathy are major obstacles to change. The problem is not anymore, a problem of foresight but rather becomes managerial and political related to a lack of adequate reactions.

Echoing Weick's suggestion (1988) to define organisation by what they choose to ignore, organizations should also be defined by what they choose to remember and forget. Barriers and failures to learn are numerous. Accident and event reports often fail to address root causes (CAIB, 2003; Dien et al., 2012) and can themselves be considered as 'fantasy' documents (Birkland, 2009). There are losses of memory and similar accidents recur even in the same organizations (e.g. NASA, BP). Organizational patterns that lead to a failure of foresight are similar (Turner, 1976, 1978). In-depth analysis of accidents already provided the "hard lessons" to be learned especially from other industries (Dien et al., 2004). These lessons are part of an international history of industrial accidents, from which can be derived some "knowledge of accidents" (Dechy et al., 2010, 2016). It can provide useful frameworks to interpret EWS and organizational weaknesses in normal operations (Dechy et al., 2016, 2018) and can develop some specific attitudes, such as vigilance, doubt and prudence as components of safety culture, and 'preoccupation with failure' (Weick and Sutcliffe, 2007). The alternative (Reason, 1997) is that managers forget to be afraid and allow drift to occur. In summary, foresight in safety relies on exploitation of existing knowledge and resources and exploration mechanisms (related to innovation, changes) (March, 1991).

After all, practically speaking, what can be done? In addition to previous remarks and suggestions, one key factor to mention here is temporality. Foresight is fundamentally about time which highlights the dynamic nature of managing risk that is either improving or eroding. Time is potentially an enemy with pressures on decision and action but is also a resource as an opportunity to investigate and collect more information about an ill-structured problem, to help make a more objective judgment, to recalibrate a risk model with new data from the real world. Engineers have to make assumptions and decisions, but they have to remain sensitive to warning signs that would confirm or otherwise the safety envelope and margins. Managers are under business and time pressure to make decisions sometimes with insufficient information to understand all implications and side-effects (Ansoff, 1975). Decisions with their rationale and information should be formatted and recorded in order to be monitored with regard to new signals and effects of actions, changes or inactions. In high-risk industries with many risk management provisions, degradation of safety can be insidious, but is announced to some extent by EWS that may be recognised by some actors and may provide windows of opportunity to take a reactive action if only those in control are listening. Will these opportunities be seized or will the reaction be 'too little or too late'?

3.8   Acknowledgements

3.9   References

Ansoff, I. (1975), Managing Strategic Surprise by Response to Weak Signals, California Management Review, Vol.18 (2), pp.21-33.

Ansoff, I., Mc Donnell, E. (1990), Implanting Strategic Management, Second edition, Prentice Hall International, United Kingdom.

Ashby, W. R. (1956) An Introduction to Cybernetics, Chapman & Hall, London, 1956.

Baker, J., Bowman F., Erwin, G., Gorton, S., Hendershot, D., Leveson, N., Priest, S., Rosenthal I., Tebo, P., Wiegmann, D., Wilson, L. (2007), The Report of the BP U.S. Refineries Independent Safety Review Panel, http://sunnyday.mit.edu/Baker-panel-report.pdf.

Boudia, S., Demortain D. (2014), La production d'un instrument générique de gouvernement. Le "livre rouge" de l'analyse des risques, Gouvernement & Action Publique 3(3): 33–53, 2014.

Boussard, V. (2001), Quand les règles s'incarnent. L'exemple des indicateurs prégnants, Sociologie du Travail, 43 533-551.

Boussard, V. (2003), Dispositifs de gestion et simulacres de contrôle, In : V. Boussard, , S. Maugeri, (Eds), Du politique dans les organisations. Sociologies des dispositifs de gestion, l'Harmattan, Paris, pp. 173-191

Birkland, T. A., (2009), Disasters, Lessons Learned, and Fantasy Documents, Journal of Contingencies and Crisis Management, vol 17, N°3.

Carroll, J. S. (2004), Knowledge Management in High-Hazard Industries, Accident Precursors as Practice, in J. R Phimister., V. M Bier., H. C. (Kunreuther Eds), Precursor Analysis and Management, Reducing Technological Risk Through Diligence, National Academy of Engineering of the National Academies

Carroll, J. S., Fahlbruch B. (2011), The gift of failure: New approaches to analyzing and learning from events and near-misses, Safety Science, 49 1–4.

Clarke, L., (1999), Mission Improbable, Using Fantasy Documents to Tame Disasters, University of Chicago Press.

Clarke L., (2006), Worst Cases – Terror and Catastrophe in the Popular Imagination, University of Chicago Press.

Columbia Accident Investigation Board (2003), Report Volume 1, National Aeronautics and Space Administration, https://www.nasa.gov/columbia/home/CAIB_Vol1.html.

COMAH Competent Authority (2011), Competent Authority for the Control of Major Hazards, Health and Safety executive, Buncefield: Why did it happen? The underlying causes of the explosion and fire at the Buncefield oil storage depot, Hemel Hempstead, Hertfordshire on 11 December 2005, Edited by HSE, https://www.hse.gov.uk/comah/buncefield/buncefield-report.pdf

Couturier J., Bruna G., Tarallo F., Chanton O., Dechy N., Chojnacki E. (2016), Quelques considérations sur le risque résiduel dans l'industrie nucléaire, in M., Merad N., Dechy L. Dehouck, M. Lassagne , Risques majeurs, incertitudes et décision : approches pluridisciplinaire et multisectorielle, Editions ESKA,.

CSB (2007). Investigation Report, Refinery Explosion and Fire, BP – Texas City, Texas, March 23, 2005, Report N°2005-04-I-TX. 2007, https://www.csb.gov/bp-america-refinery-explosion/.

Cullen, W. D. [Lord] (2000), The Ladbroke Grove Rail Inquiry, Part 1 & Part 2 Reports, HSE Books, Her Majesty's Stationery Office, Norwich, 2000. [Part 2: 2001], https://www.jesip.org.uk/uploads/media/incident_reports_and_inquiries/Ladbroke%20Grove%20Rail%20Inquiry%20Report%20Part%201.pdf, https://www.railwaysarchive.co.uk/documents/HSE_Lad_Cullen002.pdf.

Dechy N., Bourdeaux T., Ayrault N., Kordek M.-A., Le Coze J.-C. (2004), First lessons of the Toulouse ammonium nitrate disaster, 21st September 2001, AZF Plant, France, Journal of Hazardous Materials 111 pp. 131-138

Dechy, N., Descourrière, S., Salvi O. (2005), The 21st september 2001 disaster in Toulouse : An historical overview of the Land Use Planning, 28th ESReDA Seminar on the Geographical component in risk management, Karlstad University, June 2005

Dechy, N., Rousseau, J.-M., Llory, M. (2011), Are organizational audits of safety that different from organizational investigation of accidents?, Proceedings of ESREL Conference, Troyes, , 18-22 September.

Dechy N., Rousseau J.-M., Jeffroy F. (2011), Learning lessons from accidents with a human and organisational factors perspective: deficiencies and failures of operating experience feedback systems, Proceedings of the EUROSAFE 2011 conference, Paris.

Dechy N., Mortureux, Y., Planchette, G., Blatter, C., Raffoux, J.-F. (2016) ; Explorer « l'imprévisible » : comment et jusqu'où ? actes de la conférence λμ20, Saint-Malo, France

Dechy N., Rousseau J.-M., Dien Y., Llory M. Montmayeul R., (2016) Learning Lessons from TMI to Fukushima and Other Industrial Accidents: Keys for Assessing Safety Management Practices, proceedings of IAEA Conference 30 Years of Safety Culture

Dechy, N., Gyenes, Z., Merad, M. (2018), A Review on Toulouse Accident Trials: Can We Learn the Lessons Despite Uncertainty on Direct Causes?Proceedings of the Hazards 28th conference, Edinburgh, U.K., Icheme.

Dechy N., Dien Y., Marsden E., Rousseau J.-M. (2018), Learning Failures As the Ultimate Root Causes of Accidents: Managing Errors in Organizations in 'How could this happen ? Managing error in organizations', Edited by Hagen J., Palgrave Macmillan

Dechy, N., Rousseau, J.-M., Largier, A., Tillement, S., Hayes, J., Journé B., (2018), Using the 'Knowledge of Accidents' in Normal Operation: a Case Study on the Framing of Organisational Diagnosis of Safety Management, Proceedings of the 55th ESReDA seminar on - Accident Investigation and Learning to Improve Safety Management in Complex Systems: Remaining Challenges, Romania, 9 – 10 October, AGIFER

Dien, Y., "Les signaux faibles à l'aune des lanceurs d'alerte" Congrès λμ19 de l'IMdR, Dijon, 21-23 octobre 2014.

Dien, Y. & Llory, M. (2004). Effects of the Columbia Space Shuttle Accident on High Risk Industries or: Can We Learn Lessons from Other Industries? Conference Hazards XVIII, November 23-25, Manchester, UK.

Dien, Y., Llory, M., Montmayeul, R. (2004). Organisational Accidents Investigation Methodology and Lessons Learned, Journal of Hazardous Materials, 111 (1-3), pp 147-153.

Dien, Y., Dechy N., Guillaume, E., (2012) Accident Investigation: from Searching Direct Causes to Finding In-Depth Causes. Problem of Analysis or / and of Analyst? Safety Science, 50 (6), , pp 1398-1407.

Dien Y., Dechy N., Llory M., (2013), Is the Complexity of Hazardous Socio-Technical Systems "Directly" Connected to Major Event Occurrence?, Proceedings of the American Nuclear Society National Meeting, November 2013, Washington, D. C. USA.

Dien Y., Dechy N., (2016), L'impensé est-il impensable ? Ce que nous apprennent les accidents industriels, in Merad M., Dechy N., Dechouck L., Lassagne M., Risques majeurs, incertitudes et décisions, Editions MA- ESKA – ISBN 978-2-8224-0430-3

Dien, Y., Maia, P., Paul S., Røed-Larsen, S., Stoop, J., Marsden, E., (2020), The Whistle-Blowers: Active Actors in Foresight for Safety, chapter 11 in "Enhancing Safety: The Challenge of Foresight", ESReDA.

Diet, (2012), "The National Diet of Japan - The official report of The Fukushima Nuclear Accident Independent Investigation Commission - Executive summary" http://warp.da.ndl.go.jp/info:ndljp/pid/3856371/naiic.go.jp/wp-content/uploads/2012/09/NAIIC_report_lo_res10.pdf, 2012.

Edmondson, A. C. (1999), Psychological Safety and Learning Behavior in Work Teams, Administrative Science Quarterly June 1999 vol. 44 no. 2 350-383

Edmondson, A., C., Roberto, M. A., Bohmer, R., M., J., Ferlins, E., M., Feldman, L., (2005), The Recovery Window: Organizational Learning Following Ambiguous Threats, in W., H. Starbuck, , and M. Farjoun, , Organization at the limit, Blackwell publishing.

ESReDA (2005), editors Roed-Larsen S. R., Stoop J., Funnemark E, Shaping public safety investigation of accidents in Europe, DNV publishing, ISBN 82 5150304 3

ESReDA (2009), Guideline for Safety Investigation of Accidents, ESReDA Working Group on Accident Investigation, (2009), disponible sur le site www.esreda.org

ESReDA (2015) Case study analysis on dynamic learning from accidents - The ESReDA Cube, a method and metaphor for exploring a learning space for safety, edited by Van der Vorm and the ESReDA Project Group on Dynamic Learning Available from www.esreda.org

ESReDA (2015), Barriers to learn, edited by Marsden E. and the ESReDA Project Group on Dynamic Learning. Available from http://www.esreda.org.

Fayol, H. (1916) General and Industrial Management. Institute of Electrical and Electronics Engineering, Paris.

Ferjencik, M. and Dechy, N. (2016). Three accidents in European dynamite production plants: An attempt to improve the external lessons learning. Journal of Loss Prevention in the Process Industries 44, 12-23.

Ferjencik, M., Dechy N. (2020), Loss of Memory as a Cause of Failure of Foresight in Safety, chapter 4 in "Enhancing Safety: the Challenge of Foresight", ESReDA

Ferjencik M., Strucic M., Tulonen T., Marsden E., Stoop J. (2020), Use of Scenarios as a support of foresight in safety, chapter 5 in "Enhancing Safety: the Challenge of Foresight", ESReDA

Feynman, R. (1986), Personal Observations on Reliability of Shuttle, Report of the PRESIDENTIAL COMMISSION on the Space Shuttle Challenger Accident, Volume 2: Appendix F.

Geertz, C (1998), La description dense, in: La description, revue enquêtes, n°6, Ed. Marseilles, pp73-105

Guillaume E. (2011) Identifying and Responding to Weak Signals to Improve Learning from Experiences in High-Risk Industry, PhD doctoral dissertation

Gyenes, Z. and Dechy, N. (2016), Risk and safety management of ammonium nitrate fertilizers: keeping the memory of disasters alive, Loss Prevention Bulletin n°251, October 2016, pp32-36, Edited by Icheme.

Hayes J., Hopkins A., (2014) Nightmare Pipeline Failures: Fantasy Planning, Black Swans and Integrity Management, Edited by Wolters Kluwer, CCH,

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

Hollnagel E. (2014), Safety-I and Safety-II, The Past and Future of Safety Management, Ashgate.

Hopkins, A., (2001) 'Was Three Mile Island a Normal Accident?', Journal of Contingencies and Crisis Management, vol. 9, no. 2, pp. 65-72.

Hopkins, A. (2010). Failure to learn: the BP Texas City Refinery Disaster, CCH Australia Ltd. Wolters Kluvers.

Hopkins, A., Maslen, S., (2015), Risky Rewards: How Company Bonuses Affect Safety, Edited by CRC Press, Taylor and Francis.

Hourtoulou, D. (2002), Projet ASSURANCE – Assessment of Uncertainties in Risk Analysis of Chemical Establishments, E. C. project ENV4-CT97-0627. rapport final d'opération a – DRA-07 – INERIS – Dho-2002-26824, 2002.

Hudson P. (2001), Safety Culture - Theory and Practice, RTO HEM Workshop on "The Human Factor in System Reliability - Is Human Performance Predictable? ", Siena, Italy, 1-2 December 1999, published in RIO MP-032.

Jouniaux P., Hadida D., Dechy N., Marle L., Billy F., Pierlot S., Parrennes F., Rouvière G., Husson D., (2014), "Détection, pertinence et amplification des signaux faibles dans le traitement du retour d'expérience", Congrès λμ19 de l'IMdR, Dijon, 21-23 octobre.

Kingston J., Dien Y. (2017). The McNamara Fallacy Blocks Foresight for Safety, Proceedings of the 53rd ESReDA Seminar "Enhancing Safety: the Challenge of Foresight" Edited A. L. Vetere Arellano, Z. Šimić, N. Dechy, 14-15 November 2017, Ispra, Italy

Kletz, T. (1993). Lessons from Disasters: How Organizations Have No Memory and Accidents Recur. Gulf Publishing Company, Houston. ISBN: 0-88415-154-9, 183 pages.

Lagadec P. (1994). Apprendre à gérer les crises. Les éditions d'organisation. Paris.

Lannoy A., (2016) Limites, insuffisances et apports des approaches probabilistes actuelles, quelles leçons tirer?. in M.Merad, N. Dechy, L.Dechouck, M Lassagne., Risques majeurs, incertitudes et décisions, Editions MA- ESKA – ISBN 978-2-8224-0430-3

Laporte, T.R. and Consolini, P.M. (1991), Working in Practice But Not in Theory: Theoretical Challenges of "High-Reliability Organizations" , Journal of Public Administration Research and Theory, vol.1, n°1, p 19-47

Largier, A. (2008). Dispositif de gestion des compétences et logique métier. Socio-logos n°3. https://journals.openedition.org/socio-logos/1323

Lesca, H. (2001), Veille stratégique : passage de la notion de signal faible à la notion de signe d'alerte précoce, Proceedings of the VSST Conference, Barcelona, Spain.

Lechaudel, J.F., Mouilleau, Y. (1991), Assessment of an accidental vapour cloud explosion - A case study: Saint Herblain, October 7 1991, France", Proceedings of the 8th International Loss Prevention Symposium, pp 333-348, Antwerp, Belgium, 1995.

Leveson N. G., Turner C. S. (1993), An investigation of the Therac-25 Accidents, Computer, 26, 7, pp18-41

Libmann, J. (1996), Éléments de sûreté nucléaire, Édité par l'Institut de Protection et de Sûreté Nucléaire. 1996.

Linkov I., Moberg E. (2017), Multi-Criteria Decision Analysis - Environmental Applications and Case Studies, CRC Press, Taylor and Francis,

Llory, M. (1996) Accidents industriels : le coût du silence, Opérateurs privés de parole et cadres introuvables, Éditions L'Harmattan,.

Llory, M. (1999),L'accident de la centrale nucléaire de Three Mile Island", Éditions L'Harmattan, .

Llory, M. Dien, Y. (2006) Les systèmes sociotechniques à risques : Une nécessaire distinction entre fiabilité et sécurité, Performances n°30 pp 20-26, n°31 pp 9-13, n°32 pp. 20-26,.

Llory, M., Montmayeul, R. (2010) L'accident et l'organisation, Éditions Préventique,.

Llory, M. & Montmayeul, R. (2018). Comprendre l'accident, la pratique de l'analyse organisationnelle de sécurité, Editions L'Harmattan.

Maguire, S., & Hardy, C. (2013), Organizing Processes and the Construction of Risk: A Discursive Approach. Academy of Management Journal, 56 (1), 231-255.

March, J. (1991) Exploration and Exploitation in Organizational Learning, Organization Science, Vol. 2, No. 1, pp. 71-87

Marlair, G. and Kordek M-A. (2005) Safety and security issues relating to low capacity storage of AN-based fertilizers, Journal of Hazardous Materials Volume 31;123(1-3), pp 13-28.

Marsden, E., Dechy, N., Vetere Arellano, A., (2020) Big data analytics and early warning signs, chapter 10 in "Enhancing Safety: the Challenge of Foresight", ESReDA.

Médard, L. (1979) Les explosifs occasionnels, deuxième édition ; Technique et Documentation, Éditions Lavoisier, Volume 2.

Merad M. (2010) Aide à la décision et expertise en gestion des risques, Editions Lavoisier

Merad, M., Dechy N., Dehouck L., Lassagne, M. (2016) Les décisions face aux risques majeurs, retours d'expérience et pistes d'améliorations, in M Merad., N Dechy., L. Dehouck, M. Lassagne, Risques majeurs, incertitudes et décisions, Editions MA- ESKA – ISBN 978-2-8224-0430-3

Merad M. and Trump B. D., (2020) Expertise Under Scrutiny 21st Century Decision Making for Environmental Health and Safety, Springer.

Merritt, C. (2007) Testimony of Carolyn W. Merritt Chairman and Chief Executive Officer U.S. Chemical Safety Board Before the U.S. House of Representatives Committee on Energy and Commerce Subcommittee on Investigations and Oversight, May 16, 2007.

MIIB (2008) The Buncefield Incident 11 December 2005, The final report of the Major Incident Investigation Board, Volumes 1 & 2. Buncefield Major Incident Investigation Board, Richmond, Surrey.

NTSB. (2011). Pipeline Accident Report: Pacific Gas and Electric Company Natural Gas Transmission Pipeline Rupture and Fire, San Bruno, CA, September 9, 2010.

Paltrinieri, N., Dechy, N., Salzano, E., Wardman, M., & Cozzani, V. (2012) Lessons Learned from Toulouse and Buncefield Disasters: From Risk Analysis Failures to the Identification of Atypical Scenarios Through a Better Knowledge Management, Journal of Risk Analysis, 32(8). pp 1404-1419.

Paltrinieri N., Dechy N., Salzano E., Wardman M., Cozzani V. (2013) Towards a new approach for the identification of atypical accident scenarios, Journal of Risk Research, 16:3-4, 337-354.

Paltrinieri, N., Khan, F., Cozzani, V., (2015) Coupling of advanced techniques for dynamic risk management, Journal of Risk Research, 18:7, 910-930,

Perrow, C. (1982) The President's Commission and the Normal Accident, , in D. L. Sills, C.P. Wolf, & V. B Shelanski, (Eds), Accident at Three Mile Island: The Human Dimensions, Westview Press, Boulder, Colorado (1982). 2. p. 173-184

Perrow, C., (1984) Normal Accidents – Living with High Risk Technologies, Princeton University Press,.

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

Power, M. (2016) Riskwork: Essays on the Organizational Life of Risk Management. Introduction." In: M. Power (ed.): Riskwork: Essays on the Organizational Life of Risk Management. Oxford, UK: Oxford University Press, 1–25.

Rasmussen, N. C. et al. (1975) "Reactor safety study. An assessment of Accident risks in U. S. commercial nuclear power plants. Executive Summary". WASH-1400 (NUREG-75/014). Rockville, MD, USA: Federal Government of the United States, U.S. Nuclear Regulatory Commission.

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

Reason, J. (1990) Human Error, Cambridge University Press.

Reason, J. (1997) Managing the Risks of Organisational Accidents, Ashgate, Aldershot.

Roberts, K. H. (1990) Some Characteristics of High-Reliability Organizations. Organization Science, 1, 160-177

Rogers, W. P. et al., (1986), Presidential Commission on the Space Shuttle Challenger Accident (1986), Report to the President by the Presidential Commission on the Space Shuttle Challenger Accident, Government Printing Office, Washington DC.

Roux-Dufort C. (2003) Gérer et décider en situation de crise : Outils de diagnostic, de prévention et de décision. Dunod. Paris

Simon, H. A. (1973) The Structure of Ill Structured Problems, Artificial Intelligence, 4 (3), 1973, p 181-201.

Short, J., F., (1984), The Social Fabric at Risk: Toward the Social Transformation of Risk Analysis, American Sociological Review, Vol. 49, No. 6 , pp. 711-725

Stark, S., (1961) Executive Foresight: Definitions, Illustrations, Importance, The Journal of Business, Vol. 34, No. 1, (Jan., 1961), pp. 31-44, Published by: The University of Chicago Press

Taleb, N. (2007) The Black Swan: The Impact of the Highly Improbable, Random House,

Thellier, S. (2017) Approche ergonomique de l'analyse des risques en radiothérapie : de l'analyse des modes de défaillances à la mise en discussion des modes de réussite. Thèse de doctorat en ergonomie. Conservatoire National des Arts et Métiers, Paris, soutenue le 12 décembre 2017, 294 p.

Thellier S., Jeffroy F., Cuvelier, L., & Falzon, P. (2018) L'analyse des risques en radiothérapie : quelle alternative à l'AMDEC ? Congrès λμ21, 16 – 18 octobre 2018, Reims, France.

Turner B. (1976), The Organizational and Interorganizational Development of Disasters, Administrative Science Quarterly, Vol. 21, No. 3 , pp. 378-397

Turner, B., & Pidgeon, N. (1997), Man-Made Disasters, Second edition, Butterworth Heinemann [1rst edition: Turner, B. (1978), Wykeham Publications].

Tversky, A., Kahneman D. (1974), Judgment under Uncertainty: Heuristics and Biases, Science, Vol. 185, Issue 4157, pp.1124-1131

Vaughan, D., (1996) The Challenger Launch Decision - Risky Technology, Culture, and Deviance at NASA, The University of Chicago Press.

Vaughan, D. (1999) The Dark Side of Organizations: Mistake, Misconduct, and Disasterr. Annual Review of Sociology, n°25, pp 271-305

Vautier, J.-F., Dechy N., Coye de Brunélis, T., Hernandez, G., Launay, R., Moreno Alarcon, D., P. (2018) Benefits of systems thinking for a human and organizational factors approach to safety management, Journal of Environment System and Decision, 38:353-366

Wagenaar W., A. and Groeneweg, J. (1987) Accidents at Sea: Multiple Causes and Impossible Consequences, International Journal of Man-Machine Studies, 27, pp587-598

Weick, K., E., (1991) The vulnerable system: an analysis of the Tenerife air disaster, in P. J. Frost, et al., eds, Reframing organizational culture, London, Sage Publications.

Weick K. E. (1998) Foresights of Failure: An Appreciation of Barry Turner, Journal of contingencies and crisis management, volume 6, number 2

Weick, K. and Sutcliffe, K. M. (2007) Managing the Unexpected: Resilient Performance in an Age of Uncertaintyy, Second Edition, Jossey Bass.

Westrum, R., (1992) Cultures with Requisite Imagination, in J.Wise, , D. Hopkin,, P. Stager, (eds), Verification and Validation of Complex Systems: Human Factors Issues: Human Factors Issues, Berlin, Springer Verlag, pp 401-16

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, Elsevier Science Ltd, pp 235-248.

Woods D. (2009) Escaping failures of foresight, Safety Science 47 (2009) 498–501

4 Loss of Memory as a Cause of Failure of Foresight in Safety

Milos Ferjencik, University of Pardubice, Czech Republic
Nicolas Dechy, Institut de radioprotection et de sûreté nucléaire, France

4.1 Executive Summary

Loss of memory that degrades foresight in safety, especially implies a loss of knowledge of Early Warning Signs. With this focus, four aspects of memory playing a crucial role in foresight are identified which enable four different abilities (awareness for vigilance, recognition for sense-making, investigation for detection, and follow-up). They are associated with three faculties of human memorisation (encoding, storing, retrieving). Thus, difficulties for maintaining the memory useful for foresight fall into twelve categories of loss of memory. The individual aspects of memory important for foresight are distributed among different humans on different levels of hierarchy in socio-technical system. Organisational memory is a complex concept, phenomenon and system which relies on several aids, functions and artefacts (documents, procedures, processes, structures). Safety management measures can reduce some kinds of loss of memory without duplicating efforts but organisations should direct proactive specific measures against relevant categories of loss of memory.

4.2 Key Messages

- Loss of memory is a recurring root cause of disasters.
- Knowledge from the past, hindsight knowledge is still useful for insight and foresight.
- Part of the foresight is 'reproductive' based on past experiences while another part is more 'creative'.
- Loss of memory represents a loss of knowledge about early warning signs resulting from learning deficiencies.

- Four aspects of memory playing a crucial role in foresight have been identified and enable different abilities (awareness for vigilance, recognition for sense-making, investigation for detection, and follow-up).
- There are three key faculties of human memorisation (encoding, storing, retrieving).
- Loss of memory useful for foresight can fall into twelve categories defined by the combination of four aspects and three faculties of memory.
- This classification applies not only to human memory but also to organisational memory.
- Organisational memory is a complex process that integrates memories of all humans within the organisation and sustain collective memory, supporting safety management aids and functions, and all the artefacts (documents, procedures, structures, processes) that integrate the experience of designers, operators and managers.
- For each part of memory, contributing persons, aids and devices and supporting functions of the local safety management could be identified.
- Some of generic safety management measures may prevent loss of memory and there is no need to duplicate efforts. However, there is still an effort to convert the question of what increases or decreases loss of memory to the measures or specific functions in local safety management that sustain overall safety performance including the memorisation aspects and faculties.

4.3 Introduction

It may seem paradoxical to address foresight, while its perspective is towards the future, with an issue such as "loss of memory" which perspective is towards the past. Indeed, knowledge from the past, or hindsight knowledge is still useful for insight and foresight. To face risks in the future, we are convinced that history still matters even with the forthcoming era of big data and artificial intelligence which implies duties for establishing some safety history to remember.

First and as a reminder according to Stark (1961), part of the foresight is 'reproductive' based on past experiences while another part is more 'creative'. For Ricoeur analysing the relationship between history, memory and forgetting (2000), "the important is that the projection towards the future is now solidary to the retrospection on the past times". In summary (Kingston and Dien, 2017), foresight

is about imagining the future possibilities based on knowledge of the past and present. In chapter No.°5 dedicated to the use of scenarios (Ferjencik et al, 2020), we have explained how the use of scenarios for foresight in safety is both prospective and retrospective.

Second, one should remind that our society that is currently driven by 'creative destruction' (Schumpeter, 1942) which puts an emphasis on innovation and adaptation and highlights how experience and knowledge can become obsolete in face of changes that accelerate. Indeed, in management science, Abernathy (1978) proposed a paradigm shift that considered that innovation and experience should be understood as rival forces and strategies for companies. This dilemma between the two rationales was later reformulated by March (1991) within organisational learning framework: one would relate to the exploration of new knowledge while the other would relate to the exploitation of the existing knowledge. However, one should acknowledge that not all changes are disruptive, some are incremental and some designs last for several years. We believe foresight in safety has to consider both dimensions as it is done in the whole book. In this chapter, we focus more on the failure to use existing safety knowledge.

Third, the quote 'Loss of Memory' (LoM) is a fairly common explanation of the reasons or the way an unwanted event recurred or occurred despite early warning signs.

In the context of this ESReDA book on the issue of Foresight in Safety, we limit our scope to situations where LoM is used to explain some failures of foresight (Turner, 1976; Kletz, 1993; Llory, 1996; Vaughan, 1996; Reason, 1997; CAIB; 2003; Woods, 2009, Dechy et al, 2011, Ferjencik and Dechy, 2016). These are situations where LoM represented a failure of foresight, or memory loss has caused some 'Early Warning Signs[16]' (EWS), 'weak-signals', 'latent flaw', 'precursors of accidents', 'near-misses', or 'alerts' to be omitted or unrecognised, during an 'incubation period' (Turner, 1978), which created an 'accident waiting to happen'.

In other words, loss of memory is a root cause of numerous disasters. Indeed, the situations we may encounter and should explain by the term LoM are diverse. Loss of Memory could actually be considered itself as a generic and transversal kind of EWS. However, it is a term that remains so vague and broad that different forms of LoM need to be analysed.

To address the challenge of making the concepts 'foresight in safety', 'EWS' and 'LoM' operational, we therefore attempt in this chapter to identify different forms of LoM and classify them. The resulting categorization can help to highlight a few specific LoM examples to better focus repetitive prevention efforts through safety management.

In this analysis of LoM forms or patterns, we assume first that LoM represents a loss of knowledge about EWS resulting from learning deficiencies (as described in Part 4.4). Accordingly, LoM forms can be divided into four groups which relate to several (in-)abilities (lack of awareness that leads to lack of attention or vigilance, lack of recognition or sense-making abilities, lack of investigation abilities for detecting EWS, and lack of follow-up after EWS). This will be explained in Part 4.4.3.The second step of the analysis (in Part 4.5) starts from the position that the term LoM tends to relate to human memory and can be approached as a process. Some definitions on human memory is provided. This leads to a distinction between three types of deficiencies in encoding, storing and retrieving information and knowledge; this represents the second dimension of the proposed classification scheme as later explained in Part 4.5.2. Combining these two dimensions together, it leads us to propose a framework of twelve categories of risks for LoM to manage (see Part 4.5.4).

In general, however, LoM does not only concern human and individual memory. For about the third generation we already live in a world where the term memory is commonly used outside the description of the human mind as a component of machines and systems and ultimately as a component of socio-technical systems (Rasmussen, 1997). Today we are not surprised to use the term loss of group memory, collective and also organisational memory - see Stemn et al. (2018). The legacy of Kletz (1993) "lessons from disasters: how organisations have no memory and accidents recur" was a triggering key message for this chapter and especially a more controversial warning that "organisations have no memory, only people

---

by someone in the system. Not all EWS are signalled by somebody. Weak or strong signals of danger can be either amplified (change of risk models and actions taken strengthen risk management measures) or weakened (no action to reassess risk assessment assumptions or to take provisions for risk management). See Part 4.4.2 for further developments and relationships between concepts.

have". This statement is therefore discussed towards foresight in safety challenges. An analysis of this 'organisational memory' term will lead us to explain the paradoxical statements about LoM that we may come across (in Part 4.6). This will complement our knowledge of LoM forms.

At the end of the chapter (in Part 4.7), we consider what can counter LoM. A few approaches and tools within safety management are shown that could be used to prevent different categories of LoM.

Though we will provide insights from different scientific literatures, our goal was still to propose to safety practitioners (analysts, managers) some guidance to tackle this challenge and to make some links with management of safety. Some industrial and day-to-day examples will be used for the purpose as well.

4.4   Loss of Memory Relates to Learning and Knowledge

4.4.1   Failures of foresight in safety due to a loss of memory

Several disasters occurred or recurred and highlighted various kinds of loss of memory. Former astronaut Dr. Sally Ride who was a member of both NASA space shuttle accident investigation teams observed that there were "echoes" of Challenger accident (in 1986) in Columbia accident (in 2003). Indeed, the CAIB noticed (2003) that "The foam debris hit was not the single cause of the Columbia accident, just as the failure of the joint seal that permitted O-ring erosion was not the single cause of Challenger. Both Columbia and Challenger were lost also because of the failure of NASA's organisational system" (CAIB, 2003, p. 195). These deficiencies to learn and to correct organisational failures have been pointed out by the CAIB: "First, despite all the post-Challenger changes at NASA and the agency's notable achievements since, the causes of the institutional failure responsible for Challenger have not been fixed. Second, the Board strongly believes that if these persistent, systemic flaws are not resolved, the scene is set for another accident. Therefore, the recommendations for change are not only for fixing the Shuttle's technical system, but also for fixing each part of the organisational system that produced Columbia's failure" (CAIB, 2003, p. 195).

Moreover, we can also mention the premonitory conclusion in 1996 of her book about the study of the Challenger accident by the sociologist Vaughan: "After the Challenger disaster, both official investigations decried the competitive pressures and economic scarcity that had politicized the space agency, asserting that goals and resources must be brought into alignment. Steps were taken to assure that this happened. But at this writing, that supportive political environment has changed. NASA is again experiencing the economic strain that prevailed at the time of the disaster. Few of the people in top NASA administrative positions exposed to the lessons of the Challenger tragedy are still there. The new leaders stress safety, but they are fighting for dollars and making budget cuts. History repeats, as economy and production are again priorities" (Vaughan, 1996, p. 422).

Unfortunately, NASA is not the only organisation that was responsible for recurring accidents. BP had several accidents with "striking similarities" in their organisational root causes (CSB, 2007; Merritt, 2007): fires and explosions in Grangemouth refinery in Scotland (in 2000), explosion and fires in Texas City refinery in USA (in 2005), pipelines leaks in Prudhoe Bay in Alaska (in 2006), Deepwater Horizon oil rig blow-out at Macondo in Gulf of Mexico (in 2010). They were preceded by several early warning signs, some of them were recognised and signalled by some workers and managers but also within internal audits, but the follow-up was poor leading to "too little, too late" reactions (Merritt, 2007).

This phenomenon of loss of memory in relationship with failures to learn (Hopkins, 2010) and limits in enhancing organisational learning (Dechy et al, 2011; ESReDA, 2015) should not be restricted to one organisation or one worldwide company with many sites. It should be addressed transversally across countries and across sectors as well.

The ammonium nitrate industry faced a long history of accidents since the beginning of 20th century (Gyenes and Dechy, 2016) and the recent tragedy in Beyruth these last days is recalling that some lessons of the most recent ones (e.g. Toulouse disaster in 2001 in France (Dechy et al, 2004) and West major accident in Texas in 2013 (CSB, 2016)) are still not learned by different stakeholders (regulator, government, judicial department, operator, subcontractor). They all showed a low awareness of explosion danger and inherent risks of ammonium nitrate fertilizers but also some lack of follow-up after EWS or alerts. Similarly, within the explosive manufacturing industry in Europe, the analysis of three accidents showed some missed opportunities for improvement after near-misses, external accidents in other countries and regulation changes, but also a loss of designers' intentions throughout the lifetime of operation of the plant, especially on critical incident scenarios (Ferjencik and Dechy, 2016). Similar observations about losses of

awareness and follow-up of previous lessons from accidents have been made on Buncefield disaster (in 2005), as similar explosions occurred in history prior to Buncefield and even recurred after (Paltrinieri et al, 2012) (see the list of accidents in chapter 3).

4.4.2 Early warning signs and foresight in safety

Early warning sign (EWS) is an event or more generally a condition that if it is not detected and acted upon can lead or contribute to an incident. EWS is a broad term that covers several others such as weak signals (Vaughan, 1996, Llory, 1996), precursors of accidents (Carroll, 2004), near-misses, latent flaw (Reason, 1990) and extends to alerts made by operators, safety analysts, managers, inspectors or whistle-blowers.

The word 'early' means that it can be early detected, prior a more severe accident, during an 'incubation period' (Turner, 1978). If EWS are correctly recognised, they can inform of 'time-bombs' in the system (Reason, 1997). Early warning signs should be detected as early as possible in order to take proper actions in due time to prevent or mitigate incident. This is one of the main purpose of foresight in safety.

Center for Chemical Process Safety (CCPS, 2012) writes about incident warning signs which are subtle indicators of a problem that could lead to an incident. Warning signs precede incidents or contribute to them as a condition.

While some EWS are weak, subtle, ambiguous, and may raise some uncertainties and are difficult to hear and filter from the background noise, some are strong signals such as near-misses, incidents, deviations which are normalised (Vaughan, 1996) or weakened (Guillaume, 2011), or external accidents lessons that should be learned. Some strong signals can be weakened and can lead to issues in the follow-up of actions, which are sometimes 'too little, too late' as remarked by US CSB former Chair C. Merritt (2007) about Texas City refinery accident in 2005 (CSB, 2007).

In conventional risk terminology, early warning sign can be understood as an indicator of hazard strengthening or of weakening of a safety measure (control) or of an increase of the vulnerability of the targets, which can result in an increase of likelihood of occurrence or in an increase of potential severity of consequences of scenarios causing damage. Simply said, we can consider that an early warning sign is an indicator of increase of risk.

The relationship between data, information and risk is sometimes not easy to establish at all. To recognise some EWS, investigation and data collection are necessary to produce pieces of information. However, all data, information and signs cannot be called EWS. An interpretation effort is required and implies to filter from the background noise, select some signs (Lesca, 2001; Rouibah and Ould-ali, 2002) and link several pieces of information "such as in a puzzle"[17] to discover a hidden form or pattern. This pattern that links several information to a risk is considered as a safety signal. This sense-making can be performed by a worker 'close to the valves' (CSB, 2007) or a manager or an analyst at different position in the sociotechnical system. This interpretation is then challenged towards its links to the risk or even to a scenario (Jouniaux et al, 2014). This process involves individual and collective expertise, which relies on knowledge (e.g. models, theories and studies of risk but also on frameworks, stories and patterns) that are learned and memorised.

4.4.3 Four aspects of memory useful for foresight in safety

Such a way, foresight in safety means capability to indicate increase of risk with the help of EWSs. In order to be able to implement foresight in safety, this capability has to be conscious and it has to be acted accordingly.

Speaking about the ability to indicate EWSs, four conditions related to memory have to be maintained that are shown in Table 1.

The four conditions from Table 1 describe four aspects of memory considered to be a necessary instrument for foresight in safety. In our view, memory plays a crucial role in foresight in safety as much as imagination addressed in other chapters of the ESReDA book. In our definition or proposal, 'loss of memory' in any of above aspects of memory corrupts foresight in safety.

---

[17] *Metaphor of puzzle for recognizing EWS in strategic management is attributed to several researchers (Rouibah and Ould-ali, 2002). Those authors propose to consider the concept of early warning sign rather than weak signal.*

Table 1. Four aspects of loss of memory related to foresight in safety.

N°	Aspect of memory to be maintained	Ability
1	Remember that such indicators of safety problems (EWSs) can arise	Awareness, attention, vigilance
2	Remember what forms of EWSs are possible (in order to be able to identify them)	Sense making, recognise
3	Remember how presence of EWSs can be detected	Method to detect
4	Remember EWSs that were detected until they are reasonably responded (in order to be able to respond)	Follow-up

Indeed, memory useful for foresight in safety means knowledge of EWSs that can be activated with abilities to be vigilant, to make sense, to investigate and to follow-up EWS. It does not matter how the knowledge of EWSs has been acquired. In any case, knowledge of EWSs is a result of lessons learning. Such a way, loss of memory relates to unlearning and forgetting. Lessons learning may be based on use of scenarios (including stories), both retrospective and prospective (see in Chapter 6 dedicated to the use of scenarios, Ferjencik et al, 2020).

4.4.4   Applied example: the kitchen dangers
To explain our proposal and related definitions, the four aspects mentioned in Table 1 can be easily illustrated using the example of Kate and William in the kitchen (Kate and William are married; they have sometimes slightly different views on what is dangerous in their kitchen; see Chapter 6 (Ferjencik et al, 2020)).

The first aspect is related to a situation when both Kate and William do not recognise their kitchen to be a dangerous place where EWSs may warn against possible incidents. They do not remember (they have not learned) that EWS can arise in their kitchen. If a person in the kitchen is not 'aware' that frying oil is combustible and frying can cause a fire, then it is natural that this person is not 'vigilant', and does not 'pay attention', because he or she does not understand that leaving the frying pan on a hot ceramic hob unattended is an early warning sign. This person probably suffered a loss of the first aspect of memory.

The second aspect can be illustrated by a situation when they both understand a possibility of fire but do not recognise that a bottle of oil standing at the hob may be a problem. A person in the kitchen who realizes that the oil is combustible and hence the pan cannot be left unattended during frying, but does not realize that

also the presence of a plastic bottle with the oil in the vicinity of hob can contribute to the fire, suffered a loss of the second aspect of memory.

The third aspect can be described as a situation when both William and Kate understand that a decreased capacity of safety valve on pressure cooker is an EWS but they do not know that or how the presence of this EWS can be detected. In this case, they suffered a loss of the third aspect of memory.

The fourth aspect means that they know about the above bottle of oil or about degraded pressure cooker safety valve but have given up all the attempts to improve the situation. A situation where a person recognizes that the presence of plastic bottle with the oil in the vicinity of hob is an EWS, but then places the bottle on the same place again, can be related to the loss of the fourth aspect of memory.

4.4.5   Other examples from industry
The first aspect of memory means that people who are important to the system are aware of facing a hazard, etc., and that, therefore, some EWSs exist at all. It seems trivial, but this triviality is sometimes forgotten in industry. It can be illustrated by the case of West in Texas in 2013 described by CSB (2016). People who are important to the behaviour of the system have not realized that their storage of ammonium nitrate represents a hazard. Therefore, it is understandable that even repeated warnings about electrical installation faults were not considered to be EWSs related to this hazard.

We can imagine examples very similar to those from kitchen in a chemical laboratory. For instance a situation in a lab where a flammable liquid is used that is volatile and has a specific odour. Odour in the lab however is not considered to be an EWS.

Other examples have been briefly described in Part 4.4.1 and cover the four aspects of loss of memory.

4.4.6   The four aspects of LoM: description related to hazards
As we earlier defined, an early warning sign is the result of an interpretation process. An EWS can be recognised as a sign of the strengthening of a hazard or the weakening of a control over the hazard or an increase of vulnerability in the targets exposed to hazards. As such, if proper studies are conducted, it can be determined as a cause of causal event (explanation see Chapter 6, Ferjencik et al, 2020) or as an indicator of causal event or of a cause of causal event.

Differences in four aspects of memory can be well explained when we realize that EWSs are associated with the existence of hazards, controls of hazards and environmental conditions that affect hazards within the system and the exposed targets. EWSs can be recognised as signals of adverse changes in hazards, controls of hazards or environmental conditions affecting hazards or of vulnerability of targets.

1. Loss of the first aspect of memory means loss of knowledge that certain hazards and relevant controls and environmental conditions are present in the system. If one is not aware, then it's natural not to know, not to pay attention and to become vigilant about EWSs that can signal their unfavourable changes.
2. The second aspect of memory can be degraded in situation where hazards, controls, and environmental conditions are known, but there is a lack of knowledge of some EWSs that may signal their degradation. The sense making abilities are poor enough and do not allow a proper recognition of the EWS.
3. Loss of the third aspect of memory would be identified if EWSs signalling adverse changes in hazards, controls of hazards or environmental conditions affecting hazards are known but means enabling their detection are not known (improper data collection, measurement of wrong indicators).
4. The fourth aspect of memory is damaged e.g. in a situation where hazards, controls, and environmental conditions are known, and the occurrence of some EWSs indicating their degradation has been identified, but the existence of these EWSs had been routinized, normalised and forgotten before they could be responded.

It is visible from the above explanations that the loss of the first aspect of memory is close to what is described as a loss of *sense of vulnerability* (see e.g. CCPS, 2007). Reason (1997) recall that sometimes, especially after success or when preoccupation with production goals is too high, actors can lose sight of dangers and forget to be afraid. Similarly, the loss of the fourth aspect of memory is close to the *normalisation of deviance* (Vaughan, 1996) or *standardization of deviance* (see e.g. Rosen, 2015).

4.5 The Process of Loss of Memory

Foresight in safety is to be realized in socio-technical systems by people. In a first step (in this part), the memory of humans working in socio-technical system is considered to be crucial for foresight in safety. In a second step (next part), the collective and organisational dimension of memory will be addressed.

4.5.1 Memory and forgetting

Memory work is directed against the forgetting. Forgetting is linked to memory, by being its negative side and it is its condition (Ricoeur, 2000). Todorov (1995) recalls that the integral restitution or reproduction of the past is impossible. Memory implies inevitably a choice (some traits of the events will be stored while others are immediately excluded or later eliminated and forgotten). However, though accessibility of the past is a requisite, past should not rule present, which questions the adequate re-use in the new context.

"To forget is sublime" was popularised by Tom Peters a management guru (in Forbes, 1994). Indeed, in intensive innovation capitalism (Le Masson et al, 2006), the value of experience over time decreases, with forms of quick obsolescence (e.g. in computer and electronic industries). "Unlearning theory" in management (e.g. Hedberg, 1981) studied the need to intentionally get rid of outdated experience, habits and routines in order to be able to innovate.

Knowledge is partly and directly acquired by learned-by-doing and repetition which materialise by a learning curve (Argote et al, 1990). Other kind of learning is indirect through learning from others, from failures and good practices.

Some risks of unintentionally forgetting, can imply a decay in the knowledge 'commodity or assets' and 'practices'. These risks were recognised as well (Easterby-Smith and Lyles, 2011) and may imply the loss of important technical knowledge and competence, the loss of identity and personal networks, the ability to learn from errors and to remain accountable from mistakes.

This can lead to knowledge losses or crash (Ermine, 2010) and costly crisis of relearning (Garcias, 2014). This phenomenon was encountered by nuclear engineering after years of downsizing in the aftermath of Three Mile Island and Chernobyl accidents.

4.5.2 The process of memorising: three key faculties

According to Sherwood (2015), memorising process is the faculty of the human mind by which information is encoded, stored, and retrieved:

- *Encoding* or registration: receiving, processing and combining of received information
- *Storage*: creation of a permanent record of the encoded information in short term or long term memory
- *Retrieval*, recall or recollection: calling back the stored information in response to some cue for use in a process or activity.

In our view, 'information' is a very broad term than encompasses data, information and knowledge learned (patterns, stories, models).

4.5.3 The process of loss of memory useful for foresight in safety

According to the above-mentioned definition that relies on a human perspective, loss of memory represents a failure in one of the three key faculties of the memorising process, for the individuals involved in the risk prevention process.

For the case of foresight in safety in a socio-technical system, according to the proposal made in the previous part, loss of memory means that information and knowledge cover four aspects mentioned in Table 1.

Therefore, for the four aspects of memory, the danger is that they are not encoded, stored, or retrieved properly by humans who can influence the design, structure and behaviour of the sociotechnical system.

This description of loss of memory is very general (more general than descriptions from ESReDA (2015)) since it encompasses not only information that was present and later was forgotten, but also the information and knowledge that should have been present but was not present.

Visibly a large variety of causes can contribute to the loss of memory. Typical causes of loss of memory represent insufficient knowledge, training or support of four aspects of memory. For instance, the refreshment of some training needed to maintain a skill is very dependent from the frequency of use in the daily practices, differentiating rescue skills from daily health and safety skills for operating (Lawani et al, 2018).

Frequently, such problems become highly visible when turnover of workers bring a loss of experience, transfers of obligations to new people, outsourcing, subcontracting, or ageing of personnel especially when the generation of 'baby-boomers' is retiring. Within the presented approach also all these problems point to the insufficient knowledge, training or support of four aspects of memory. For example, the subcontracting issue was raised for Toulouse disaster about the likely lack of awareness about the danger of chemical incompatibility between ammonium nitrate based wastes and chlorinated compounds. In addition, LoM happens between groups of workers and phases of operations, especially when modifications are made by operators forgetting the designers' intentions and limits of the system (Ferjencik and Dechy, 2016). LoM can occur as a consequence of changes and transitions (e.g. technological, regulatory, cultural and societal) (Mangeon et al, 2020), which are especially at work these years.

Such a way, not only situations when a specific knowledge or ability has been forgotten (i.e. it was not possible to be retrieved) are covered by these four aspects of loss of memory. In addition, situations are covered when this knowledge or ability has never been present (encoded and stored) in memory or has been present (and retrieved) but the ability and will to use it has not been observed.

4.5.4 Twelve categories of loss of memory

The three human faculties of memorising – encoding, storing and retrieving – allow to distinguish three kinds of failures leading to LoM. Failures of individual faculties can occur in each of four aspects of memory mentioned in Table 1 (awareness for vigilance, recognition for sense-making, investigation for detection, and follow-up). Thus, this leads to consider a two-dimensional scheme that is depicted in Table 2 that classifies LoM into twelve options, each representing a failure of one faculty in one aspect. E.g. LoM marked L2C3 represents failure of faculty #3 in aspect #2 or information is not retrieved on what forms of EWSs are possible. This division of LoM forms may help to sort various difficulties for maintaining the memory and possible safety management actions against the loss of memory.

Table 2. Twelve categories of loss of memory.

		Memorising faculties		
	Memory of information/knowledge	is not encoded	is not stored	is not retrieved
Aspects of memory related to foresight in safety	that Early Warning Signals can arise (awareness)	L1C1	L1C2	L1C3
	what forms of Early Warning Signals are possible (recognition and make sense)	L2C1	L2C2	L2C3
	that and how presence of Early Warning Signals can be detected (method)	L3C1	L3C2	L3C3
	about EWSs that were detected until they are reasonably responded (follow-up)	L4C1	L4C2	L4C3

4.5.5 Applied example: loss of memory in the kitchen

As long as Kate or William are alone in the kitchen, it may seem that all the LoMs that may occur are truly connected to their individual memories, of their individual experiences. Even with Kate and William both in the kitchen, we can only associate memory with the brains of these two people. Notice that they can develop a collective memory of their experiences in this kitchen or previous ones, with their parents or as students, and make it a topic of discussion. In the end, several hazards and practices to manage risks become implicit and do not need an explicit communication and analysis of the risks in the kitchen and on how to handle them. But this approach becomes unsustainable, especially when Kate and William want to involve their teenagers in the kitchen. The process of learning and memorising has to start all over again.

It cannot be assumed that from the very first moment they enter the kitchen, George, Charlotte or Louis will have the same information about EWSs as their parents. William and Kate will teach and train their children, on the relevant information to consider and the knowledge to learn that is important for managing risks and foresight in safety. And they may use some tools to do this - for example,

a list of EWSs detectable in preparing the pressure cooker for use. What if William and Kate fail to put an item on the list and Charlotte or Louis get an unwanted event? In this case, it was probably the loss of human memory classified in Table 2 as L3C1. But what if the critical item was originally listed but ceased to be readable, for example due to the effects of kitchen fumes? In this case, it could be an L3C2 LoM, but only on condition that we accept the list of EWSs as part of the memory.

This example shows that in organisations (even a couple and a family can be considered as a group or as sub-unit of an organisation) it makes sense to extend the term memory beyond the human mind to address the aiding mechanisms that support and/or replace human memory (especially considering turnover risk and retirement perspective). The existence of such aids is a normal situation. Memory support is one of the tasks of safety management systems. Management system and culture should support memory of personnel. Hardware and software of socio-technical system may contain tools that help memory of humans (e.g. proper databases and/or procedures for identification of EWSs). We will address some features of organisational memory in the next chapter and some functions of the safety management system to support it (in Part 4.7).

4.6 Loss of Memory and Organisational Memory

4.6.1 Extension from human to organisational memory

The previous daily-experience example of the kitchen can serve as an explanation for the need to extend the perimeter of memory of risks that is to be managed and applied to an organisation. Obviously, organisational memory includes memories of all humans and groups within the organisation, who contribute to make the information and knowledge on all four aspects of EWSs encoded, stored and retrieved as described in Table 2. It includes the collective memory of some experiences and tacit rules of groups of workers related to danger (or within the family for the kitchen example). But in addition to human minds, it is also necessary to consider in organisational memory all the safety management aids and functions implemented by an organisation in relation to the aspects and faculties of Table 2. Indeed, the memory of many kitchens risks faced over decades and centuries should be somehow treated by the designers of kitchen, and they should implement man-machine interface to secure some processes (e.g. confirmation

request when programing, alarm, fail-safe,…). To us, this is a rather radical extension.

Indeed, Kletz (1993) already warned that if organisations had no memory, accidents would recur. ESReDA (2015) underlined that the loss of knowledge is one of symptoms of the numerous barriers and failures to learn listed in that report: "There is a natural tendency that memory fades over time. People forget things. Organisations forget things. The lessons learned from incidents and accidents are slowly lost if no measures are taken to make them alive." This quote already associated LoM with the organisation. Not only people, but organisations lose information from classes L1C1 to L4C3 as shown in Table 2.

In addition, ESReDA (2015) stressed the importance of keeping memory: "Keeping memory of past events, lessons to be learned or not forgotten, and changes made is a key learning function". Expanding ESReDA (2015) and other proposals (Ferjencik and Dechy, 2016), organisations must therefore set-up a dedicated process to remember important safety issues and struggle actively against natural tendency to loss of memory that increases with time (erosion of memory, turnover of people, loss of designer intentions…). The management of LoM require some functions that we associate, with concerns of synergy and limited resources in real world context, with safety management as well as functions that support the retention of information from classes L1C1 to L4C3.

4.6.2   Nature of organisational memory

Organisational memory is not a simple concept. The human memory metaphor with its three faculties is often used, but it is criticised for the anthropomorphism risks it generates when addressing organisational memory. Organisational memory is defined (Walsh and Ungson, 1991) as "stored information from an organisation's history that can be brought to bear on present decisions". Another author insisted more on knowledge than information (Stein, 1995) "Organisational memory is the means by which knowledge from the past is brought to bear on present activities, thus resulting in higher or lower levels of organisational effectiveness". Often seen as internal, some researchers (De Cuffa et al, 2018) proposed to widen its scope as practitioners mobilise other means such as social networks outside the organisation's control and responsibility.

Knowledge is also a much debated concept that can be seen in three different ways according to scientific disciplines. Some scholars in knowledge management (combining cognitive and economic resources based approaches) consider knowledge as a commodity with a "stock" of knowledge and often focus on the codification of the information content and the flow of its dissemination. Organisations can be seen as collections of knowledge assets (Wenger, 1998). But this approach is criticised for the reification of knowledge and inadequate understanding of 'situated action" and impact of context. A few scholars in organisational learning approach it as a capability for organising (Carroll, 2004). While scholars in theories of action study practices and approach knowledge rather as a process embedded in practices of learning and doing, called knowing, that is acquired by doing. It is also acquired by sharing experience and appropriating knowledge in context of use with community members (Cook and Brown, 1999; Gherardi, 2006)). Knowledge is dynamic, is a process and an outcome. "Knowledge is a flux mix of framed experiences, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers. In organisations, it often becomes embedded not only in documents or repositories but also in organisation routines, processes, practices and norms" (Davenport and Prusak, 1998).

Polanyi (1975) insisted on the personal character of knowledge "all knowing is personal knowing", especially in relationship to the concept of tacit skill. Bell (1999) considers that "knowledge is the capacity to exercise judgment", especially "to make competent use of categories and distinctions constituting that domain of action" (Wenger, 1998). But the judgment and competent use are influenced by a collective and organisational dimension that provides shared understandings, heuristics, knowledge repertoires and rules of action, which then relates to organisational memory. These ideas enabled Tsoukas and Vladimirou (2001) to propose that: "knowledge is the individual capability to draw distinctions, within a domain of action, bases on an appreciation of context, or theory or both. Organisational knowledge is the capability members of an organisation have developed to draw distinctions in the process of carrying their work, in particular concrete contexts, by enacting sets of generalisations whose application depends on historically evolved collective understandings". Within similar focus on knowledge for action, Gherardi (2006) addresses safety (in the construction industry) as an object of knowledge, the result of a practical activity of knowing, and the context in which that activity is performed and institutionalised as 'organisational practice'.

Knowledge categories often distinguished (Nonaka and Takeuchi, 1995) are explicit, tacit but also declarative and procedural, and even judgmental (Girod-Séville, 1996). In the famous SECI model (Nonaka and Takeuchi, 1995), they require four different process for being extracted, transferred and used (e.g. socialisation, externalisation, internalisation and combination).

The organisational memory systems are "sets of knowledge retention devices, such as people and documents, that collect, store, and provide access to organisation's experience" (Olivera, 2000). Some means are tangible artefacts (intranets, databases, smartphones, procedures, products, people, social networks) and less tangible (culture). They are more or less dispersed and centralised. Three levels can be distinguished: individual, groups dispersed, centralised (Girod-Séville, 1996).

However, when conceived as a passive depository of data, knowledge and information is doomed to remain of little use (Bannon and Kuuti, 1996). The key issue is that the use relies on an interpretative process that takes into account the context of use. Codification removes context and the challenge to codify without knowing the context of use is not obvious (Koornneef and Hale, 2004). While use of organisational memory highlights processes of memorising and remembering that should compensate loss of context, but should also co-produce complementary knowledge adequate to the new context, implying that is not just an application of knowledge. In other words, there will always be some improvisation in context.

Their accessibility (retrieval in our labelling) is a key variable for their use. Indeed, some research show that people were considered as the most effective memory system, especially for easy access, specific filtering of content (Olivera, 2000) and for experiential knowledge (Arasaki et al. 2017). While dependence to technological means has increased, the reliance on "who knows what" is still a key (Jackson and Klobas, 2008). The speed in access to information and the nature of activity shape a distinction between administrative jobs which rely more on centralised technological systems of explicit organisational memory while operational jobs prefer access means with tacit content (De Cuffa et al, 2018). External social network means are more used than expected blurring the frontier between internal and external organisational memory means. While French nuclear power plants started to invest in the nineties (Girod-Séville, 1996) some resources in official memory (made of centralised memories of declared,

procedural and judgmental), the research showed that in daily activities, the organisational memory was rather coming from the unofficial underground individual and collective memory (declarative, procedural and judgmental). This parallel and redundant system remains a key but is under-invested by management. It relies on key experts (Girod-Séville, 1996), and pillars of experience (Llory, 1996) that have a lot of judgmental experience, especially about networks of people and incidents. It is vulnerable to turnover. The official memory is rather used for legitimising the management choices towards the control authorities. Both systems interacts.

According to Lévy (1990), the new technologies of information and communication are creating a third pole "computer and medias" that follows the writing and oral poles. From integration and incarnation by humans in oral mode, it shifted to written archives, losing its connection to individuals and context, being objectivised in technical provisions and artefacts. To counter the danger of forgetting, among the combination [individual-writing-computer], companies invest too much in technology. While doing so, they promote an official memory that is not so used in daily practice. In this sense, Tsoukas and Vladimirou (2001) warn that in knowledge management, digitalisation cannot be a substitute for socialisation.

An effective memory relies on interactions and a good combinations of the three modes (oral, written, computer and medias). It also relies on the connexions between people, units that build the meta-memory on the location of its memory resources. Expert-systems (popular in nineties) failed to capture the know-how of experts. A better reliance on key individuals, old and pillars of experience, or "filters" that compensate loss of context (Koornneef, 2005) remains an important strategy. Also, investing on formalisation of organisational history, on stories which are valuable and convey culture, beliefs, values and help decision-making (Girod-Séville, 1996; Hayes and Maslen, 2014; Duffield and Whitty, 2015). Communities of practice with focus groups are also a key to develop knowledge sharing about skills and know-how. The role of facilitators, mentors is a key (Duffield and Whitty, 2014).

4.6.3   Extension accords with nature of organisational memory

The foregoing overview of the nature of organisational knowledge and memory has accumulated a number of perspectives that would require further explanation and commentary. For example, differences in memory versus knowledge,

relationships between the human mind, written documents and computerized media, or the boundaries between external and internal.

All this scientific knowledge informs us how complicated is the field we enter when considering loss of memory in socio-technical systems. However, they do not suggest that the description of the organisational memory that we have proposed here in a safety practitioner perspective is fundamentally incorrect.

In summary, we consider that organisational memory includes:

- memories (that are encoded, stored and retrieved) of all humans within the organisation, who contribute to make and communicate the information and mobilise their knowledge on all four aspects of EWSs,
- the collective memory of experiences and tacit rules of groups of workers related to danger,
- the safety management and safety culture aids and functions implemented by an organisation in relation to the twelve categories of loss of memory,
- all artefacts that directly support human actions (documents, databases,…) or indirectly influence activities and interactions where organisational memory has been embedded into the equipment design, organisational structures and processes.

In this chapter, we assume that the decisive role in this description is played by the description of the safety management system, its tools and functions. To refine and use this description it is necessary to use a modern and up-to-date description of the safety management system (see Parts 4.7.5 and 4.7.6).

4.6.4   Paradox of organisational memory
An important starting point of this ESReDA book chapter, is to discuss the legacy of Trevor Kletz and its provocative statement: indeed, Kletz (1993) has warned that 'organisations have no memory, only people have'."

To our interpretation, Kletz did not want the notion of memory to be exclusively associated with the human brain, but at the same time realised the needs for organisations to work actively on their memory. In addition to the above strongly critical statement about organisations that have no memory Kletz (1993) also stated: "the leitmotiv of this book: the need for organisations to learn and remember the lessons of the past", which is coherent with our findings and

research (Ferjencik and Dechy, 2016; Dechy et al, 2016). So Kletz (1993) was warning us about the paradox that is also mentioned in ESReDA (2015): organisations do not strictly have human memory and at the same time need to remember.

4.6.5   Complexity and fundamental difficulty for maintaining the memory
The extended description of organisational memory clearly shows how complicated the organisation's memory is and exposes how fundamentally difficult it is to maintain the memory. It will likely remain a root cause of accidents if this issue is not address by a strong program and without specific management actions.

The fundamental difficulty for maintaining the memory is that the individual aspects of memory important for foresight in safety are distributed among different humans and groups on different levels of hierarchy in socio-technical system.

In addition to people and groups, other devices and tools (databases, documents…) may be involved and help to centralise and sustain part of the official memory. Machines can also play a supporting role in foresight. In a simpler case they help the humans to remember, identify, by providing aids and cognitive prosthesis like databases, etc. Basically, any aspect of memory (be aware - identify - detect – respond) can nowadays be accomplished by or with machines, and men-machine interfaces.

But we should consider that all system meta-components such as machines, equipment, procedures, processes and structures of organisations, are all embedding some of the memory of their designers, their operators, their managers and some of the organisational memory. Some memory is internal while other is external; some is formal while other is informal (e.g. kept within a group oral culture, such as a story on an incident), and their accessibility varies according to the work situations where it could be used. In summary, any organisational artefact integrates some organisational memory and by interacting with other meta-components of the sociotechnical system is more or less making alive the organisational memory.

The interactions of meta-components are permanently formed and influenced by formal and informal organisational interactions, rules and approaches (management system and culture). Hence all actors (operators closest to the valves (e.g. in Texas City), safety analysts such as auditors, managers) are

concerned by the four aspects of memory (awareness, making sense, investigating, follow-up) and may be exposed to difficulties in the three faculties of memorisation process.

The role of all is important towards treatment of EWS for foresight in safety. However, we want to insist on the point that the knowledge on the highest levels of hierarchy where local safety culture is promoted, is regulated, is scrutinized for the alignment it implies, is of extraordinary importance. As it is suggested above: humans on the highest levels of hierarchy need the knowledge about the EWS that can arise in order to motivate their vigilance and their commitment to identify EWSs, detect EWSs and respond to EWSs. Indeed, the notion of weak signals in safety has been inspired by one of its origin in strategic management (Ansoff, 1975; Ansoff and McDonnell, 1990) and military management ('intelligence' in Turner, 1978), especially when uncertainty is high and to avoid strategic surprises. Loss of this knowledge and commitment would represent a fundamental loss of memory in the system.

4.7   Activities against the Loss of Memory

4.7.1   Use of extended description of memory
Though we previously understood that taking care of human, group and organisational memory useful for foresight is quite complicated, we assume that in many ways it merges with general care of good safety culture and safety management.

However, to avoid the danger of dilution and lack of will if one considers that LoM is already addressed by existing provisions in safety management, the foregoing text also offers one aid that may be useful in identifying more specific measures that may decrease LoM. Using the scheme in Table 2, local organisational memory can be divided into twelve categories. The scheme can be applied to any activity within the organisation.

For each Table 2 category separately, the following questions can be asked:

- What is the memory of important hazards according to the different actors?

- Who are the pillars of experience holding critical knowledge and memory about risks and EWS?
- Are there some processes to formalise some of this knowledge and to train newcomers on an informal basis within day-to-day practice?
- Is the continuity and stability of groups and communities of practice considered?
- Are the lessons from other plants, units, competitors, countries, sectors learned?
- What are the threats in the near and longer term to human, group and organisational memory owned by the organisation?
- Are some persons from the local hierarchy involved in actions to maintain organisational memory?
- Which aids and devices contribute to make available this memory?
- How much organisational and human memory is embedded in the designs, structures, and processes and how is it known to daily users?
- Which functions of the local management system do support it?
- Are the operators and managers trained to methods to detect EWS?
- Is the follow-up of critical risk reduction actions adequate?
- …

Of course, the list of questions remains open and can be further developed. Such an analysis could help detect, for example, under-covered parts of local organisational memory.

As an illustration, example of cooking in the pressure cooker can be used (see Parts 4.4.4 and 4.5.5). If we focus on detection categories L3C1 and L3C2, then the above question can draw our attention to the fact that there is a threat that newcomers will not realize that detection of safety valve throughput is necessary before cooking. An aid reminding this act as inherently as possible should be incorporated.

4.7.2   Against the loss of first aspect of memory
Examples of approaches that can help against the loss of individual aspects of memory are given in this and next sections.

Kletz (1993) already suggested some ways of improving the organisational memory. Starting from the belief that organisations had no memory, he tried to promote the memory of insiders especially with regard to safety in design and production. Writing about memory, he concentrated on the need to remember

the lessons learned from the past undesirable events. He already recommended to processing the lessons learned into the form of short messages (stories) which are spread, discussed and taught. These lessons should be referenced anytime they caused a change of a code, standard or operating instructions. Old messages should be made permanently accessible.

The CAIB (2003) compared NASA practices with US submarines. Indeed, the US Navy submarine had two major submarine loss in 1963 (the Thresher) and 1969 (the Scorpion) which resulted in renewed efforts to prevent accidents. Some of them are redirected towards loss of memory (pp183-184). CAIB noticed: "The submarine Navy has a strong safety culture that emphasizes understanding and learning from past failures. NASA emphasizes safety as well, but training programs are not robust and methods of learning from past failures are informal" and "The Navy implements extensive safety training based on the Thresher and Scorpion accidents. NASA has not focused on any of its past accidents as a means of mentoring new engineers or those destined for management positions".

All these ideas seem to be perfectly right if the first aspect of loss of memory - missing knowledge, awareness and therefore vigilance about EWS that can arise in the system - is to be prevented. Such messages can vividly and convincingly remind the existence of EWSs, support this knowledge and motivate the will and commitment to identify EWSs and respond to EWSs.

Nevertheless, such activities cannot be considered as sufficient with regards to prevention of other aspects of loss of memory. Activities recommended by Kletz (1993) provide examples of EWSs arising in the system but do not warrant that any possible EWS will be identified and responded during the learning from undesirable events, both internal, and external. Completeness of identification is not warranted.

4.7.3 Against the loss of second aspect of memory
Woods (2009) states that the safety field still lacks the ability to identify EWSs. He also recommends the alternative way to desired indicators – adaptive stance –, but this is not the main issue here, except if some EWS invite to set-up an adequate and adaptive response to prevent further safety degradation. Contrary to this opinion, CCPS (2011) means that warning sign surveys are feasible and offers a tool to identify EWSs – a list of catastrophic incident warning signs that flows from the description of safety management system in book (CCPS, 2007) and which is

supplemented by a list of physical warning signs. Based on this list classes of EWSs that can be distinguished and most probable classes of EWSs may be identified for a specific socio-technical system. This list seems to represent a tool that may help to identify EWS in a socio-technical system when it arises. Such a way it gives a strategy on how to identify EWSs during learning from experience and at the same time the ability to prevent second aspect of loss of memory.

4.7.4 Learning acts against the loss of memory
Since LoM related to foresight in safety is a matter of forgetting the safety lessons learned by people and organisations, suitable recommendations can be found in texts about learning (e.g. ESReDA, 2015). For example, if we share the empirical statement that learning is quite often not satisfactorily efficient (ESReDA, 2015 ; Dechy et al, 2011) and is among the major recurring root cause (Hopkins, 2010, Dechy et al, 2018), For the sake of completeness, it should be noted at this point that Marsden (2017) suggests other possible root causes of LoM in addition to ineffective learning. He summarizes them as short-termism, loss aversion, regret aversion, ambiguity-driven indecisiveness, and dilution of responsibility.

The paper by Lawani et al. (2018), on the other hand, makes recommendations on how to make people's training more successful with a better memorisation if the frequency of use of the knowledge and skill is low (e.g. contrasting daily safety practices and emergency and rescue practices only performed during refreshing exercises).

Stemn et al. (2018), in a retrospective analysis of a fairly large set of articles, tried to describe how and why failure to learn from safety incidents do occur. To understand where breakdowns in learning from incidents are occurring, a bowtie analysis was used to organise the literature on failure to learn from safety incidents in a way that informs researchers and practitioners of priority areas. Stemn et al. confirmed in their retrospective analysis the importance of safety management functions for learning and memory.

In our analysis of three accidents in a very particular type of industry in Europe (Ferjencik and Dechy, 2016), we have seen the importance to learn the "hard lessons" "from others", especially from accidents that occurred abroad. This is not a new idea but still there remains huge margins for progress. This remark is to be extended to generalizable lessons coming from accidents in other industrial sectors (e.g. Can we learn and transfer the lessons learned from Columbia space

shuttle loss to others sectors in Dien and Llory, 2004). Again, the knowledge basis to consider for memorisation should integrate the history and knowledge of recurring patterns of accidents (Dien et al, 2004; Dechy et al, 2016). We are convinced that this knowledge and the stories of accidents are useful for the four aspects of memory described in Table 1. The CAIB (2003) noticed "Recurring Training and Learning From Mistakes: […] For example, since 1996, Naval Reactors has educated more than 5,000 Naval Nuclear Propulsion Program personnel on the lessons learned from the Challenger accident."

In addition, lessons from good, reliable and safe practices or new requests from regulations appear from time to time as well and challenge the status quo of the knowledge base. If they are not seized as opportunities for improving safety management and preventing next accidents, they can be considered as a flawed process of management of actions and to some extent of memorisation.

Last but not least, the importance to keep memory of designers' intentions (Ferjencik and Dechy, 2016), while taking care of the danger of loss of information, knowledge and memory at every step of the life-cycle of the sociotechnical system (Stoop, 1996; Leveson, 2004) remains a key issue. In this perspective, CAIB (2003) observed that "Both Naval Reactors and the SUBSAFE have "institutionalized" their "lessons learned" approaches to ensure that knowledge gained from both good and bad experience is maintained in corporate memory. This has been accomplished by designating a central technical authority responsible for establishing and maintaining functional technical requirements as well as providing an organisational and institutional focus for capturing, documenting, and using operational lessons to improve future designs. NASA has an impressive history of scientific discovery, but can learn much from the application of lessons learned, especially those that relate to future vehicle design and training for contingencies. NASA has a broad Lessons Learned Information System that is strictly voluntary for program/project managers and management teams. Ideally, the Lessons Learned Information System should support overall program management and engineering functions and provide a historical experience base to aid conceptual developments and preliminary design".

The importance to formalise stories of incidents and accidents in order to facilitate the lessons sharing and remembering has especially been invited by Kletz (1993). Incidents databases have been established and the development of computers (and in the future with big data) has provided many new possibilities for better storing and retrieving data lessons (ESReDA project group in the nineties focused on the databases issue, produced several deliverables and organised a few ESReDA seminars). More recent work has highlighted the role played by stories in the daily decision-making processes (Hayes, 2013) but also in the mentoring of younger colleagues (Hayes and Maslen, 2014; Maslen, 2014). Beyond the formal training, and the databases, stories are particularly efficient from the memorisation standpoint and for transferring skills and values. As we recalled, US Navy implemented this lever and NASA did not (CAIB, 2003).

A last aspect deals with human resources management, to retain knowledge from critical competence (Ermine, 2010) by codifying, but also socialised 'pillars of experience' (Llory, 1996) with novices, or build through career paths management and other provisions. CAIB (2003) highlighted some provisions to retaining knowledge within US submarines: "Naval Reactors uses many mechanisms to ensure knowledge is retained. The Director serves a minimum eight-year term, and the program documents the history of the rationale for every technical requirement. Key personnel in Headquarters routinely rotate into field positions to remain familiar with every aspect of operations, training, maintenance, development and the workforce. Current and past issues are discussed in open forum with the Director and immediate staff at "all-hands" informational meetings under an in-house professional development program."

Among all these knowledge sources, some may be enough as they are to raise awareness, to be used as cognitive resource to recognise EWS, to investigate and to maintain rigour on the follow-up. However, with the goal of enhancing foresight in safety and proactivity in safety management, we believe that some dedicated efforts to extracts from this knowledge, the frameworks useful for awareness, interpretation of EWS are still necessary, as well as socialisation processes with pillars of experience and management of communities of practices.

In addition, not all EWS will be predefined. Some detection, investigation, sense-making processes will require imagination, further investigations to collect additional data and discussions with peers, analysts and managers to converge towards new type of dangers and EWS. To some extent, discussion with whistleblowers will provide an opportunity to revise beliefs, existing dangers and EWS lists. (see. Chapter 12, Dien et al, 2020).

4.7.5 Role of safety management system functions

Two ideas emerged in the previous parts:

1) Loss of memory can be divided into four aspects (awareness for vigilance, recognition for sense-making, investigation for detection, and follow-up) and three faculties (encoding, storing, and retrieving). Combination of both results in division of loss of memory into twelve categories.

2) Speaking about memory, not only human memory but also a group and organisational memory have to be considered; in operational perspective, they combine a careful resource management of human brains and bodies (skills) with aids from safety management system including safety culture.

The above ideas will be further developed in this and following parts. The close relationship between categories of loss of memory and functions of safety management system will be made visible. To be operational and synergistic with resources constraints, we will show that the question on what decreases certain category of loss of memory can be converted into the question of what improves the performance of specific functions in local safety management. Beware that a complementary strategy is however needed to address tacit skills and personal knowledge and knowing (Polanyi, 1975). It would require long term human resource management that includes socialisation processes (Nonaka and Takeuchi, 1995), invest in key experts (Girod-Séville, 1996) that can share judgmental experience within communities of practices (Wenger, 1998).

Safety management has many definitions as well. Harms-Ringdhal (2004) formulates probably the simplest one: 'Safety management is a way of managing the hazards (safety risks) of a company.' Number of more detailed descriptions of safety management can be found in literature. They differ from a brief information in Meyer and Reniers (2013), formal description in BSI (2007) to very detailed explanation in CCPS (2007). Their recommendations can be considered valid despite the fact that completeness of even the most detailed description is not warranted. For example, Broadribb (2018) shows that CCPS (2007) may lack recommendations for "Sharing and Learning Lessons".

For purposes of this chapter, we need a systematic and detailed description of safety management system that is generally respected. Suitable safety management system has to cover much broader range of functions than mere risk management (see Meyer and Reniers, 2013). It has to include even functions considered usually to be a part of safety culture.

Risk Based Process Safety (RBPS) model of safety management system by CCPS (2007) is used in this chapter. Article by Frank (2007) can be recommended for the introduction to this model. RBPS model consists of twenty elements. The elements are grouped into four main themes that is to say four safety management pillars: Commit to Process Safety, Understand Hazards and Risk, Manage Risk, and Learn from Experience. Relations between pillars can be illustrated by Figure 1.

For each of twenty elements, the RBPS model identifies key principles associated with the implementation of the element. Further, essential features required to support each key principle are identified. Such a way, a four-level hierarchical description of functions of safety management system is created that identifies several hundreds of safety management system functions.
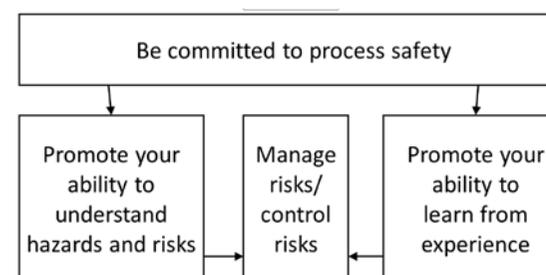
Figure 1. Relations among four safety management pillars from the Risk Based Process Safety model (CCPS, 2007).

Failures of safety management system functions are often considered to be among the underlying causes of undesirable events. Since EWS can be determined as a cause (including root cause) of causal event, links between safety management system and loss of memory start to be visible: early warning signs may be interpreted as failures of safety management system functions. This idea complements the two from the start of this part such a way that decrease of LoM can be converted to the improvement of the performance of specific functions in local safety management.

4.7.6 Safety management against the loss of memory

RBPS model of safety management system systematically describes functions and tools that may be applied as a protection against incident scenarios. Scenarios are always connected to hazards.

In previous Part 4.4.6 we explained how EWSs are associated with the existence of hazards, controls of hazards and environmental conditions that affect hazards. EWS represent causes or indications of failures in these hazards, controls of hazards and environmental conditions that affect hazards.

Altogether, we have strong arguments that the improvement of specific safety management system functions decreases specific categories of loss of memory. The practical problem that remains is to identify those particular relationships between functions and categories. These relationships provide an opportunity to convert a problem from the area of loss of memory to the area of safety management. In addition, if both the loss of memory category identification and the safety management system function identification are satisfactorily detailed, these relationships will provide useful support against the loss of memory.

The RBPS model is so detailed that for a particular activity it enables to precisely target safety management system measures against a particular category of loss of memory.

We will illustrate this with an example of a specific activity from our kitchen example, which was already used in sections 4.4.4 and 4.5.5. As shown in Table 3, we focus on detection categories L3C1 to L3C3 during preparation of pressure cooker for use.

In our example, we consider three possible findings of loss of memory, which we denote a) to c). We assume that we have found that information about what can detect a failure of the safety valve that protects against a specific hazard (overpressure during cooking in a pressure cooker) was not a) encoded, b) stored, c) retrieved. Each individual function of safety management system is described as a chain composed of a pillar, an element, a key principle, and an essential feature from the RBPS model.

Table 3. Conversions of categories of LoM to functions of safety management.

Failed activity	Kitchen example: Preparation of pressure cooker for use		
Category of Loss of Memory	Memory of information/knowledge that and how presence of Early Warning Signs can be detected is not		
	a) encoded: L3C1	b) stored: L3C2	c) retrieved: L3C3
Functions of safety management system described in Risk Based Process Safety model	Three alternatives: (i) in case that parents ignored to prepare the information: Commit to Process Safety - Process Safety Competency - Maintain a dependable practice - Develop a learning plan; (ii) in case that parents did not encode due to poor understanding: Understand Hazards and Risks - Process Knowledge Management - Protect and update process knowledge - Ensure accuracy; (iii) in case that parents forgot an item: Manage Risk - Asset Integrity and Reliability - Address equipment failures and deficiencies - Promptly address conditions that can lead to failure	Understand Hazards and Risks - Process Knowledge Management - Catalog process knowledge in a manner that facilitates retrieval - Protect knowledge from inadvertent loss	Two alternatives: (i) in case that children were not able to retrieve: Understand Hazards and Risks - Process Knowledge Management - Catalog process knowledge in a manner that facilitates retrieval - Document information in a user-friendly manner; (ii) in case that children did not try to retrieve: Manage Risk - Asset Integrity and Reliability - Develop and maintain knowledge, skills, procedures, and tools - Train employees and contractors.

4.8   Conclusions

Foresight is not only about imagination and exploration, but it also relies on reproduction and exploitation of existing knowledge; prospective views remains linked to retrospective views.

In an era that shows an acceleration of innovations and changes (technological such digitalisation trend, industry 4.0, but also societal, such as with short and potential long term impacts of COVID-19 pandemics), the value of experience in the time dimension is threatened and may become obsolete. However, history and memory still matters for expert knowledge at the age of big data and artificial intelligence, especially in systems governed by incremental changes.

If we take the current disaster of Beyruth, some authorities lost awareness and follow-up capabilities after alerts received. If we look at the major scale crisis of COVID-19, we forgot some lessons even from the 1918-1919 Spanish-flu, the worst pandemic on 20th century which killed 50 to 100 millions of people, 2 to 5% of population, much more than the world war one (Vinet, 2018). And we repeated patterns of minimising threats, late reactions to EWS, normalised health system vulnerabilities.

At the end of this journey, we are more than ever convinced that conducting generic and specific actions integrated in management of safety against loss of memory is a strategic investment to prevent accidents. One should remain humble in front of the complexity of "managing" human and organisational memories, and in front of the huge and everlasting tasks that require a dedicated dynamic and recurring program. The difficulty is not so much in the imagination, but it is rather in the complexity of the distributed meta-components of the system that should be used to lever organisational memory. In addition, it requires a dedicated commitment and constant attention. We believe that top management can provide organisational resources to pay attention to key safety issues, such as EWS, but also they can lever the practices if they are committed as exemplary leadership.

The accidents we referred in the dynamite and ammonium nitrate industries, the case of NASA and US Navy, and the pandemics are tragic stories but perfect reminders of how these issues matters. They also provide an alert to remain proactive towards the external lessons to learn and extend the perimeter of 'organisational memory' to other industries and countries. The everyday practice in your kitchen can be also the reminder of hidden and embedded aspects of individual, group and organisational memory and trigger your thoughts on how to make the EWS alive in daily practices with family and friends and on how to transfer those practices at work.

We are conscious that a strategy that would use the synergy with safety management actions for preventing loss of memory is necessary but not enough. There is also a general tendency to emphasize the knowledge management tools especially at the time of digitalisation, big data and artificial intelligence. One should accept that not all context can be captured in formal tools of knowledge management when codifying and storing information. In addition, the re-use will always imply a critical distance and adaptation to the context of use, within local practices that require a form of creativity. Indeed, "all knowing is personal knowing" (Polanyi, 1975) which invites to consider individual memories related to tacit skills and abilities to make judgment, which are coupled to socialisation mechanisms. This is done by the people who filter, who are the pillars of experience, within groups or communities of practice. So, on the one hand, a greater care and respect to worker expertise should be acknowledged, especially in face of particular turnover and skills losses with retirement waves. And on the other hand, organisational learning and making alive memory of others 'hard lessons' from failures is not trivial as well.

4.9   Acknowledgments

4.10   References

Abernathy, William J. (1978). The Productivity Dilemma: Roadblock to Innovation in the Automobile Industry (Vol. 35): Johns Hopkins University Press Baltimore, MD.

Ansoff, I. (1975) Managing strategic surprise by response to weak signals, California Management Review, Vol.18 (2), 21-33.

Ansoff, I. and Mc Donnell, E., (1990), Implanting strategic management, Second edition, Prentice Hall International, United Kingdom

Argote, Linda, & Epple, Dennis. (1990). Learning curves in manufacturing. Science, 247(4945), 920-924.

Arasaki, P. H. K., Steil A. V., Santos, N. (2017) Sistemas de memória em organizações intensivas em conhecimento: um estudo de caso, Espacios 38(4), 1.

Bannon, L. J. and Kuutti, K. (1996), Shifting perspectives on organizational memory: from storage to active remembering, In proceedings of the 29th IEEE HICSS Vol III, Information Systems – Collaboration Systems and Technology, pp. 156-167, IEEE Computer Society Press, Washington.

Bell D. (1999), The axial age of technology foreword: 1999, in The coming of post-industrial society. New York, Basic Books, Special anniversary edition, ix-lxxxv.

British Standards Institution BSI (2007), BS OHSAS 18001:2007 Occupational Health and safety management systems - Requirements.

Broadribb, M. P. (2018). And Now for Something Completely Different. *Process Safety Progress* 37(1) 25-30.

Columbia Accident Investigation Board (2003), Report Volume 1, National Aeronautics and Space Administration, the Government Printing Office.

CCPS (2007). *Guidelines for Risk Based Process Safety.* John Wiley & Sons, Hoboken, New Jersey, USA. ISBN: 978-0-470-16569-0, 698 pages.

CCPS (2012). AIChE Center for Chemical Process Safety. *Recognizing Catastrophic Incident Warning Signs in the Process Industries.* John Wiley & Sons, Hoboken, New Jersey, USA. ISBN: 978-0-470-76774-0, 227 pages.

CSB. (2007) Investigation Report, Refinery Explosion and Fire, BP – Texas City, Texas, March 23, 2005, US CSB Report N°2005-04-I-TX. 2007.

CSB (2016). U.S. Chemical Safety and Hazard Investigation Board. West Fertilizer Company Fire and Explosion, Final Investigation Report, REPORT 2013-02-I-TX, January 2016, 267 pp. Accessible from http://www.csb.gov. Accessed 5. 3. 2018.

Davenport, T. H., Prusak, L. (1998). Working knowledge, Cambridge, MA: Harvard University Press.

De Cuffa D., Kraemer R., Steil A. V. (2018), Use of organizational memory systems in a police organization, International journal of Knowledge Management, volume 14, issue 3, July-September.

Dechy N., Bourdeaux T., Ayrault N., Kordek M.-A., Le Coze J.-C. (2004) First lessons of the Toulouse ammonium nitrate disaster, 21st September 2001, AZF Plant, France, Journal of Hazardous Materials 111 pp 131-138

Dechy N., Rousseau J.-M., & Jeffroy F. (2011). Learning lessons from accidents with a human and organizational factors perspective: deficiencies and failures of operating experience feedback systems. In: Proceedings of the EUROSAFE 2011 conference, Paris. Available from http://www.eurosafe-forum.org/.

Dechy N., Rousseau J.-M., Dien Y., Llory M. Montmayeul R., (2016) Learning lessons from TMI to Fukushima and other industrial accidents : keys for assessing safety management practices, Proceedings of the IAEA International Conference on Human and Organizational Aspects of Assuring Nuclear Safety –Exploring 30 Years of Safety Culture, 22-26th February, Vienna, Austria

Dechy N., Dien Y, Marsden E., Rousseau J.-M. (2018), Learning failures as the ultimate root causes of accidents, in Hagen J. U. How could this happen? Managing errors in organizations, Palgrave Macmillan

Dien, Y. & Llory, M. (2004). Effects of the Columbia Space Shuttle Accident on High Risk Industries or: Can We Learn Lessons from Other Industries? Conference Hazards XVIII, November 23-25, Manchester, UK.

Dien, Y., Llory, M. & Montmayeul, R. (2004). Organizational accidents investigation: methodology and lessons learned, Journal of Hazardous Materials, 111 (1-3), pp 147-153.

Duffield, S. and Whitty S. J. (2015), Developing a systemic lessons learned knowledge model for organisational learning through projects, International Journal of Project Management, 33, pp 311-324.

Easterby-Smith, M., & Lyles, M. A. (2011). In praise of organizational forgetting. Journal of Management Inquiry, 20(3), 311-316.

Ermine, J.-L., (2010). Knowledge crash and knowledge management. International Journal of Knowledge and Systems Science, 1(4), 79-95.

ESReDA, Ed. (2015), Barriers to learning from incidents and accidents, Marsden E., Dechy N., Dien Y., Druspsteen L., Felicio A., Cunha C., Roed-Larsen S., Tulonen T., Stoop J., Strucic M., Vetere Arellano A.-L., Van der Vorm J. www.esreda.org

ESReDA (2015) Project Group on Dynamic Learning. Guidelines for Preparing a Training Toolkit in Event Investigation and Dynamic Learning. Available from http://www.esreda.org.

Ferjencik, M. and Dechy, N. (2016). Three accidents in European dynamite production plants: An attempt to improve the external lessons learning. *Journal of Loss Prevention in the Process Industries* 44, 12-23.

Frank, W. L. (2007). Process Safety Culture in the CCPS Risk Based Process Safety Model. Process Safety Progress 26 (3) 203-208.

Garcias F. (2014), Apprentissage, désapprentissage et réapprentissage organisationnels – Le cas d'une activité d'ingéniérie de grands projets complexes, Thèse de doctorat ParisTech, Sciences de gestion.

Gherardi S. (2006), Organizational knowledge: the texture of workplace learning. Blackwell publishing.

Girod-Séville M. (1996), La mémoire des organizations, Editions L'Harmattan, ISBN 2-7384-4885-2.

Guillaume, E. (2011), Identifying and Responding to Weak Signals to Improve Learning from Experiences in High-Risk Industry, PhD thesis, Delft University, The Netherlands.

Gyenes, Z. and Dechy, N. (2016), Risk and safety management of ammonium nitrate fertilizers: keeping the memory of disasters alive, Loss Prevention Bulletin n°251, October 2016, pp32-36, Edited by Icheme.

Harms-Ringdahl, L. (2004), Relationships between accident investigations, risk analysis, and safety management, Journal of Hazardous Materials 111 13–19.

Hayes, J. (2013). Operational Decision-making in High-hazard Organizations: Drawing a Line in the Sand. Farnham: Ashgate.

Hayes, J., Maslen, S. (2014). Knowing stories that matter: learning for effective safety decision-making, Journal of Risk Research.

Hedberg, B., (1981), How organizations learn and unlearn, in Handbook of organizational design, Nystrom P. C, Starbuck W. H., Oxford University Press.

Hopkins, A. (2010). Failure to learn: the BP Texas City refinery disaster, CCH Australia Ltd. Wolters Kluvers.

Hutchins E. (1994), Comment le "cockpit" se souvient de ses vitesses, Sociologie du travail, 36ème année, n°4, Octobre Décembre 1994, Travail et cognition, pp451-473.

Jackson, P., Klobas, J. (2008), Transactive memory systems in organizations: implications for knowledge directories. Decision Support Systems, 44.

Jouniaux P., Hadida D., Dechy N., Marle L., Billy F., Pierlot S., Parrennes F., Rouvière G., Husson D., (2014) "Détection, pertinence et amplification des signaux faibles dans le traitement du retour d'expérience", Congrès λμ19 de l'IMdR, Dijon, 21-23 octobre 2014.

Kingston J., Dien Y. (2017). The McNamara Fallacy Blocks Foresight for Safety, Proceedings of the 53rd ESReDA Seminar "Enhancing Safety: the Challenge of Foresight" Edited A. L. Vetere Arellano, Z. Šimić, N. Dechy, 14-15 November 2017, Ispra, Italy

Kletz, T. (1993). *Lessons from Disasters: How Organizations Have No Memory and Accidents Recur.* Gulf Publishing Company, Houston. ISBN: 0-88415-154-9, 183 pages.

Koornneef, F., and Hale, A. (2004), Organizational memory for learning from operational surprises: requirements and pitfalls, in Andriessen J. H., Fahlbruch B., How to manage experience sharing – From organisational surprises to organisational knowledge, Elsevier science, Amsterdam

Le Masson, P., Weil, B., & Hatchuel, A. (2006). Les processus d'innovation : conception innovante et croissance des entreprises: Hermes science publ.

Lesca, H., (2001), Passage de la notion de signal faible à la notion de signe d'alerte précoce, Proceedings of the VSST conference, Barcelona, Leveson, N., (2004), A new model for engineering safer systems, journal of Safety Science, vol 42, pp237-270.

Lévy, P. (1990), Les technologies de l'intelligence: l'avenir de la pensée à l'ère informatique., Points science

Merritt, C., W., (2007). Testimony of Carolyn W. Merritt, U.S. CSB, before the U.S. House of Representatives, Committee on Energy and Commerce, Subcommittee on Investigations and Oversight, May 16, 2007.

Meyer, T. and Reniers, G. (2013). Engineering Risk Management. Walter de Gruyter, Berlin/Boston. ISBN: 978-3-11-028515-4, 284 pages.

Lawani, K., Hare, B., Cameron, I. (2018). Integrating early refresher practice in height safety and rescue training. *Safety Science* 110 411-417, https://doi.org/10.1016/j.ssci.2018.03.029

Llory, M. (1996). Accidents industriels : le coût du silence, Opérateurs privés de parole et cadres introuvables, Éditions L'Harmattan.

Mangeon, M., Dechy N., Rousseau J.-M., (2020), The 1969 and 1980 nuclear accidents in Saint-Laurent-des-Eaux: When transition generates forgetting (in french). Proceedings of the Conference λμ22, IMdR.

March, James G. (1991). Exploration and exploitation in organizational learning. Organization Science, 2, 71-87.

Marsden, E. (2017). Justifying safety interventions based on uncertain foresight: empirical evidence. In *53rd ESReDA Seminar*, Ispra, Italy, November 2017.

Maslen, S. (2014). "Learning to Prevent Disaster: An Investigation into Methods for Building Safety Knowledge among New Engineers to the Australian Gas Pipeline Industry." Safety Science 64: 82–89.

Olivera, F. (2000), Memory systems in organizations: an empirical investigation of mechanism for knowledge collection, storage and access. Journal of Management Studies, 37 (6), pp811-832

Nonaka, I.,Takeuchi, (1995), The Knowledge-Creating Company: How Japanese Create the dynamics of Innovation. New York, NY: Oxford University Press.

Polanyi, M. (1975), Personal knowledge. In Polanyi L. and Prosch H. (Eds), Meaning, Chicago, IL, University of Chicago press, pp22-45

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

Reason, J. (1990). Human Error, Cambridge University Press.

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

Ricoeur P. (2000), La mémoire, l'histoire, l'oubli., Editions Points, Paris.

Rouibah K, Ould-ali S. (2002), PUZZLE: a concept and prototype for linking business intelligence to business strategy, Journal of strategic information systems, 11, pp133-152

Rosen, R. (2015).The Ten Commandments of Risk Based Process Safety. *Process Safety Progress* 34(3) 212-213.

Sherwood, L. (2015). *Human Physiology: From Cells to Systems*. Cengage Learning. pp. 157–162. ISBN 978-1-305-44551-2.

Stark, S., (1961), Executive Foresight: Definitions, Illustrations, Importance, The Journal of Business, Vol. 34, No. 1, (Jan., 1961), pp. 31-44, Published by: The University of Chicago Press

Stein E. (1995), Organizational memory: review of concepts and recommendations for management. International journal of information management, 15, 17-32.

Stemn, E., Bofinger, C., Clift, D., Hassall, M. E. (2018). Failure to learn from safety incidents: Status, challenges and opportunities. *Safety Science* 101 (2018) 313–325.

Stoop, J. (1996), Design Control Practice scheme. See in ESReDA Guidelines for safety investigation of accidents (2009), www.esreda.org

Todorov I. (1995), Les abus de la mémoire, Editions Arléa.

Tsoukas H. and Vladimirou E. (2001), What is organizational knowledge? Journal of management studies, 38:7, November 2001, 0022-2380, pp 973-993.

Turner, B. (1976), The Organizational and Interorganizational Development of Disasters, Administrative Science Quarterly, Vol. 21, No. 3 (Sep., 1976), pp. 378-397

Turner, B. & Pidgeon, N., (1997) Man-Made Disasters, Second edition, Butterworth Heinemann [1rst edition: Turner, B. (1978), Wykeham Publications]

Vaughan, D.. The Challenger Launch Decision. Risky Technology, Culture, and Deviance at NASA, The University of Chicago Press. 1996.

Vinet, F. (2018), La grande grippe. 1918. La pire épidémie du siècle. Editions Vendémiaire.

Wenger E (1998), Communities of practice. Cambridge University Press.

Wikipedia (2018) Available from https://en.wikipedia.org/wiki/Memory. Accessed 25. 4. 2018.

Woods, D. D. (2009). Escaping Failures of Foresight. *Safety Science* 47, 498–501.

5 Use of Scenarios as a Support of Foresight in Safety

Milos Ferjencik, University of Pardubice, Czech Republic
Miodrag Stručić, EC Joint Research Center, the Netherlands,
Tuuli Tulonen, Tukes, Finland
Eric Marsden, FonCSI, France
John Stoop, Kindunos Safety Consultancy Ltd., the Netherlands

5.1 Executive summary

Incident scenarios are a practical tool for thinking about risk. Scenarios may be results of prospection or retrospection. Both prospective and retrospective scenarios can be used for lessons learning.

Any incident scenario can be reduced to a set of causal events. Lessons learning can reach Early Warning Signs (EWS) through the identification of causal events. EWSs are causes and indicators of causal events.

This chapter shows that results of lessons learning via scenarios can be used:

- to prevent loss of memory,
- to list all possible EWSs,
- to identify whether a failure/error/condition represents an EWS,
- to prioritize EWSs.

All preceding claims are illustrated by examples.

5.2 Key Messages

Foresight requires determining the events/conditions that are to be considered EWSs. The incident scenarios may play useful roles since EWSs can be determined from scenarios obtained by both prospective and retrospective analysis. The path to determine EWSs leads via the determination of causal events.

With the use of incident scenarios, both identifying and prioritizing the EWSs is possible. They help make visible the EWSs, and select EWSs that deserve special attention (e.g. real-time monitoring).

Scenarios as an investigation component of lessons learning help to determine sets of EWSs that should be searched for and tracked during the analyses.

Scenarios as a documentation component of lessons learning help the determination of whether a specific failure/error represents an EWS.

5.3 Introduction

This chapter is based on ideas presented in the paper by Ferjencik (2017). The text uses terminology that is standard in publications issued by the American Institute of Chemical Engineers. Readers interested in a reminder of the meanings and relations of terms such as hazard, control, initiating event, scenario etc. may find instructive illustrations in articles by R. F. Blanco, e.g. in Blanco (2014). A new term, i.e. 'causal event', is introduced in this chapter.

When scenarios are discussed, both accident or incident scenarios are implied throughout this chapter. The word 'scenario' has the same meaning as the term 'accident sequence' in Benner's paper (1975), i.e. a possibly multilinear sequence of events representing individual actions of animate or inanimate actors that leads to an injury or damage. If one of the actors fails or is unable to adapt, the perturbation starts the accident sequence. Thus, the scenario begins with a perturbation (initiating event) and ends with the last injurious or damaging event in the sequence.

Scenarios represent a tool for lessons learning. Both prospective and retrospective scenarios can be used for lessons learning. Lessons learning in this chapter focuses on early warning signs (EWS). EWSs are part of lessons learned resulting from the lessons learning process. The main concept introduced in this chapter states that the EWSs can be identified with the use of scenarios via the identification of 'causal events'.

Additionally, scenarios can be used as an investigation and documentation tool. Use of scenarios as an investigation component of lessons learning helps to identify sets of EWSs that should be searched and tracked during the analyses. As a documentation component of lessons learning, it helps to determine whether a specific failure/error represents an EWS.

This chapter shows that results of lessons learning via scenarios can be used:

- to prevent loss of memory;
- to list all possible EWSs;
- to identify whether a failure/error/condition represents an EWS;
- to prioritize EWSs.

All preceding claims are illustrated by examples.

Moreover, the list of attributes necessary for the tools of lessons learning according to Benner and Carey (2009) is reproduced in this chapter. The question is discussed whether and under what conditions the scenarios and EWSs can carry all these attributes.

5.4 Early warning signs

5.4.1 Definition of EWSs

In this part, only a brief introduction into the concept of EWSs is sketched. CCPS (2012) writes about incident warning signs, which are subtle indicators of a problem that could lead to an incident. Warning signs precede incidents or contribute to them.

In conventional risk terminology, early warning signs can be understood as an indicator of strengthening a hazard or of weakening a safety measure, which can result in an increase of frequency or severity of consequences of scenarios causing damage. Since both an increase of frequency and an increase of consequence severity cause an increase of risk, then, briefly, an early warning sign is an indicator of an increase of risk.

Outside the risk based schemes of thinking, but not in contradiction with them, the occurrence of EWSs could be interpreted as an increase of vulnerability. Thus, foresight in safety could mean the capability to flag an increase of risk with the help of EWSs.

General explanation of foresight can be found in Chapter 2, Røed-Larsen et al., 2020.

5.4.2 EWSs are part of lessons learned and a result of lessons learning

Within the analysis of lessons learning system functions, processes and practices, Benner and Carey (2009) observe that divergent views exist about whether lessons learned are causes, cause factors, conclusions, findings, issues, statements, recommendations or scenarios described in text in narrative reports.

Clearly, they are right. Large accessible literature about incident investigations and lessons learning is not consistent in terminology and approaches. Nevertheless, in this text it is considered that identification of early warning signs is part of lessons learned. EWSs are considered here to be a desirable result of lessons learning. Consequently, lessons learning tools, like scenarios, are expected to detect EWSs.

5.4.3 Examples of EWSs: Kitchen

A simple example shows that in a known environment, some people tend to identify EWSs intuitively.

Kate and William are married; William is taking a parental leave from work. He takes care of the children and also he cooks. He likes cooking. In connection with cooking, he frequently makes small changes – hopefully improvements – in the kitchen.

Kate is glad that William likes cooking; however, she does not agree with all his improvements in the kitchen. For instance, she does not like the bottle with oil in close proximity to the stove, or a heavy bowl in the shelf above the ceramic hob. In addition, she hates William's habit of leaving the frying pan on the stove unattended.

When they had a disagreement over this the last time, William argued that nothing had happened. Kate answers that all these changes are indicators of problems that could lead to an incident. In accordance with CCPS (2012) she calls them warning signs or early warning signs (EWS) and insists that William should avoid making changes in the kitchen that could lead to increasing the risk.

5.5 Scenarios represent a tool for lessons learning

5.5.1 Example: Intuitive use of scenarios

William, in our example, replies that he does not see anything serious in the changes he made in the kitchen. Kate states that this is because he is not

intentionally imagining any incident scenarios. Thinking about danger with the help of scenarios comes natural to Kate. The experience gained through the realisation of hazards serves as a stimulus to develop this skill that Kate has. The experience does not need to be personal; knowledge-based experience will be enough. When Kate, for instance, sees a picture where a ceramic hob from a kitchen is damaged by a fall of canned food, she realises that any heavy object above the ceramic hob is a hazard, and starts thinking about scenarios initiated by falls of heavy objects, and about relevant preventive/mitigating controls.

5.5.2   Hypotheses about roles of scenarios

This is quite a common way of thinking. Information about incident serves as an empirical information about a hazard and its behaviour. The term behaviour is used here in accordance with Benner and Carey (2009). When they write about behaviour, they mean actions of animate and inanimate actors (examples, in case of Kate and William's kitchen, could be William's behaviour or behaviour of ceramic hob).

It is possible that the ability to spontaneously develop incident scenarios based on experience gained from observing hazard behaviour is a result of evolutionary selection. For example, we know that for our ancestors living in the cave, the presence of the sabre-tooth tiger in the neighbourhood represented a hazard. It is undeniable that the ability to imagine a scenario initiated in this hazard (ability to predict what can happen if a tiger lurks in front of the cave) and the ability to prepare appropriate preventive/mitigating controls in order to minimise the damage caused by the realisation of this hazard was an advantage during human evolution.

Kate bases her identification of EWSs on the idea of possible incident scenarios. She imagines the scenarios of possible fires in the kitchen and therefore she perceives the above-mentioned EWSs as unacceptable. Kate actually says what is well known from risk analysis:

- Scenarios make it possible to foresee the risk comprehensively.
- Scenarios are a practical tool for thinking about risk.

In addition, since the EWSs are indicators of increased risk described by scenarios, it is expected that Kate may add:

- Early warning signs (EWSs) can be derived from scenarios.

- Scenarios are a practical tool for identifying and prioritising the EWSs.

This set of statements or hypotheses about roles of scenarios will be used as milestones in the following text. First, the usefulness of scenarios for lessons learning will be highlighted. Then it will be shown that (i) lessons learning using scenarios can reach EWSs through the identification of causal events, (ii) results of lessons learning via scenarios can be used for various purposes, and (iii) the use of scenarios as a tool to obtain EWSs has many of the required attributes of lessons learning tools.

5.5.3   Scenarios make it possible to foresee the risk comprehensively

Origins of danger are called hazards. Definition from CCPS (2008) states that hazard is a physical or chemical condition that has the potential for causing harm. Hazards in the industrial environment is usually associated to the presence of a dangerous substance or a possibility of an undesirable reaction or an accumulation of energy.

In case of William's kitchen, the three hazards identified are the following: bottle with oil close to the stove, potential for oil in the frying pan to ignite, and heavy bowl on the shelf above the ceramic hob falling. In case of an industrial plant, the three hazards may be the following: presence of volumes of explosives, potential of decomposition reaction in the explosive, and the energy of compressed air in piping of filling machine.

Hazards can be systematically identified. Several suitable techniques were developed for this purpose. Probably the most universal techniques for hazard identification in industrial installations are FMEA and HAZOP (See CCPS 2008).

Mere identification of hazards however does not say too much about the risk that is connected with a process or with an operated system. Presence of the bottle with oil in the kitchen means only that the risk connected with the use of kitchen cannot be zero. Three reasons exist why mere knowledge about present hazards is not enough:

- The article by Kaplan and Garrick (1981) reminds us that risk increases with the increasing presence of hazards, but it also decreases according to measures which are intended to keep control over hazards. Some of such measures may prevent realisations of hazards, and others may mitigate the effects of realisations. Various types of these measures are called barriers, safeguards, regulations, or layers of protection. Here we will mostly use the

term controls or preventive/mitigating controls, which seem to be the most general.

- The risk is not only influenced by the interaction of hazards and controls, but also by the interaction of hazards among themselves. This refers to the terms domino effect or knock-on effect. For example, the ignition of the oil in the pan can develop into the ignition of the oil inside the bottle.
- The magnitude of the risk is also influenced by local environmental conditions that change, regardless of hazards and controls. For example, the development of a fire in the kitchen may be different depending on whether the door and/or the window are open. The risk of the industrial plant varies according to the propagation of the shock waves and the gas clouds.

All three reasons mentioned above explain that scenarios describe the complexity of real danger much better than mere hazards. Kaplan and Garrick (1981) consequently argued for this and defined risk as a set of scenarios s_i, each of which has a probability p_i and a consequence c_i. Although this approach has its limits which are discussed in (Aven, 2008), it is preferable when we think about the use of scenarios.

5.5.4 Scenarios are a practical tool for thinking about risk

Crowl and Louvar (2011) state that scenario is a description of the events that result in an incident or accident. According to Marshall and Ruhemann (2001) scenarios describe how the situations can develop when a hazard starts to realise. The above verb "realise" means the process of an event or events by which the *potential* in a hazardous system becomes *actual*. In accordance with this idea, the scenarios are sequences of events in which the first event (initiating event) starts the realisation of a hazard. The sequence can, but does not have to, include other - developing - events in addition to the initiating event. See Figure 1. Developing events may be undesirable events in the hazard, failures or successes of different controls, application of different environmental conditions, or escalation of development to other hazards present.
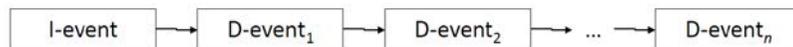
where n is any natural number or zero

Figure 1. Scenario.

In the kitchen, Kate thinks about fire scenarios; in the plant, she would imagine explosions related to the production of emulsion explosive charges. For example, in the kitchen, a scenario may start by the ignition of the oil in the frying pan; followed by extinguishing of fire or by escalation of fire triggering other hazards, including the oil bottle in the vicinity of the stove; and develop until the fire spreads to the entire fire load in the kitchen.

Such scenarios are called incident scenarios since they cause non-negligible damage. Such scenarios have two substantial properties:

1. Each scenario represents one possible interaction of real conditions in the process/system. The scenarios not only take into account the hazards in the process/system but also the ways in which these hazards are realised, how the controls fail or succeed, how the hazards interact and how environmental conditions contribute to the development of the incident.

2. Each scenario represents one contribution to the risk of process/system. Each incident scenario represents one possibility of how damage may arise in the process/system. Or each scenario represents one part of the risk according to the classical definition by Kaplan and Garrick (1981).

Kate obviously has in mind both these two properties when saying that scenarios make it possible to see the risk comprehensively. In accordance with the article by Kaplan and Garrick (1981), the risk of process/system is for her a set of all conceivable incident scenarios in the process/system.

Kate also feels how important the description of scenarios is for thinking about risk. If the scenario describes a specific accident/incident that happened in the past, its description will contain the information relevant to the understanding of its origin, i.e. the origin of this specific part of risk. If a scenario describes a generic incident/accident that may happen in the future, it in fact represents a group of similar specific scenarios. It is accordingly called a representative scenario and explains the origin of a subset of risk.

Having in mind all the preceding properties of scenarios, we start to be aware of another very important feature of scenarios: their clarity and transparency makes them very powerful in explaining the risk to the general public. Scenarios may serve as an extremely useful communication tool with the general public.

5.5.5 Incident scenarios may be results of prospection

Prospective scenarios arise by developing initiating events in hazards. Event trees are commonly used to represent and create them as it is described e.g. by CCPS (2000, 2008). An example event tree is in Figure 2. Figure 3 contains the same list of scenarios as the event tree in Figure 2.

When an analyst constructs an event tree, he starts from a known initiating event in a hazard, knows the behaviour of hazards, and is aware of controls and environmental conditions. He usually begins by considering how and in what order after the initiating event, the controls and environmental conditions should be applied to minimize the damage caused. This sequence of events is called success scenario. Success scenario defines heading of event tree. In Figure 2 it consists of the initiating event and three developing events.

The analyst then considers what the negations of controls and environmental conditions may cause in the development of an incident. He records the findings in the tree graph below the heading. This way he creates a list of prospective incident scenarios, which start with the selected initiating event.
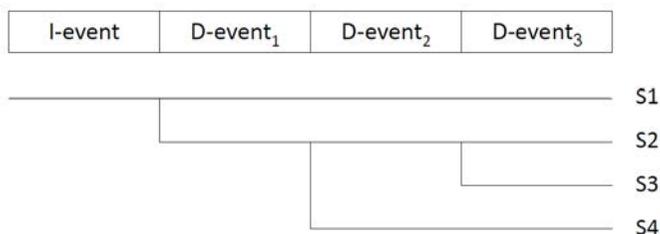
Figure 2. Event tree.
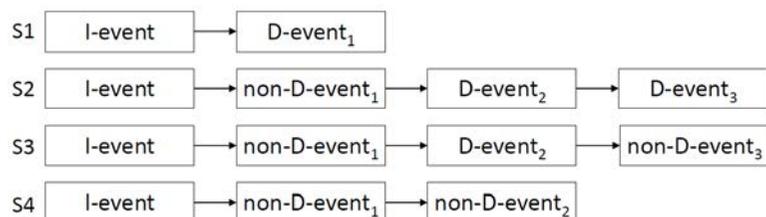
Figure 3. List of incident scenarios from event tree in Figure 2.

Regarding risk analysis, which is essentially a list of scenarios, sometimes it is said that classical approaches to the identification of possible scenarios, which are described by CCPS (2000, 2001, and 2008), do not necessarily reveal all possible scenarios. Scenario-based techniques such as red-teaming (DoD 2003) and anticipatory failure determination (Kaplan et al 1999) can also be used to challenge existing safety cases, attempting to find gaps in the accident scenarios that have been analysed (Masys 2012).

Scenario-based exercises can also be used for simulation-based training exercises which aim is to improve system resilience by strengthening operators' knowledge of system and safety barriers operations.

5.5.6 More about prospective scenarios

Event trees do not represent the only way to identify the prospective scenarios.

Event sequence in an event tree that starts by an initiating event and resulting in an outcome, may be a relatively long and detailed. But it may be simplified and reduced to a mere pair of initiating event and related outcome. This approach represents a starting point for layer of protection analysis (LOPA) described by CCPS (2001).

A bow tie according to CCPS (2008) or according to Hatch et al. (2019) represents another alternative to an event tree. Bow tie is more detailed than the event tree, since the initiating event is expanded into a tree of event causes.

In Part 5.5.4. it can be noticed that a single prospective scenario usually represents a group of similar specific scenarios and is accordingly called a representative scenario. In majority of cases, when individual prospective scenarios are mentioned they could be replaced by sets of scenarios. If bow-ties were used instead of event trees to illustrate scenarios, this would be evident.

Extensiveness, complexity, and level of detail of representative scenarios depend substantially on how the individual events in the sequence are described. Above all, the resolution whether the sequences are characterised in terms of (i) fulfilment of safety functions, (ii) intervention of protection systems or occurrence of physical phenomena, (iii) successes and failures of individual components, may substantially influence the extensiveness and specificity of scenarios. Zio (2007) discusses this problem in more detail.

However, whether event trees, LOPA pairs or bow-ties are used to represent scenarios, it can always be said that scenarios can be used in fully quantitative, semi-quantitative and fully non-quantitative modes. The first one is suitable for quantitative risk estimation, the latter for communicating on risk with non-specialized people, which is mentioned at the end of Part 5.5.4.

5.5.7 Incident scenarios may be results of retrospection

Retrospective accident scenarios are created as a result of the reconstruction of incidents in the process/system. According to Johnson (2003), such reconstruction is always necessary during the investigation regardless of the method used to analyse the causes of the incident.

Retrospective scenario is a sequence of events. But its first event does not necessarily have to be identical with the initiating event that starts the realisation of a hazard. The sequence can, but does not have to, include developing events. Developing events are not limited to undesirable events in hazards, failures or successes of different controls, applications of different environmental conditions, or escalations of development to other present hazards. In addition, the most surprising and unpleasant difference of retrospective scenarios from prospective scenarios is that they do not consist only of one line of events but may variously branch and splice. Example of a retrospective scenario is shown in Figure 4.
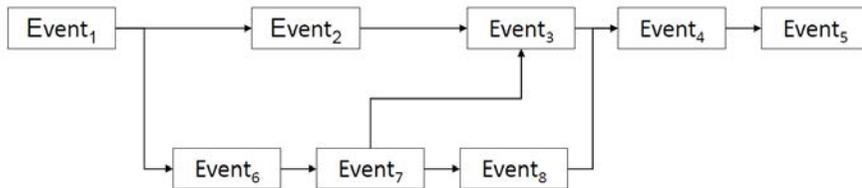
Figure 4. Example retrospective scenario.

Retrospective scenarios generally represent multilinear sequences as described by Benner (1975). Excessive events result from a descriptive effort that does not take into account only undesirable events in hazards, failures or successes of different controls, applications of different environmental conditions, or escalations of development to other present hazards. Branching and splicing (examples can be seen in Figure 4) is result of taking into account different actors, as described by Benner (1975).

In the kitchen, $Event_1$ can be "Start of frying in the pan", $Event_2$ "Chicken breast is fried in oil", $Event_6$ "William attempts to take the bowl from the shelf above pan", $Event_7$ is "Bowl falls down on the pan and ceramic hob", $Event_8$ may be "Ceramic panel above electric heaters is broken". $Event_3$ may be "Frying pan overturns" and $Event_4$ may be "Spilled oil ignites". In such a case $Event_7$ will be determined as an initiating event according to Benner's (1975) definition. $Event_8$ and $Event_3$ are failures of controls, and $Event_4$ is an undesirable event in hazard.

Again, the picture resulting from the retrospection may be more complicated as events in the diagram (for instance $Event_7$), may be expanded into trees of their causes. The overall picture may then resemble a bow-tie diagram.

5.5.8 Both prospective and retrospective scenarios can be used for lessons learning

Today's designer or an operator of an industrial system, or for instance, a food safety regulator (see Afonso at al., 2017) may think about the realisation of hazards just like Kate thinks about heavy objects over a ceramic hob or like a cave dweller thought about a lurking tiger. For such thinking, it is necessary to know the behaviour of the relevant hazards, to understand them based on natural science or to have experience with them. Scenarios can then be used as a tool that supports the thinking. The effectiveness of such thinking can be enhanced by adopting appropriate techniques.

Benner and Carey (2009) do not limit the use of the term "investigation-oriented lessons learning data sources" only to retrospection related to experienced accidents or incidents, but also to potential or hypothesised accidents or incidents originating from hazard and risk analyses.

Similarly, also here investigation is related both to retrospection and prospection. Incident scenarios can arise in two ways: as a result of retrospection (incident analysis) or prospection (risk analysis). These two options will be discussed in detail in Part 5.

5.5.9 Desirable attributes of scenarios as tools for lessons learning

Based on the preceding considerations, together with Kate we would like to use the scenarios as a tool for lessons learning, namely for the development of EWSs. Benner and Carey (2009) analysed desired attributes of lessons learning tools. They showed that the development of lessons learned can be divided into investigation and documentation. Investigation should support production of lessons-to-be-

learned source data. Documentation should facilitate satisfaction of desired user criteria.

If scenarios are to be used in investigation, they need attributes like:

1. A stated investigation purpose of providing lessons learned leading to changed future behaviours.

- For example, Kate could intend to investigate accidents in all the kitchens of all the Williams living throughout the UK, to improve their safety during cooking.

2. An input-output framework for describing what happened, enabling lessons learned data sets, to describe behaviours to change in non-judgmental and logically verifiable terms.

- For example, Kate would require every kitchen accident to be described as a finite set of sequences of simple sentences beginning with an initiating event and ending with a damage description.

3. A focus on behaviour data acquisition and processing, to enhance efficient documentation of lessons learned from accident-generated lessons learning source data.

- For example, the sentences in the sequences would have to describe actions of animate or inanimate actors, or behaviour.

4. Specifications for behavioural building block structure, grammar, syntax, and vocabulary and a structure for input data documentation, to ensure data consistency and economy, and facilitate data coupling and support for documenting lessons learned.

5. Machine support for input data sequencing, parsing, coupling, concatenation, data set display and expansion capabilities, to facilitate lessons learned processing and dissemination, and to reduce latency.

6. Objective quality assurance and validation process for behavioural data sets.

The three above-mentioned examples show that scenarios are able to fulfil the requirements of attributes #1 to #3. Attribute #4 requires the introduction of certain standards on how the events are described, and how the conditions are replaced by the events. Vocabulary can be limited to definite lists of actors and/or

actions using checklists. Such a standardization is easier to reach in the industrial environment than in a kitchen since in the industry it may be supported by a marking system. In the kitchen it would need to use, e.g. for the bottle with oil, always the same term.

Attribute #5 is connected with the use of computers for recording the scenarios, which is achievable especially when the attribute #4 is fulfilled. Attribute #6 states only that the use of scenarios for lessons learning cannot be considered satisfactory if it is not subjected to quality assurance.

5.5.10 Attributes of scenarios desirable for documentation part of lessons learning

According to Benner and Carey (2009), desired attributes for documentation would include:

1. Efficient tools to facilitate documentation of behaviour data sets, and reduced latency.

2. Specifications for lessons learned behavioural data outputs meeting users' needs, harmonized with other learning organisation lessons learned sources or knowledge management artefacts, with maximised signal-to-noise ratios, providing context, minimising interpretive and analytical workload for users, and reducing latency.

3. Machine lessons learned processing support and repository uploading capabilities, to accelerate lessons learned documentation and deployment into all repositories.

4. Internet lessons learned output data repository and notification capabilities, to facilitate "push" or "pull" lessons learned data dissemination, enable wide deployment, and minimise latency.

5. Rapid repository access, search and filter capability, to minimise user access time, cost and workloads.

6. Objective lessons learned quality assurance and validation functions, to enable developer to ensure lessons learned quality before entry into repositories.

7. Lessons learned repository modification or updating capability, to ensure lasting lessons learning quality.

The second set of attributes by Benner and Carey (2009) seem to be very demanding, much different from the state existing in many industrial environments. Evidently, they saw the attributes from a very general perspective and created a description of "an ideal state". Doubtlessly, without the use of computerised database tools they cannot be achieved nor even approached. But hopefully in very simple cases as our kitchen case, the documentation of scenarios may be reasonably used for lessons learning even if it does not achieve the most ambitious standards.

5.6 Lessons learning can reach EWSs through the identification of causal events

5.6.1 Causal events in retrospective incident scenarios

The main purpose of lessons learning is to identify what behaviour of actors was wrong and what behaviour has to be improved in order to prevent or mitigate the recurrence of an incident. Identification of wrong and improvable behaviour is possible, as soon as the incident scenario is reconstructed. The reconstructed scenario usually contains not only individual events, but also a description of context in which the events occurred. Context is described as a set of conditions.

If the analysis of the retrospective scenario is aimed at preventing the repetition of the same or similar scenarios, it must focus on those events in the scenario that worsen the control over a hazard. The events in the sequence have to be identified that influenced unfavourably the behaviour of actors and thus, contributed to the incident.

These events are often called causal factors. CCPS (2003) defines causal factor as a negative event or undesirable condition that, if eliminated, would have either prevented the occurrence (= incident scenario) or reduced its severity or frequency. Since this definition permits causal factor to be a condition, we will modify it slightly. We will require the conditions always to be linked to events which context they describe (we suppose that such a state is always achievable). Then we leave the term causal factor and define causal event as a negative event, including its context that if eliminated would have either prevented the occurrence or reduced its severity or frequency.

Let us suppose that $Event_5$ in Figure 4 is "William uses the fire extinguisher". This event itself does not seem to be a causal event. But if $Event_5$ happened in the context that William was not able to initially use the extinguisher and hence, the extinguishing started much later than possible, then the $Event_5$, including its context, would visibly be a causal event.

Another approach to retrospective scenarios requires to replace all the conditions within the chart by (sequences of) events. This approach relates to the explanations added as Epilogue to the original article by Benner (1975). The advantage is that constraining the flow chart to events is always possible and solves the problem. (Way to the exclusion of conditions is commented in the end of Part 5.6.6) Causal factors will be then represented only by events and be identical to causal events.

Ideally, the set of causal events represents the set of necessary and sufficient events explaining HOW the incident occurred, while the scenario itself explains WHAT occurred. A reconstructed incident scenario is reduced to a set of causal events during the retrospective incident analysis.

For the incident scenario that might be represented by Figure 4, the set of causal events is $Event_7$ (bowl falls on the pan and ceramic hob), $Event_8$ (ceramic panel above electric heaters is broken), $Event_3$ (Frying pan overturns), $Event_4$ (spilled oil ignites) and $Event_5$ (William uses fire extinguisher later than possible).

5.6.2 Causal events in prospective incident scenarios

Analysis of incident scenarios using event trees uncovers possible interactions of real actors in the system, i.e. interactions of present hazards, controls and environmental conditions. For most of the events in the tree, it is valid that they can change within a certain range without changing the scenario. For example, if in the tree in Figure 2 the initiating event is the ignition of oil in the pan, and the first developing event is a fire intervention with a lid, then the fire intervention can take place at any time within a certain time interval of about tens of seconds without changing the course of the scenario. An event tree analyst considers the ranges within which the events can be changed. Individual scenarios from the tree thus represent whole classes of somewhat different scenarios, which however do not differ in qualitative terms, i.e. by the type of events involved. The event tree thus contains representative incident scenarios. For more details see, for example, article by Kaplan et al. (2001).

Prospective scenario analysis can be used even before the precise form of the individual conditions in the process/system is known. Once an initiating event is defined, all the safety functions that are required to mitigate the incident must be defined and organised according to their time of intervention as Zio (2007) describes it. In the case of ignition in the frying pan, we could consider immediate firefighting, limitation of propagation, delayed firefighting, and extinguishing by an external fire brigade. Defining safety functions can be very useful in the design phase because it can be used to define controls.

Prospective analysis typically seeks to investigate systematically all representative initiating events and related incident scenarios. Scenarios created by prospective analysis take the form of conjunctions of events from which no event can be removed. When thinking about risk, events in scenarios that represent degradation of control over hazards are at the heart of interest. If the convention is kept that the tree heading contains a success scenario, then events that represent degradation are both initiating events and all events that negate successes from the heading, i.e. all the events starting in Figs 2 and 3 with the word "non".

Above we defined causal event as a negative event including its context that if eliminated it would have either prevented the occurrence or reduced its severity or frequency. This is the exact description of both initiating events and negating events in the event tree. Thus, initiating event and negating events in the event tree can be called causal events.

Therefore, prospective analysis using event trees can serve as a tool for the systematic identification of all possible (representative) causal events in the process/system. Visibly, this conclusion does not depend on the form of scenarios mentioned in Part 5.5.6.

In addition, it can be shown that the simplified prospective scenarios used in the layer of protection analysis by CCPS (2001) can serve as a tool for the identification of possible causal events in the process/system. In this case, failures of layers of protection can be identified as causal events.

Similarly, we suppose that all other methods of identification of prospective scenarios can be used to identify causal events.

5.6.3   Comparison of role of causal events in prospection and retrospection

While prospective analysis attempts to predict all possible causal events that might occur, retrospective analysis identifies the combination of causal events that actually occurred. If analyses are flawless, then retrospective analysis should result in one of the scenarios created by prospective analysis.

Nevertheless, if we have a set of possible incident scenarios created by a prospective analysis for the process/system, it is not certain that the scenario generated by the incident retrospection in this process/system can be quickly identified with one of the prospective scenarios. There may be several reasons for unsuccessful identification:

(i) Retrospective analysis may mix several scenarios that took place concurrently;

(ii) Scenario events in retrospective analysis are determined in more detail than those in prospective scenarios;

(iii) Certain conditions that worsen the control over a hazard in real undesirable event in the process/system may be omitted in prospective analysis.

These practical findings represent some of the motivations for achieving the attributes quoted in Parts 5.5.9 and 5.5.10 when using scenarios as a lessons learning tool. Theoretically, such problems should not arise if all the attributions according to 5.5.9 and 5.5.10 are reached. Nevertheless, we know that reality still is quite far from Benner and Carey's (2009) ideal.

Nevertheless, it is true that the most important common finding is as follows: in both prospective and retrospective scenario analysis, the main outcome in terms of safety is always a set of events that represent a worsening of control over the hazards to which our attention should be focused. In other words, in both cases our interest focuses on events called causal events.

5.6.4   Scenarios make visible the threatening conditions in the process/system

The previous parts have shown that *any incident scenario can be reduced to a set of causal events*. The set of causal events represents a combination of events worsening the control over the hazards. They are at the same time the combination of necessary and sufficient conditions for consequences and frequency of this incident scenario. The causal events can be represented by the following:

- initiating event in the hazard, or
- failures of the measures intended to mitigate the realisation of a hazard, or
- failure of the measures intended to prevent the realisation of additional hazard, or
- events adversely affecting the environmental conditions influencing the realisation of hazards.

This result shows that the scenarios make visible the ways in which hazards realise (come to be) in a particular process/system. They visualise the real role of hazards and related controls and environmental conditions in a particular process/system. This visualisation is the basic purpose of both risk analysis and undesirable event analysis.

5.6.5 Better than prospection or retrospection is the combination of both

Retrospectively, i.e. based on experience with specific undesirable events, only specific accident scenarios can be revealed within the incident cause analysis. From a logical point of view, this is an inductive process. Its advantage is that it identifies the real weaknesses of control over the hazards, usually the most likely ones. It may also reveal weaknesses that within risk analysis remain hidden from our eyes for their delicacy. The disadvantage is that it reveals only some weaknesses and scenarios, not necessarily those that most contribute to the risk. The disadvantage may also be that, in the analysis, causal events are not identified in a sufficient manner. The results may mistakenly adhere only to the partial weakness, which is only a contribution to the general causal event.

Prospectively, i.e. based on a process/system analysis, the risk analysis can reveal theoretically all possible incident scenarios. From a logic point of view, this process is deductive. (This means, of course, that it also contains the inductive component - general rules on behaviour of hazards and controls based on experience). The advantage of this approach is that it systematically searches for all weaknesses in the control across all the hazards. It is able to reveal all the weaknesses and scenarios, including those with low frequencies. It can also reveal weaknesses that, by mere application of experience, remain hidden from our eyes. The disadvantage of the prospective approach, however, is that the analysis cannot avoid various neglects and simplifications because of which some substantial interactions of hazards and controls may be omitted. Hence, the outcome of the prospection may appear to be complete, but in reality, substantial scenarios are missing.

Since it is difficult to avoid the above-mentioned errors when using these approaches, the combination of a prospective and a retrospective approach seems to be a practical and realistic approach to identifying scenarios.

5.6.6 Early warning signs are causes and indicators of causal events

We realised in the previous parts above that *a set of scenarios makes the risk of the process/system visible as a set of sets of causal events*. As we have already mentioned in Part 5.4.1, the essence of foresight is the capability to see EWSs or indicators of problems that could lead to an incident, or the indicators of risk increase. In the context in which risk is decomposed into incident scenarios, and incident scenarios are in turn decomposed into causal events, foresight thus, means the ability to see the signs that some identified causal events could actually occur. In particular, we would like to be able to see signs of possible occurrence of causal events that contribute most importantly to the risk.

It follows from the previous paragraph that the EWSs can be identified as the causes of causal events including causal events themselves, or indicators of causal events, or indicators of causes of causal events. (Among indicators, the leading indicators are preferred.)

This finding means that a correct and complete identification of causal events is of essential importance. An unidentified causal event (CE) represents an invisible set of EWSs, existence and importance of which stay unknown.

In this context it has to be strongly recommended to follow the Epilogue by Benner (1975) and exclude any possibility that a causal event would stay hidden given the presence of conditions within the scenario (retrospective) description. There are various ways how to do this. An approach shown by Accou and Reniers (2018) is a promising way that excludes conditions from a descriptive chart of a scenario and replaces conditions by events. The universal model of safety management activities (safety fractal) promises to help identify a possibility and sort of such a replacement for any condition within the chart.

Also it is recommended to perform a check of identified causal events by rewriting each of them as an adverse influence acting on a vulnerable target due to missing barriers or regulations. This approach originates in MORT by Johnson (1973) and warrants that all CEs identified correspond with the definition of causal events.

5.6.7 Possible approaches to identification of event causes

Unfortunately, the concept of causes does not have clear and unambiguous content. If we talk about the causes, we can talk about many kinds of events and ideas. Nevertheless, it can be repeated here, that the EWSs represent causes of causal events, *whatever the causes mean*.

In technical practice, at least direct causes and underlying causes are usually distinguished. Smaller differences exist with respect to direct causes. They are physically detectable failures, errors, states, conditions, the combination of which leads to an occurrence of causal event.

But there are quite different ideas in various approaches to incident analysis about what are the underlying causes. In the relatively common root cause analysis (RCA) methods, the underlying causes are called root causes and represent deficiencies in the implementation of a safety management system. They could also be referred to as organisational causes.

Verschueren (2018) is focused on the organisational causes and their relation to EWSs. He underlines the importance of organisational dysfunctionalities. According to Verschueren (2018) organisational dysfunctionalities can be detected and can act as EWSs.

General acceptance of contemporary focus on organisational causes is confirmed in Hollnagel (2014):"In the thinking about types of causes, we see a development that goes from technology to the human factor and, most recently, to the organisations and to culture." In accordance with this, most part of contemporary methods of cause analysis agrees with the idea that organisational causes have to be searched for. They attempt to identify them.

A hierarchy of checklists, called root cause map, is often used to determine underlying causes in RCAs and improved RCAs. Such an approach is described in CCPS (2003). An example of elaborated analysis method that is nowadays used in industry can be found in a paper by Nicolescu (2018) where the method of the Investigation Body of Norway, AIBN, is applied to identify causal events (direct causes) and a tool named SMS wheel is applied in order to identify underlying causes.

Improved RCAs such as described by Ferjencik (2014) would include also the underlying causes in safety culture or attitudes of local management. As shown in the article by Ferjencik, guidelines by CCPS (2007) are useful for this purpose.

There are at least two examples of alternative methods to determine direct and underlying causes. Symptoms would be determined with analysis by ESReDA (2009). Failing processes would be identified instead of root and underlying causes in an analysis by Leveson (2004). Nevertheless, for both approaches, the identification of causal events according to the definition used here would be the necessary starting point. For Leveson's approach, it is shown in Stoop and Benner (2015). Leveson's analysis process starts with a step *Identify the systems and hazards involved in the loss*. This requirement can be translated into *Identify the controls and hazards involved in the loss*. Causal events point to such an identification.

This diversity means that EWSs and searching for EWSs can have very variable forms. While these differences in our understanding of causes can discourage us, they all point to the same general fact: EWSs can be determined from incident scenarios as (partial) causes of relevant causal events.

Example: The determination of causal events is very easy in conventional event trees. Four causal events are present in Figure 2 according to Figure 3: I-event, non-D-event$_1$, non-D-event$_2$, and non-D-event$_3$.

Various techniques and approaches can be used for the identification of causes of causal events. Fault tree analysis (FTA) that is recommended in book CCPS (2003), is very productive in prospective analysis. Figure 5 shows possible results of application of FTA to two causal events. It can be observed from Figures 2, 3 and 5 that cause1, cause2, and cause3 represent EWSs for all scenarios S1 to S4. Cause4 and cause5 are EWSs only for scenario S3. Cause2 indicates the possibility of formation of both causal events at the same time. Cause2 may represent a sort of common cause failure. Typically, the EWSs with common-cause nature may be the deficiencies in the local safety management system i.e. underlying causes.
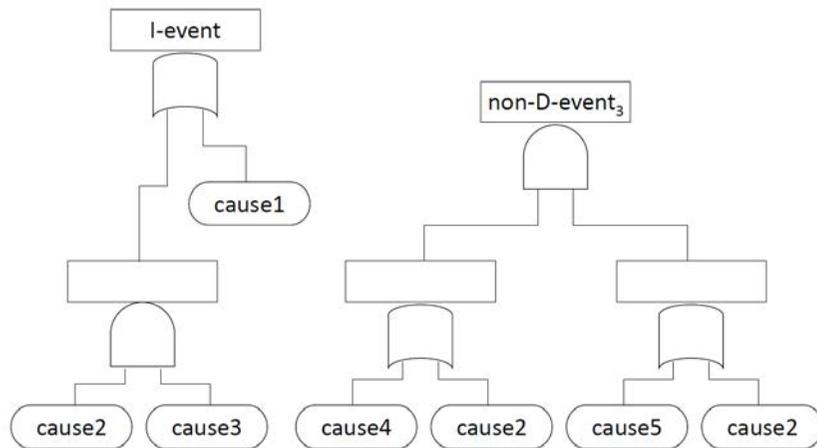
Figure 5. Causes of two causal events from Figure 2 and 3.

5.6.8   Steps to the identification of EWSs

The identification of EWSs begins when the incident scenarios are constructed. They make visible the realisation of hazards, which is the main purpose of the construction of incident scenarios. Scenarios allow the identification of causal events. They make visible the roles of hazards and controls of hazards. Once causal events are known, a way to make EWSs visible is open. Therefore, the visibility of the EWSs emerges through the visualisation of the role of hazards and controls of hazards.

Scenarios can help see the EWSs in two steps. In the first step, we determine the causal events in the incident scenarios; in the second step we determine the causes of the established causal events and indicators of causal events and their causes.

In prospective analysis, causal events are determined as:

- events in hazards that initiate realisation of hazards,
- events in hazards that escalate damages,
- events that represent failure of controls over realised hazards,
- events that allow damage escalation by setting up adverse environmental conditions.

In a retrospective analysis, causal events are selected as events that meet the definition of causal event.

In case of prospective analysis both the elaborated form of scenarios that is used within quantitative risk analysis (and modelled with the help of ETA, FTA, and HRA or with the help of bow ties), and the simplified form of scenarios typical for layer of protection analysis (modelled as an initiating event – consequence pair) can be exploited.

Practical note:

To be used efficiently as a tool for EWS identification, the set of scenarios should not be excessively wide, the scenarios should not be too specific, and the identification of EWS should not be limited to specific direct causes represented by the failures/errors. The set of scenarios should:

- be limited to selected critical scenarios,
- have description of scenarios that prefer functions (not elements),
- have EWSs that are identified in underlying layers, too, i.e. as deficiencies of the safety management system.

5.6.9   Variability in identification of EWSs

As it is visible from preceding parts of the chapter, the detailed understanding on what are early warning signs can be substantially variable. There is no single possible way to identify EWSs. Nevertheless, a few important findings can be stated:

- Based on the definition and procedure for identification of EWSs, it can be understood why we may have known and unknown EWSs and what their existence may signify.
- Based on the understanding of relations of causes, it can be understood why there may be synergic relationships among EWSs.
- Based on the fact that EWSs may be causes it can be understood that Cube (Chapter 8, Stoop et al., 2020) may be applied to identifications and checks of EWSs.

5.7 Application of scenarios as a tool for lessons learning

5.7.1 Example: Kitchen prospection

A frying pan filled with oil is a hazard in the kitchen. Kate worries that the oil in the pan may ignite - she considers the ignition of the oil in the pan to be a possible initiating event. Rapid extinguishing by laying the lid on the pan minimizes damage after the initiating event. If this does not happen, further development depends on whether there is another hazard near the pan - a plastic bottle of oil. If it is not there, the damage is minimized, i.e. it can be expected that the oil in the pan will burn out, the smoke will cause damage, but the fire will not expand further. If the bottle with cooking oil is present and stays nearby, it is a matter of time when a large amount of burning oil is spilled on the stove and on the floor. At this point, the rapid use of a suitable fire extinguisher can minimize damage. If the extinguisher is not used quickly, the fire will spread across the room. Further development depends on whether the door is opened into the adjoining dining room or whether it is closed. Closed door minimizes damage, in the sense that when the fire breaks out, the window and becomes noticeable from the outside of the house, no further rooms are hit so far. If a fire-fighting car arrives in time, it will save most of the house from the fire. The success scenario consists of an initiating event and five developing events.
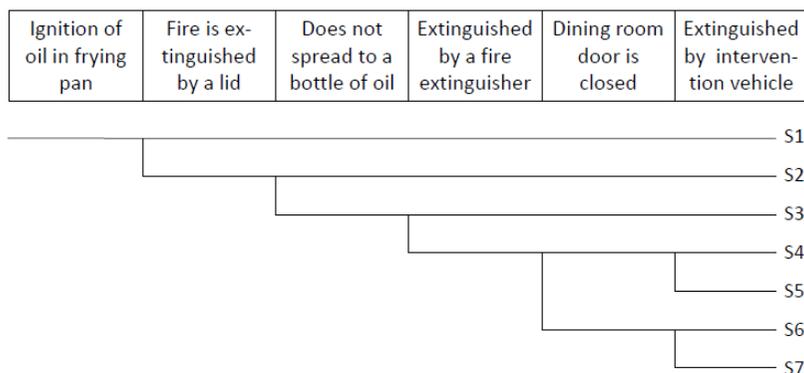
Figure 6. Analysis of possible developments of ignition of oil in frying pan.

Three of the developing events are the use of controls, one event is the realisation of another hazard, and one can be considered to be the application of the environmental condition. The entire event tree (Figure 8) contains seven incident scenarios.

Six causal events are determined: ignition of oil in frying pan; fire is not extinguished by a lid; fire spreads to a bottle of oil; fire is not extinguished by a fire extinguisher; dining room door is open; fire cannot be extinguished by intervention vehicle. EWSs in the kitchen can be determined as analysis results of possible causes of individual causal events. For example, William's habit of leaving the frying pan unattended may contribute to the causes of the initiating event and is the cause of the failure of the first developing event. It is therefore a clear early warning signal. Presence of the bottle of oil in close vicinity of ceramic hob, as well as the absence of fire extinguisher in the kitchen will be identified among EWSs.

5.7.2 Example: Kitchen retrospection

Let's imagine that a fire broke out in the neighbourhood of William and Kate. The fire destroyed the neighbour's kitchen. Investigations have shown that the real incident scenario in the kitchen took place as the scenario in Figure 7 shows.
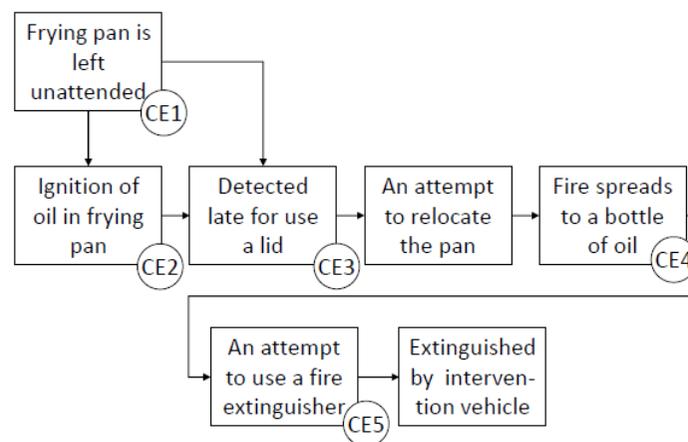
Figure 7. Scenario of real incident in William's neighbour's kitchen.

The scenario recalls scenarios S5 and S7 from Figure 6. Because the information about the dining room door status is missing in the scenario, it is not to be expected that this reconstruction of the incident could identify the EWSs causing the door to be opened. On the contrary, the causal event CE1 is identified in the

reconstructed scenario, which is missing in the scenarios in Figure 6. The presence of the extra causal event in the scenario can be explained by point (ii) in Part 5.6.3 Causal event CE1 is the cause of causal events that we find in scenarios S5 and S7 from Figure 6. Therefore, this external incident does not bring Kate any new facts she would not know from the prospection. On the other hand, this retrospection makes William change his undesirable habit.

5.7.3 Example: Industrial unit prediction

Let us move from the kitchen into the industrial environment. As an example we will use a unit for production of emulsion explosive charges. (The example is inspired by Ferjencik and Dechy, 2016.) Figure 8 shows a basic arrangement of this plant. Protective walls surround a light building inside of which the automatic filling machine produces explosive charges from the explosive paste. In this environment, William may play a role of personnel and Kate represents his manager.

Initiation of detonation during the start of filling machine represents a possible initiating event in unit for production of emulsion explosive charges. In case that the individual events in the sequence are described as fulfilment of safety functions (see (i) in Part 5.5.6), resulting event tree may look as it is shown in Figure 9.
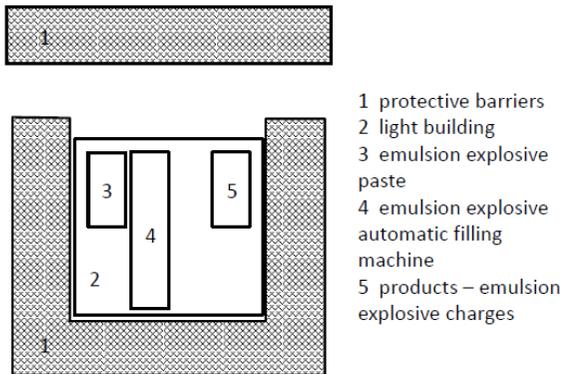
Figure 8. Unit for production of emulsion explosive charges (bird's-eye view).
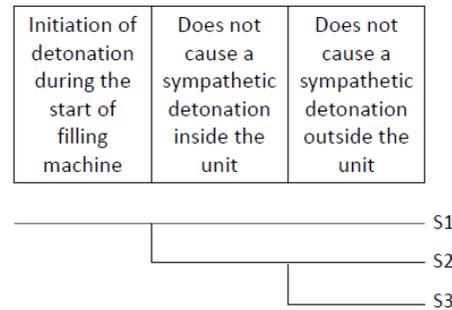
Figure 9. Scenarios of possible incidents in unit for production of emulsion explosive charges.

Causes of sympathetic detonations (inevitable transmission of detonation) have to be analysed in order to identify early warning signs. Typical EWSs that correspond with the second and third identified causal events are excess amount of explosives, inappropriate deployment of explosives, and insufficient resistance of unit.

5.7.4 Example: Industrial object retrospection

Figure 10 shows the scenario that actually occurred in unit for production of emulsion explosive charges.
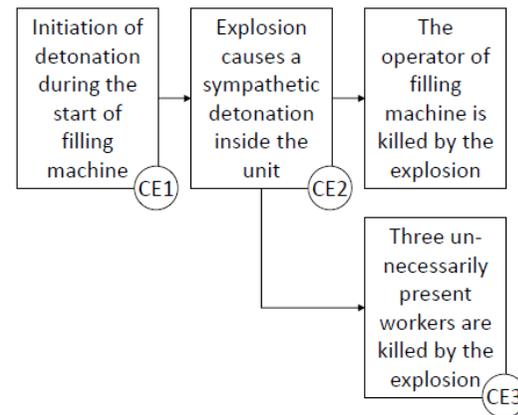
Figure 10. Scenario of real incident in unit for production of emulsion explosive charges.

The scenario contains causal event CE3 that is not identified in the event tree in Figure 9. CE3 represents a fatal impact of a shock wave on three persons present in the he building due to missing protective barriers that could warrant their protection and due to missing regulations that could warrant their absence in the room. This reformulation in accordance with MORT (Johnson, 1973) shows that CE3 really represents a causal event. In this case, the real event revealed a deficiency in the prospective analysis of Figure 9. As stated in Part 5.6.3, point (iii), it was overlooked that controls during the start of filling machine should also include the care of ensuring the absence of surplus persons in the building. In this case, the retrospective analysis reveals EWSs that prospective analysis was not able to detect. An event tree suitable for the identification of relevant EWSs would have to be created by extending the event tree of Figure 9. Its head would include the third developing event "Presence of personnel within the reach of detonation effects is minimised". This example, which is taken from real experience, illustrates the opinion that the combination of both prospection and retrospection is the better approach, rather than either of them on their own.

5.7.5   Example: NPP retrospection

Figure 11 is reproduced from the paper by Strucic (2017). The figure describes a real incident scenario from a nuclear power plant. The failure of the chiller condenser coil led to the shutdown of all three units at the site. Colour conventions are applied in the scenario. Events are represented by green rectangles, conditions by blue ovals. Red rhombus represents a causal event. In this case, only one causal event is identified (which is identical with an initiating event). The brown circle describes scenario consequences.

The identification of early warning signs requires, in this case, the analysis of causes why the fouling of the control bay chiller outlet condenser coils resulting in a high temperature of the outlet water may occur.
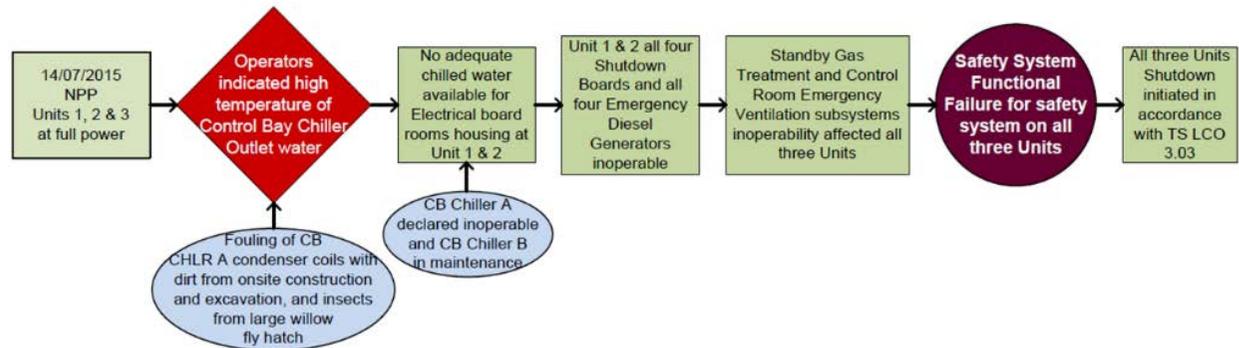
Figure 11. Incident scenario from NPP according to Strucic (2017).

Figure 12 reproduces the summary of the analysis of causes of the causal event from Figure 11 according to Strucic (2017).
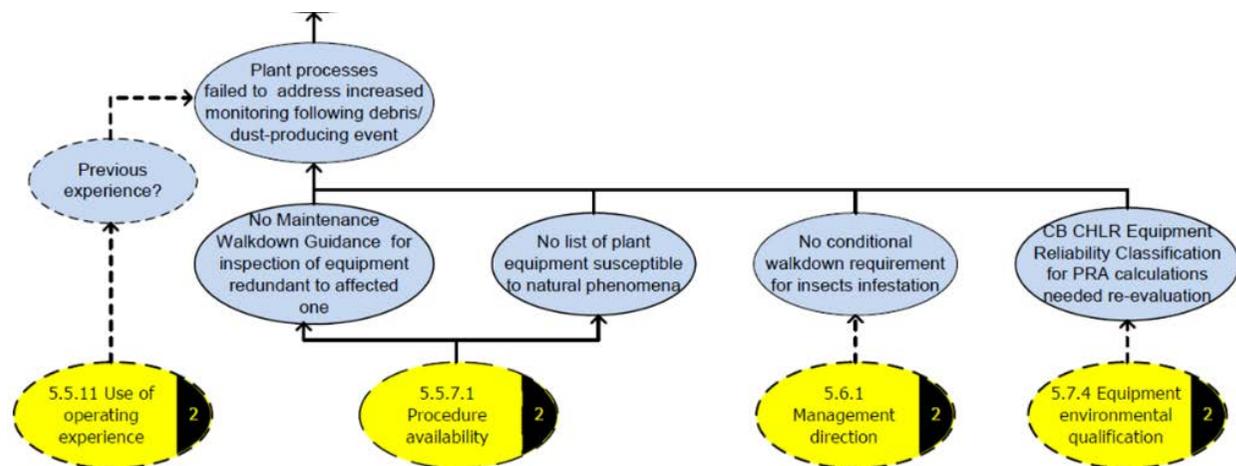
Figure 12. Causes of a causal event from NPP according to Strucic (2017).

Undoubtedly, this graphical summary originates from the analysis that belongs into the family of RCA. This analysis classifies the causes into categories of causes (yellow ovals) in accordance with a manual (i.e. with a list of checklists) and visibly the categories refer to organisational problems. Hence, the analysis uses a prefabricated list of root causes. Terminology of the original analysis is, however, different from the terms used here. Our causal event is called deviation and the term causal factor is reserved for classified root causes.

This cause analysis could be complemented by a fault tree analysing all possible technical causes that can cause fouling of the chiller coil. Some of relevant causes are listed in the description of the condition in Figure 11. Indicators of the causal event or causes of the causal event can be added to the list of EWSs. Increase of temperature on the outlet of chiller or increased presence of willow fly can be examples of such indicators.

5.7.6   Results can be used to list all possible EWSs

Identification of EWSs may result from the identification of causes of specific undesirable events called causal events. Although none of the above examples contains a satisfactory detailed cause analysis, the examples show the way how the identification of early warning signs can be performed and demonstrate that scenarios are suitable for this purpose.

Prospective scenarios are appropriate for the identification of EWSs. They are suitable for a comprehensive analysis of all generic causal events and for the identification of all possible EWSs. This analysis may represent a modification of a systematic risk analysis. Analogous obstacles endanger completeness of the identification of EWSs as completeness of any risk analysis.

Examples also show that retrospective scenarios are appropriate for the identification of EWSs. Nevertheless, the retrospective scenarios cannot be expected to provide a complete list of possible EWSs. They focus our attention only on a certain segment of the complete list. On the other hand, retrospection may easily draw our attention to imperfections and inconsistencies of the specific attempts to identify a complete list of EWSs. This is why in real industrial environment the combination of prospective and retrospective analysis should be considered to be the best possible way to identify early warning signs. Inspiring explanations related to the identification of EWSs in industrial environment can be found in Chapter 7, Strucic, 2020.

5.7.7   Results can be used to prevent loss of memory

Loss of memory means that information, such as:

- the fact that EWSs can arise,
- possible forms of EWSs,
- methods to detect EWSs,
- EWSs that were detected until they are reasonably responded,

(four aspects of memory) is not encoded, stored, or retrieved in minds of humans who can influence form and behaviour of the system.

In accordance with the above description of four aspects of memory, three forms of loss of memory may be distinguished:

- missing knowledge that an EWS can rise in the system,
- missing ability to identify an EWS when it arises,
- missing ability to respond to the EWS that was identified.

These three forms of loss of memory cover all situations, not only those when a specific knowledge or ability has been forgotten (i.e. it was not possible to be retrieved). In addition, situations are covered when this knowledge or ability has never been present (encoded and stored) in memory or has been present (and retrieved) but the will to use it has not existed.

Encoding, storing, and retrieving the above-mentioned information about EWSs represent the documentation of EWSs resulting from lessons learning. Documentation of EWSs can build on the scenario documentation. Documentation of scenarios forms the basis of risk analysis documentation. There is a number of risk analysis documentation technologies ranging from simple procedures to software and database tools.

Kate probably will not require formalised documentation of EWSs resulting from examples in Parts 5.7.1 and 5.7.2. Results of examples in Parts 5.7.3 and 5.7.4 however require proper documentation. The stored results may represent a subset of prioritized process safety information for given socio-technical system according to Wincek (2011). The highest documentation requirements are expected in extremely complicated and sophisticated systems, such as NPP from example in Part 5.7.5. In this case, for instance, a software like Risk Spectrum (refer to www.riskspectrum.com) is conceivable as a documentation tool. More information can be found in Chapter 12, Simic, 2020.

Encoding, storing and making the knowledge about EWSs retrievable are supposed to be actions against the loss of memory. Scenarios as a generally utilised tools make the realisation of EWSs documentation possible.

5.7.8 Results can be used to identify whether a failure/error/condition represents an EWS

Any form of scenarios between an elaborated form of scenarios typical for quantitative risk analysis (modelled with the help of ETA, FTA, and HRA) and a simplified form of scenarios typical for layer of protection analysis (modelled as an initiating event – consequence pair) is expected to be documented.

Regardless of what form of scenarios are being used, the documentation of lessons learning is expected to contain, in addition to scenarios, the causes of causal events. Comparing a specific failure/error/condition with the recorded EWSs makes it possible to determine whether the failure/error/condition represents an EWS.

5.7.9 Results can be used to prioritize EWSs

If the list of identified EWSs is compared with prioritized process safety information for a given socio-technical system, it may serve as a simplified risk analysis of the relevant safety impact or accident near-miss potential of EWSs in other circumstances. As a result, the list of EWSs may be divided into two categories: (a) EWSs, which are important from risk perspective, and thus worth responding; and (b) EWSs not important and not worth responding.

The stored results of risk analysis (critical scenarios) should be characterised in terms of fulfilment of safety functions. Relevant systems, components and/or failure modes or classes of EWS should be readymade within a database serving as a tool for simpler realisation of the task.

Prioritisation of EWSs can be done analogously as determination of quantitative importances of components according to Vesely et al. (1981). Let us assume that prospective analysis of the process/system results in the list of incident scenarios s_i, where $i = 1$ to N. Let us assume that point estimates of frequency f_i and of damage x_i are determined for each scenario. Point estimates of scenario frequencies are determined with the use of point estimates of frequencies of causes of individual events in scenarios. Point estimate of risk of the process/system R can be determined as a sum of all products $f_i \times x_i$ for i =1, ..., N. Let us determine a modified point estimation of risk R(EWS) as a sum of products

f_i(EWS) $\times x_i$ for i =1, ..., N, where frequencies f_i(EWS) are determined with the use of point estimate of frequency of EWS = 0/year. Priority of cause EWS is p(EWS) = R - R(EWS). The higher the priority, the greater the risk reduction can be achieved by suppressing the occurrence of the EWS.

5.7.10 Resulting EWSs may have many of required attributes

Kate may require an ambitious investigation purpose of providing lessons learned. For example, it was suggested in Part 5.5.9 that she may intend to investigate accidents in all the kitchens of all Williams living throughout the UK, to improve their safety during cooking.

In such a case, she needs corresponding software tools, such as Risk Spectrum mentioned in Part 5.7.7, which includes both graphical editors to facilitate investigations and databases to facilitate documentation. With such a powerful tool, it can be expected that resulting EWSs will have many of attributes required in Parts 5.5.9 and 5.5.10.

At this moment, it can be supposed that Investigation attributes 1-3 are easily present and attributes 4-6 can be present. Similarly, Documentation attributes 1-7 can be present.

5.8 Conclusions

This presents scenarios as an extremely useful tool that may support foresight in safety. The concept of scenario is understood here as general as possible.

Both prospective scenarios, mainly derived from risk analyses, and retrospective scenarios, mainly derived from undesirable event investigations, are usable for foresight.

There are no restrictions on how the scenarios are presented. All methods are available, from the simplest pairs of initiating event - consequence to relatively complex bow-tie diagrams.

The chapter introduces the concept of casual events, representing a 'skeleton' of a scenario, no matter the type and form of the scenario. Casual events enable visibility of early warning signs, which is a condition for foresight in safety. The path to the determination of EWSs leads through the determination of causal events.

The examples presented in the chapter show:

(i) Scenarios as an investigation component of lessons learning help the determination of sets of EWSs that should be searched and tracked for during the analyses, and

(ii) Scenarios as a documentation component of lessons learning help the determination, whether a specific failure/error represents an EWS.

5.9   Acknowledgements

5.10   References

Accou, B. and Reniers, G. (2018). Analysing the depth of railway accident investigation reports on overspeeding incidents, using an innovation method called "SAFRAN". In *55th ESReDA Seminar (Session 5),* Bucharest, Romania, October 2018.

Afonso, A., Garnett, K., Noteborn., H., and Maragkondakis, P. A. (2017). Emerging Risks in Food and Feed, the Importance of Foresight. In *53rd ESReDA Seminar*, Ispra, Italy, November 2017.

Aven, T. (2008). *Risk Analysis*. John Wiley & Sons, Chichester, England. ISBN: 978-0-470-51736-9, 194 pages.

Benner, L. (1975). Accident Investigations: Multilinear Events Sequencing Methods (modified by Epilogue added), accessible from ludwigBenner.org, accessed on March 27th, 2018, original article *Journal of Safety Research*, June 1975, Vol. 7, No. 2.

Benner, L. and Carey, W. D. (2009). Lessons Learning System Attributes: An Analysis. In *36th ESReDA Seminar*, Coimbra, Portugal, June 2009.

Blanco, R. F. (2014) Understanding Hazards, Consequences, LOPA, SILs, PFD, and RRFs as Related to Risk and Hazard Assessment. *Process Safety Progress* 33, 208-216.

CCPS (2000). *Guidelines for Chemical Process Quantitative Risk Analysis, 2nd edition.* American Institute of Chemical Engineers, New York, USA. ISBN: 0-8169-0720-X, 754 pages.

CCPS (2001). *Layer of Protection Analysis – Simplified Process Risk Assessment.* American Institute of Chemical Engineers, New York, USA. ISBN: 0-8169-0811-7, 270 pages.

CCPS (2003). *Guidelines for Investigating Chemical Process Incidents, Second Edition.* American Institute of Chemical Engineers, New York, USA. ISBN: 0-8169-0897-4, 452 pages.

CCPS (2007). *Guidelines for Risk Based Process Safety.* John Wiley & Sons, Hoboken, New Jersey, USA. ISBN: 978-0-470-16569-0, 698 pages.

CCPS (2008). AIChE Center for Chemical Process Safety. *Guidelines for Hazard Evaluation Procedures, Third Edition.* John Wiley & Sons, Hoboken, New Jersey, USA. ISBN: 978-0-471-97815-2, 542 pages.

CCPS (2012). AIChE Center for Chemical Process Safety. *Recognizing Catastrophic Incident Warning Signs in the Process Industries.* John Wiley & Sons, Hoboken, New Jersey, USA. ISBN: 978-0-470-76774-0, 227 pages.

Crowl, D.A. and Louvar, J.F. (2011). *Chemical Process Safety, Fundamentals with Applications, 3rd edition.* Pearson Education, Boston, MA, USA. ISBN: 978-0-13-278283-8, 705 pages.

Defence Science Board Task Force (2003), The Role and Status of DoD Red Teaming Activities, Defence Science Board Task Force, Washington, DC.

ESReDA (2009). Guidelines for safety investigation of accidents. Technical report, ESReDA.   Available   from http://www.esreda.org/Portals/31/ESReDA_GLSIA_Final_June_2009_For_Download.pdf

Ferjencik, M. and Dechy, N. (2016). Three accidents in European dynamite production plants: An attempt to improve the external lessons learning. *Journal of Loss Prevention in the Process Industries* 44, 12-23.

Ferjencik, M. (2014) IPICA_Lite—Improvements to root cause analysis. *Reliability Engineering and System Safety* 131, 1–13.

Ferjencik, M. (2017). Roles of Incident Scenarios in Foresight. In *53rd ESReDA Seminar*, Ispra, Italy, November 2017.

Hatch, D., McCulloch, P. and Travers, I. (2019) Enhancing PHAs: The Power of Bowties. *Chemical Engineering Progress*, February 2019, 20-26.

Hollnagel E. (2014). *Safety-I and Safety-II. The Past and future of Safety Management.* Ashgate Publishing Ltd., Surrey, England. ISBN: 978-1-4724-2305-4, 187 pages.

Johnson, C.W. (2003) *Failure in safety-critical systems: a handbook of incident and accident reporting.* Glasgow University Press, Glasgow, UK. ISBN: 0-85261-784-4.

Johnson, W.G. (1973) *The Management Oversight & Risk Tree – MORT*. SAN 821-2, U.S. Atomic Energy Commission, Division of Operational Safety. Available from www.nri.eu.com.

Kaplan, S. and Garrick, B.J. (1981). On the Quantitative Definition of Risk. *Risk Analysis* 1, 11-27.

Kaplan, S., Haimes, Y. Y. and Garrick, B. J. (2001) Fitting Hierarchical Holographic Modeling into the Theory of Scenario Structuring and a Resulting Refinement to the Quantitative Definition of Risk. *Risk Analysis* 21, 807-819.

Kaplan, S., Visnepolschi, S., Zlotin, B., Zusman, A. (1999). New Tools for Failure and Risk Analysis. Anticipatory Failure Determination (AFD) and the Theory of Scenario Structuring. Detroit: Ideation International Inc.

Leveson, N. (2004) A new accident model for engineering safer systems. *Safety Science* 42, 237–70.

Marshall, V. and Ruhemann, S. (2001) *Fundamentals of Process Safety*. Institution of Chemical Engineers, Rugby, UK. ISBN: 0-85295-431-X, 298 pages.

Masys, A. J. (2012). Black swans to grey swans: revealing the uncertainty. Disaster Prevention and Management: An International Journal, 21(3), 320–335. doi: 10.1108/09653561211234507

Nicolescu, M. (2018) A freight train derailment analyses using Accident Investigation Board Norway method and Safety Management wheel tool. In *55th ESReDA Seminar (Session 6)*, Bucharest, Romania, October 2018.

Stoop, J. and Benner, L. (2015). What do STAMP-based analysts expect from safety investigations? *Procedia Engineering* 128 (2015) 93-102.

Strucic, M. (2017). Use of Event and Causal Factor Short Cart Reports to Assess and Simplify Accident Reports. In *53rd ESReDA Seminar*, Ispra, Italy, November 2017.

Verschueren, F. (2018) Learning from organisational dysfunctionalities. In *55th ESReDA Seminar (Session 2)*, Bucharest, Romania, October 2018.

Vesely, W.E., Goldberg, F.F., Roberts, N.H., and Haasl, D.F. (1981) *Fault Tree Handbook* NUREG-0492. U. S. Nuclear Regulatory Commission, Washington DC, USA.

Wincek, J. C. (2011). Basis of Safety: A Concise Communication Method for Critical Process Safety Information. *Process Safety Progress* 30, 315-318.

Zio, E. (2007). *An Introduction to the Basics of Reliability and Risk Analysis.* World Scientific Publishing, Singapore. ISBN: 978-981-270-639-3, 222 pages.

6 Visibility of Early Warning Signs

Miodrag Stručić, Joint Research Centre, European Commission, The Netherlands

6.1 Executive summary

Visualisation of early warning signs is critical for clear understanding of existing weaknesses' roots and to define effective actions to prevent their escalation. The process of visualisation is described by use of fictional devices and features to emphasize the importance of different phases in the process.

In the first phase of the visualisation process, "detectors" are used for detection of warning signs - "signals". Their "output" is further "modulated" through reporting systems to provide a robust repository of important facts, but also attributes about issues, as well as to increase awareness in our socio-technological Organisation. These signals are then "amplified" to an appropriate level of visibility, by which we are able to fully understand vulnerabilities of the Organisation. These insights, combined with the knowledge and experience of operations and design, provide necessary ingredients for smart decisions in fighting potential threats.

6.2 Introduction

Warning Signs, like the signs defined by road traffic regulations, provide information about immediate, delayed or potential danger. In every case they should be treated as Early Warning Signs (EWS), and every warning sign, if processed in the right way, will eliminate the possibility of its escalation. This is true even if an immediate event occurs – registered and processed through an efficient Operational Experience Feedback system can help affected stakeholders and those responsible for similar Organisations to prevent similar events.

There are many Early Warning Signs (ESW) that can be detected and adequately treated before they transform into a bigger problem. Some EWS however are too weak to be recognised as a threat to safety, but are still detectable. Once detected, those signs-signals should be "modulated" and "amplified" to an appropriate visibility level that can be justified as a treat by stakeholders. When signals are visualized and presented as a real threat, they can be efficiently treated to prevent further development into new incident or accident.

The main objective of this chapter is to define or give directions to define how to reveal or visualise EWS in socio-technological environment. Not only in-house, but also external and publicly available data, should be used to help in determination of EWS. Better understanding i.e. visualisation of this data can help Organisations that didn't experience the same or similar incident to foresee safety hazards, assess the risk of its occurrence, and initiate adequate measures.

It is not easy to quantify the contribution to safety processing of external experience, but due to similarities of hazardous industries in their safety concepts, the basic reasons for reported deviations should be examined with the same respect and effort as for their own.

6.3 Detection

Signals of decreased safety – Warning Signs in Hazardous Organisation - should be recognised and confirmed by any individual regardless of his/her organisational level (IAEA, 2018). Reporting these signals is of the highest importance for Organisation, not only because of immediate prevention of an undesirable event, but also to enable a system of multi-dimensional assessment whose main purpose is to improve the safety level of the Organisation i.e. reduce the probability of severe accidents. Detecting these signals is easier if they are obvious, but signals identified by "out of routine" practice could be very difficult to detect.

However, information about deviation, no matter how it is detected (Stručić., 2016), should activate the system which is able to "visualize" any warning signal. Good definition of EWS could improve their detection and treatment. For this purpose, it is equally important to define main detection modes and categorise them. To better understand detection modes it is necessary to distinguish different types of "detectors".

6.3.1 Built-in / Surveillance Detector

Most EWS are discovered by design. These include all forms of alarms and annunciators, as well as indicators and records required to monitor processes during all phases of operation. Also check/verification-lists predefined in written

form (audits, QC, operating procedures) present an efficient tool to detect warning signals.

Good examples of built-in or Surveillance detectors are annunciators' panels in Nuclear Power Plant (NPP) Main Control Room (MCR). Even in the case that some "unimportant" signal is created in one of the less important locally controlled systems, the alarm from local panel will send a signal to the MCR and the Main Control Board alarm for local panel will go off. MCR operators usually ask the local operator for signal confirmation and react to this alarm in accordance with a well defined procedure. Process failures that often occur are treated instantly but they are also recorded in Log Books, Process Information System or reported through Corrective Action Program - CAP (IAEA, 2005).

Furthermore, some non-wired detection processes, e.g. auditing, are strictly following a prepared plan and procedure so that all non-conformances are recorded and reported through a final report. The same is true of Quality Control inspections which are performed in accordance with strictly defined check-list items. In the case of a verification procedures (e.g. Line-up checks) performed by an operator, the response to a discovered deviation is similar to the response to an automatically detected deviation.

Detection of these deviations usually brings immediate solutions and they are not followed by deeper investigations. These approaches produce large numbers of recorded items, but in our moment of Organisational and technological progress, it is still acceptable to leave most of them for later deeper assessments. In other words, there is still no available automated system[18] to connect all detected items, compare them, use available "experience" and immediately process them in the best way for optimized and safe further operation. Although technological design experts are striving to build a system that can predict all deficiencies and warn operators or auto-correct them, would it be ever possible? Certainly it would never be possible using only built-in/surveillance detection as the abandoned IAEA ASSET program showed.[19]

6.3.2   Advanced systematic approach Detector

There are also Early Warning Signs which could be detected by assessment tools (IAEA, 1997) - Focused Self-Assessments, Safety Assessments (Risk Analyses), Peer Reviews, Advanced Surveillance programs[20], Performance Indicators programs, as well as Preventive and Predictive Maintenance. These tools enable discovery of the EWS in a systematic, predefined way, many times by teams with expertise and experience in specific areas. Although some of these approaches overlap with Built-in approaches, they enable discoveries of deviations thanks to performers' incisive and experienced approach, i.e. they give freedom to performers to examine a wider range of assessed items.

The advantage of the team approach lies in the fact that experts, who are usually not part of the specific assessed process, reveal deviations using their experience, knowledge, skills and techniques, and many times seeing things with "different eyes". They can reveal hidden deviations not visible to staff doing routine work in their field. It could be e.g. human or organisational issue easily visible from team member chosen from management or non-related domain experts.

Advanced surveillance programs, i.e. non-built-in detectors, give freedom to performer(s) to find unexpected inconsistencies or discover that faulty equipment or systems point to the deeper reason for a given anomaly. A typical example would be detecting a fault induced by ageing of a component[21] which is, almost by default, applicable for all other items in the corresponding group or system. Thus, Operating Organisation will not just fix the problem, but also launch actions to assess the system and find optimal solutions.

EWS discovered by one of these approaches is not always pointing to the main problem, therefore the assessments' results should be analysed for deeper causes. E.g. Performance Indicator detects negative trend of "Number of Overdue Work Requests", but additional effort is needed to find deeper reasons for this trend.

---

[18] *Or Artificial Intelligence system*
[19]*E.g. abandoned IAEA ASSET program for prevention of incidents and accidents where the main input was surveillance data which was not sufficient to foresee enough incidents to prove effective.*

[20] *Various surveillance techniques and tools that are not originally installed in the Organisation. Those include e.g. advanced electronic diagnostic equipment components, new process data analysis programs etc.,*
[21] *Note that Ageing monitoring scope can overlap with Built-in detectors*

As mentioned, these approaches can overlap with Built-in ones, but their results are less predictable. This also means that trends of number of detected issues can vary significantly over time. In the case of a decreasing trend, an Organisation may think that further use of one of these tools is unnecessary, but in fact it is just the opposite – the declining trend should stimulate the Organisation to improve their detection program (explained more in Trends in detection of deviations subchapter – 6.3.5).

6.3.3 Analysis of Operational Experience and Technological-Organisational-Human performance Detector

A hidden EWS could be detected by revealing latent weaknesses using Causal analysis of internal and other industry's events. Good examples can be found conducting Root Cause Analyses which compile event investigation results. Thorough investigation followed by systematic definition of causes reveals warning signals which contributed to the evolution of an analysed event. E.g.
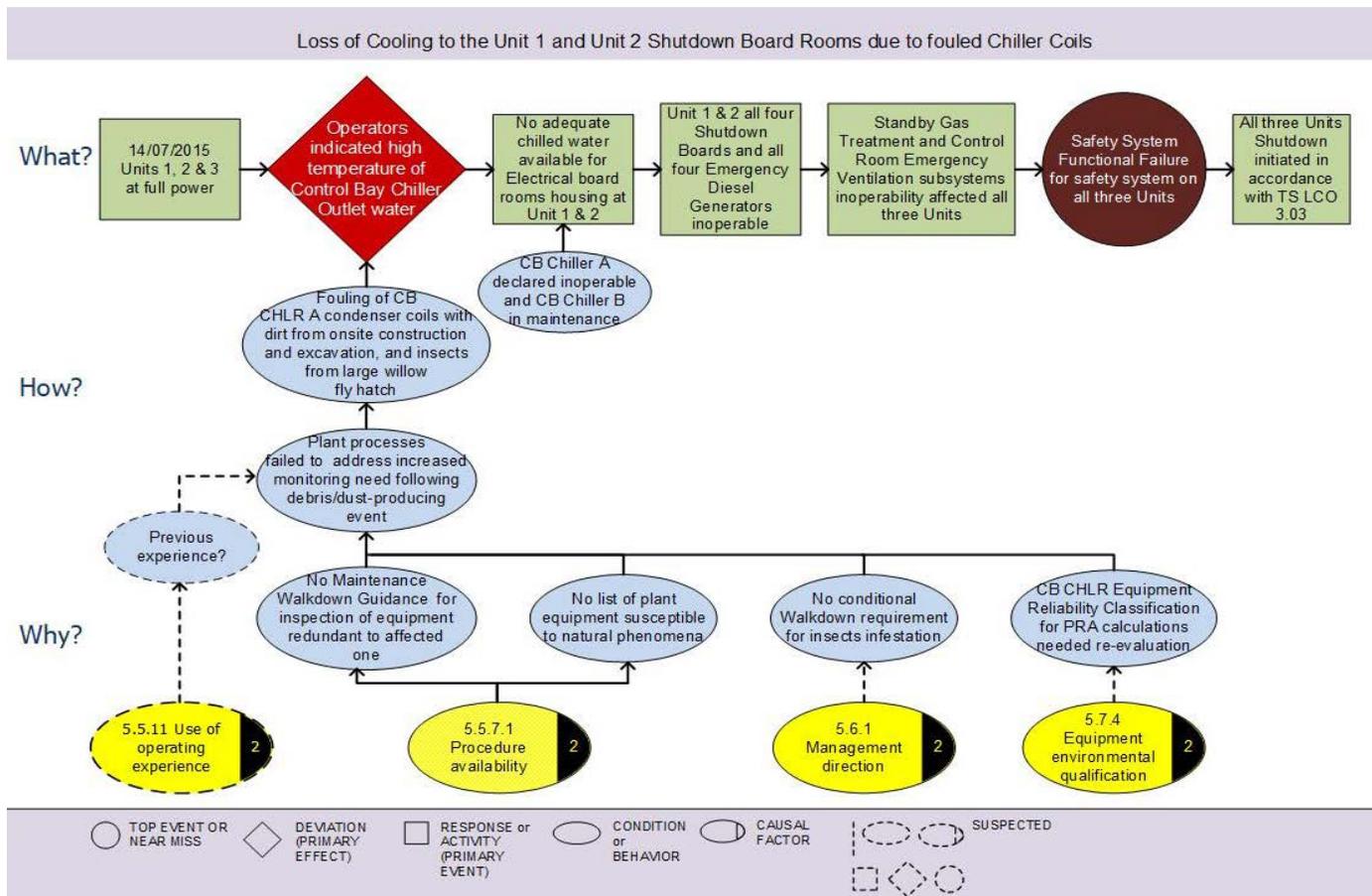
Figure 1. Event and Causal Factor Short Chart – Loss of Cooling event

factors connected with Personal Work Practice could reveal that Lack of Peer Check in Critical Task is one EWS.

There are many approaches, techniques and tools used in analysing events (Ziedelis and Noel, 2011) or Near misses (subchapter 6.4.1), but one technique often used for events in NPPs is Event and Causal Factor Charting (E&CFC). An important advantage of this charting technique is that visual presentation of analysis transparently shows the causes of an event that can then be easily used to define adequate action plans. This advantage could be used also for the purpose of presenting events after the analysis is performed. In two examples (fig 1. - 4.) of not-so-significant events in NPPs some not very transparent issues are highlighted (Strucic, 2017).

Example in Fig 1. presents an event in one NPP that had no adverse safety consequences but its Root Cause Analysis revealed some potential causes which the affected plant then eliminated by adequate actions to prevent a more serious event. What is not explained in the analysed report, is how one of the revealed problems was treated in the past. In this concrete case, non-use of operating experience or ineffective corrective actions are typical possible causes of repeated problems. It is visible in the highlighted part in Fig. 2.

The second example in fig 3. shows hidden Human and Organisational Factor (HOF) problems not revealed in the published report and most likely not in the original one either. While the Organisation could be satisfied that the revealed cause points to deficiency in the specific software, behind this cause there are possible warning signals of deficiencies in HOF domain which could create more serious problems. It is visible in the highlighted part in Fig. 4.

These examples show how to reveal additional or hidden warning signals which could be ignored or missed in routine Condition/Deviation report processing. This approach can be used by any similar Organisation interested in safety improvement of their own installations. The Operating Organisation where the event occurred can benefit the most, but all other similar Organisations can also find it useful.
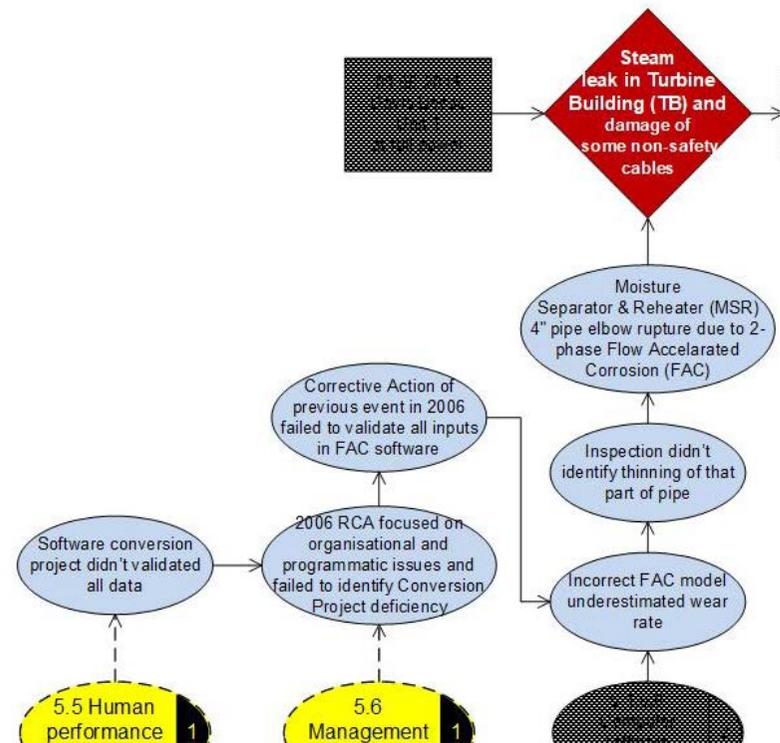
Figure 2. Event and Causal Factor Short Chart – Loss of Cooling event highlight
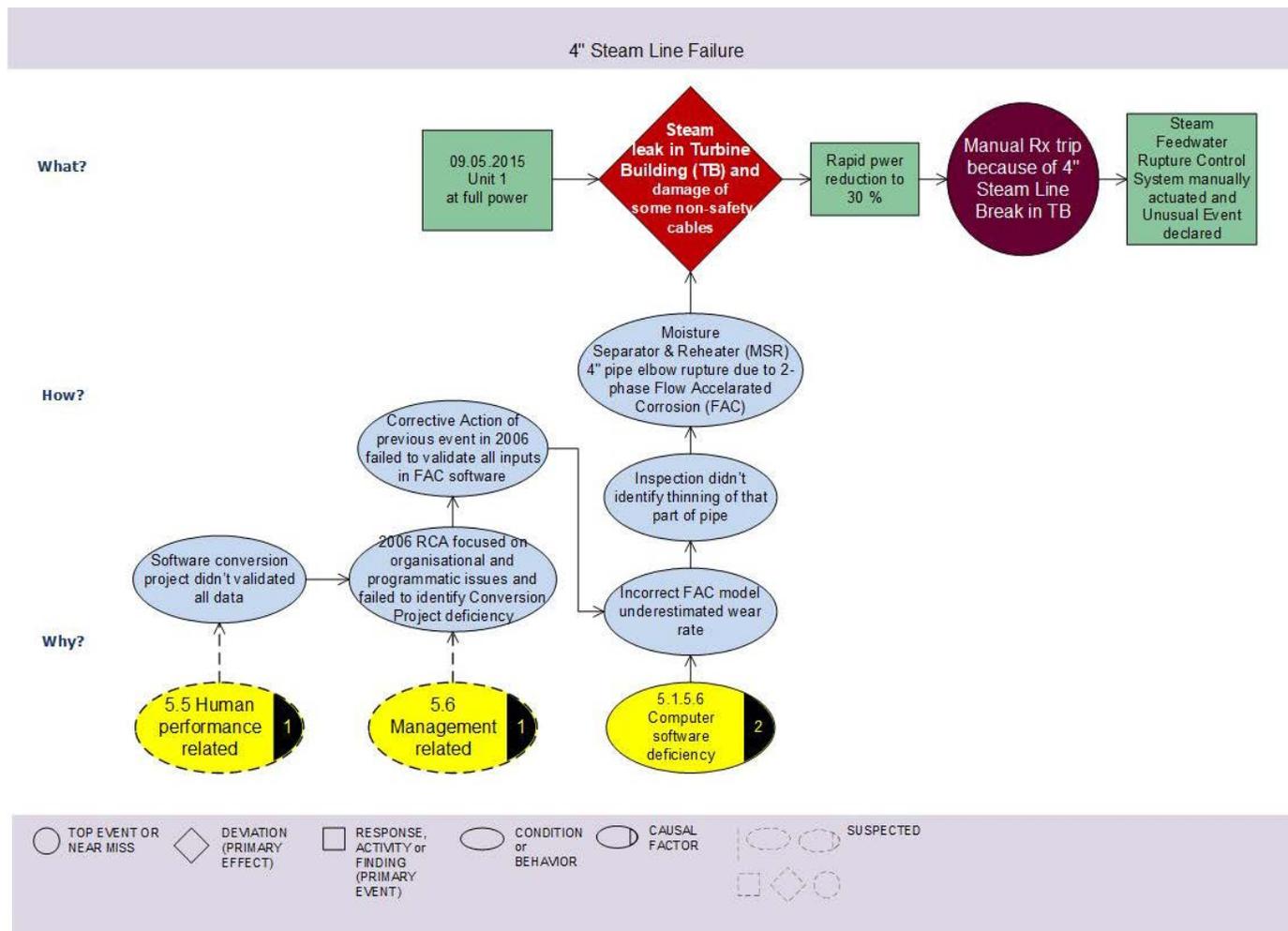
Figure 3. Event and Causal Factor Short Chart – Steam Line Failure event

It should be mentioned that short E&CFC diagrams (fig. 1-4) can quickly give to external operating Organisation the essential information needed for fast qualitative risk assessment of a potential similar problem (Strucic, 2017). Evolution of event is presented in the Primary event line – in the upper horizontal row. An experienced and knowledgeable stakeholder can find out if the event is applicable to his/her Organisation and what is the potential hazard. The mechanism of each defined deviation (red rhomb) genesis is explained in below connected boxes by the most appropriate answers to how and why questions.
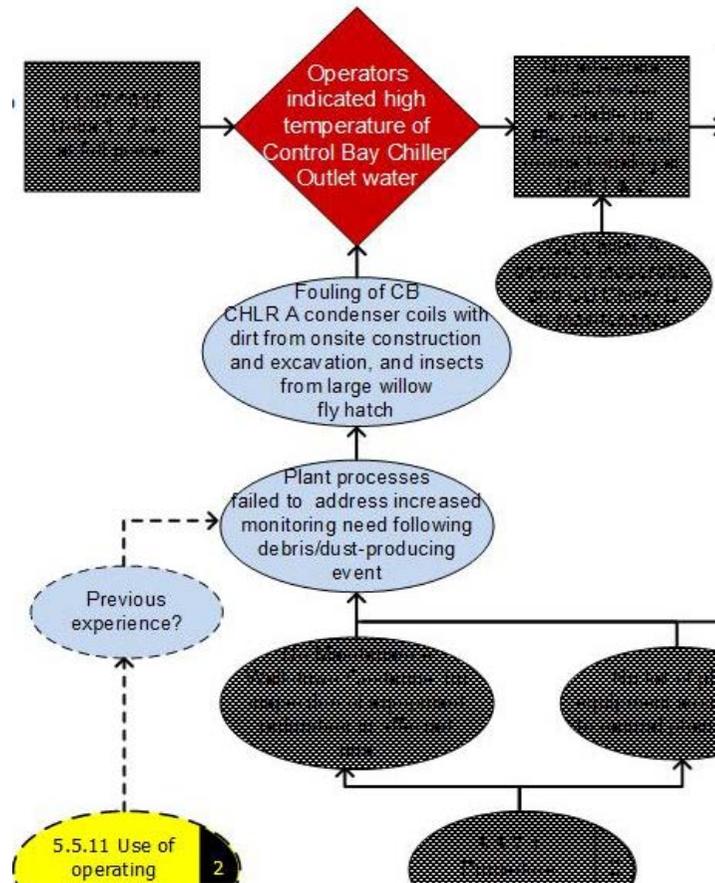
Figure 4. Event and Causal Factor Short Chart – Steam Line Failure event

Understanding causality of deviations gives the stakeholder an idea of vulnerability of the operating Organisation under his/her responsibility. Knowing potential hazard and how vulnerable the Organisation is to similar or same deviation provides stakeholder the rough estimation of their own risk. Therefore, stakeholders can easily decide if action for changes or further self-assessment in the Organisation is needed.

6.3.4   Detection by "chance"

Many times, the problems, especially of human or Organisational nature, are discovered during activities different than ones described in previous subsections. Those could be just a "side product" of activity like meetings, trainings, work-related trips or even some activities outside the regular working hours or Organisation premises.

Also "Equipment related" deficiencies can be detected by "chance". As a concrete example, when one NPP unit experienced an unplanned Reactor Shutdown and activated some systems important to safety (IAEA, 2015), two major systems faults were discovered. Fortunately, the safety system activation wasn't required and these deficiencies didn't result in unsafe conditions. This case clearly showed that surveillances, tests or any other defined assessment approach could not detect these problems. A good lesson from this event teaches operating Organisations to thoroughly examine all their major unplanned transients and find deficiencies that couldn't be found by regular established processes.

Since reporting requirements are not able to address items detected by "chance", they may not be processed and easily can be forgotten. However, deviations found "by chance" could be of high importance since other systematic and well-defined processes didn't detect them and probably won't. Therefore, Organisations should increase their awareness of all minor, hidden or accidentally found warning signals and encourage employees to report them too.

6.3.5   Trends in detection of deviations

In any discussion of detectors, it is important to mention the possibility of trending detection modes. The Topical study of Nuclear Power Plants design deficiency (Stručić, 2016) reviewed the worldwide operating experience from NPP events where design deficiencies are addressed. One of the outcomes, the trend graph, has been created during this study (Figure 5.).

The Detection Mode trend graph, based on information provided in IAEA International Reporting System (IRS) (IAEA, March 2020, https://nucleus.iaea.org), shows that the number of events with actual consequences is increasing over time, which suggests both that other detection approaches should be employed and existing detection should be improved. A sudden drop in the number of events that are detected by surveillance and reviews may be an indication of obsolete or inadequate deficiency detection methods.
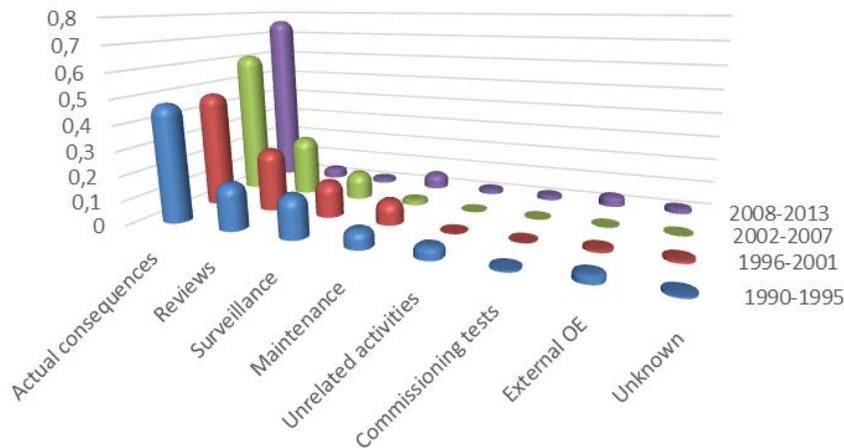
Fig 5. Trend graph of detection modes of design deficiency in IRS

The graph of detection mode trends could indicate if the main tools for detection of deviations are becoming inefficient. If there is no appropriate recording and coding/categorisation of events in operating Organisations, it could be difficult to create trend graph, otherwise this could be used as one Performance Indicator as well.

6.4 Reporting system

The simple concept of Problem-Screening-Analysing is a natural approach for handling any problem. Hence, it can be very practical to use Corrective Action Program (IAEA, 2005) concept for any kind of problem (IAEA, 2012).

It worthwhile to note the definition of USA Nuclear Regulatory Commission - *CAP is the system by which a utility finds and fixes problems at the nuclear plant. It includes a process for evaluating the safety significance of the problems, setting priorities in correcting the problems, and tracking them until they have been corrected* (USA Nuclear Regulatory Commission, March 2020, https://www.nrc.gov/reading-rm/basic-ref/glossary/corrective-action-program.html).

6.4.1 Deviation Report

Whatever the source or detection mode of EWS is, the first step in any CAP process is to record the detected deficiency (acquisition phase in fig. 6) through a Condition/Deviation Report Form. Practically, the whole process of visualisation is useless without input – acquired anomalies, errors, mistakes, discrepancies, wrongdoings, or simply, any deviation or problem. Operating Organisations should promote a Reporting Culture (Reason, 1998) and encourage all employees to report all noticed deviations and potential problems.

Furthermore, there are many definitions of Near Misses but are they really necessary? "Things are never so bad they can't be made worse", a favourite Humphrey Bogart quote (Brainyquote.com, March 2020, https://www.brainyquote.com/quotes/humphrey_bogart_108860) is relevant – it reminds us that all events, major or minor, need only some small effort of evil or bad luck to become worse. By this interpretation all events can be considered as Near Misses. And, needless to say, any event labelled Near Miss can easily propagate under some realistic circumstances to become a disaster. Thus, all undesirable events, regardless of the consequences, should be reported and analysed.

Organisations should be aware and avoid traps of non-reporting. The first notice or record of any deficiency or irregularity is the crucial step which is easily missed, e.g. because the person who discovers it cannot foresee all possible consequences, interactions or users of reported/recorded information, and doesn't consider it important to report through the official reporting system.

The reporting system should be easily accessible and user friendly for all employees, regardless of their position or duty. The reporting form should include questions about:

- What happened?
- Time, location and involved subject(s); and
- Author – This is important for evaluators to know whom to ask for possible additional question. But it must be emphasised that anonymous reporting should be encouraged as well and always be an option;

The author should be encouraged to include all information necessary for further processing (equipment number, coordinates, time of discovery, procedure used, circumstances, misbehaviour, outcome, proposed action etc.). Reporting

application could ease this input with predefined lists of equipment or persons, as well as other relevant categories.

Once the record is created, screening of the deficiency should be performed as soon as possible.

6.4.2   Screening and categorisation

When the first key step is done, i.e. the deviation is registered in the Reporting System, the responsible personnel should take care of further processing of the report as soon as feasible. Besides technological classification (equipment numbers, mode of operation, failure mechanism…), additional categories should be entered: If there is an urgency to react; what corrective measures are taken; what could be possible consequences; what is a safety significance level; which department and person is responsible; dissemination list etc. - information necessary to effectively fix the problem, transfer lessons and experience, and provide basis for possible deeper evaluation.

This phase is partially automated, and may be further automated in future, but human involvement is indispensable. Having a good knowledge of design, operation and Organisation increases the accuracy of categorisation and enables more efficient further processing. A broad understanding of possible data used by different stakeholders or by other programs in the Organisation is one of the benefits of having experienced and knowledgeable screeners.

A good example of an efficient screening process highlights two levels of screening phase. All reported deviations from previous day are discussed at morning multidisciplinary Screening Committee meeting where decisions are made how to treat the reported issues and who will be responsible for further actions. It is usually organised immediately after Operations' meeting where often first-hand information about safety and technical issues are given. In parallel, the Independent Safety Engineering Group is taking care of coding, administration and coordination of all CAP items. This experienced staff is further involved in analyses and trending of events, which is a highly important part of whole CAP process (Bach, 2007).

6.5   Visualisation

As explained in the previous sub-chapter, the destiny of each item in the CAP depends, in the first place, on screening. Typically, after screening of the Deviation/Condition Report, Apparent Cause Analysis is performed, tasks are sent to responsible persons and information is disseminated. Very few of events are immediately investigated and more deeply analysed by Root Cause Analysis. Whatever method or tool is used, it is important that the deviation's causes are registered and can be further processed.
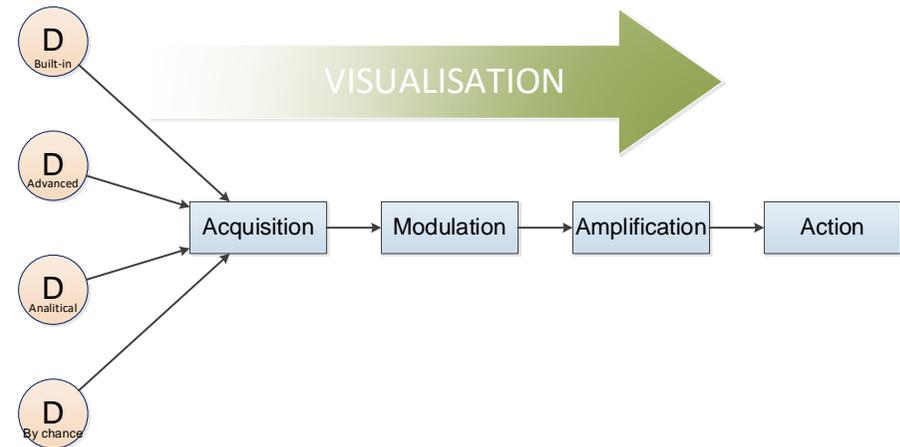
Fig 6. Simplified diagram of whole process from signal detection to action

All deviations present Warning Signals of unsafe or unreliable operation. Obviously, causes are not all deeply investigated for all reported deviations and, accordingly, there are also many Hidden Warning Signs which look "invisible", but still, all of them are producing Weak Signals that can be captured, modulated[22], amplified and treated (Fig 6). Fortunately, all conditions are kept in a Reporting System database and could be assessed any time and analysed by some additional assessment processes.

Typically, Common Cause and Trend analyses (IAEA, 2012) use CAP database for later identification and assessment of latent deviations. Furthermore, there are other processes like Focused Self Assessments (Stručić et al., 2006) or Peer Reviews

---

[22] *Note that "Modulation" of signal in this metaphorical model is mainly used in sense of transforming EWS to the factual "signals" – Cause or Probable Cause's Effect – input to "Amplifier" module (6.5.1).*

that always look in CAP for analysis inputs. Correctly performed, all these assessment tools would reveal hidden signals and find solutions for each. Together with use of Root Cause Analysis tools, they should bring to light all Warning Signals and their causes that can be efficiently treated to prevent recurrence of events or occurrence of similar events, and their escalation into accident.

Therefore, Weak or Hidden Warning Signals can be and should be visualized to the level where effective corrective actions can be easily defined and implemented.

6.5.1   Amplification
Corrective Action Program (IAEA, 2005) or any other similar approach that requires further processing of internally reported deviations, events and Near Misses is intended to correct their causes. Depending on the Organisation's policy and procedures, screening process usually results in Direct Action, request for Apparent or Root Cause analysis, or just report closure without corrective actions. Whatever is the outcome of screening, it is necessary to optimise resources and ensure that all information is preserved for later use. Fear of "feeding the beast" with a large amount of reported items should be alleviated through policy, to support a strong reporting culture. In the case that the reported item is not immediately fully processed, i.e. the report is closed, it can be managed in another self-assessment process such as Performance Indicators program or Common Cause analysis to assess the causes of reported deviations.

In any case, each revealed "Cause", which produced, contributed or might produce an undesirable Effect, is a Warning Signal too. The same "Cause" could be also the Weak or Hidden Warning Signal of deeper problem inside the Organisation. Therefore, it should be amplified to become widely visible and manageable.

Figure 7 presents the process of amplification of this deficiency signal, i.e. cause. Each signal (cause) should be tested for necessity of deeper investigation. That is usually true for direct causes, which if corrected will not necessarily prevent reoccurrence of the event or creation of a bigger accident (Ziedelis and Noel, 2011). Thus, it would be necessary to investigate it and find deeper cause, i.e. to find why this cause existed. This cause, in this metaphorical model, is processed as effect of deeper cause. Note that C/E (Cause/Effect) converter is used for this transformation.

Hence, a "signal" is "amplified" through the "Amplifier" module, i.e. investigated one level deeper to define the immediate cause of this effect. The new cause is

tested if it is deep enough and manageable by Organisation (Roed–Larsen et al. 2005). If not, it should be investigated further. After some iteration, Deepest Manageable Cause is defined as final output of the Amplifier Module. This process reveals "invisible and unknown" facts, which are basically unrecorded deviations of the Organisation. Note that this iteration process can go in many directions: from finding specific cause that could be eliminated with surgical precision, to discovering generic problems that need a more comprehensive action plan.
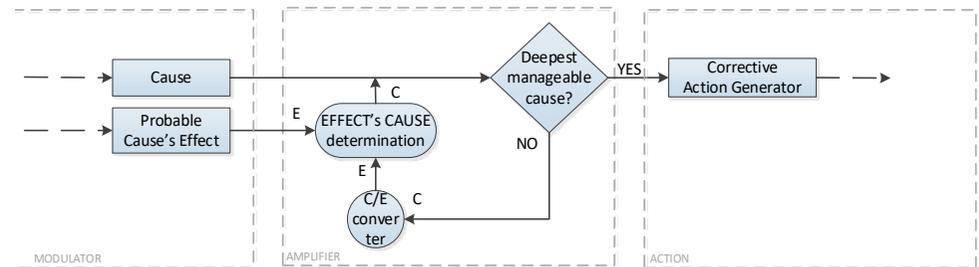
Figure 7, Amplification of Weak or Hidden Warning Signals

In addition to different types of "Causes", such as Root Cause, Direct Cause, Apparent Cause or Contributing Cause, it is important to recognise the Probable Cause too (fig 2 and 4). Many times, guided by Cause Analysis procedure, and requested by authority or required by legal requirements, some Facts are not required to be deeper examined although they are indicating existence of other important causes. Although not investigated yet, it can be assumed by our best judgement why these Facts exist. Since these causes are not proved, we can call them Probable Causes (Stručić, 2017). In figure 7, a Fact generated by a Possible Cause is presented as an Effect. Needless to say this Effect is a Warning Signal too. It is also important to note that the module, which processes other causes, processes Probable Cause too, but since it is just assumed, one step back is needed to extract its Effect and process it through an "Amplifier".

6.5.2   Elements of Amplifier
The purpose of good event analysis is to prevent recurrence of event i.e. to find the proper cause and enable its elimination. Therefore, the question "Deepest Manageable Cause?" is used in subchapter 6.5.1 to emphasise the main characteristic of the hunted cause - it examines the cause and compares the implied action based on that cause to the ability of the Organisation to efficiently

implement that specific action. If this action is too demanding: e.g. exceeding the Organisation's resources; cannot be arranged due to external subject's unavailability or just missing external approval, then the corresponding cause should be further examined. Furthermore, the same should be done to "Shallow" causes which produce work requests usually "just to fix the problem" (Ziedelis 2014).

To examine or re-examine cause, the other fictional element "Cause/Effect converter" (C/E) is used to transform cause into effect. Since every cause is an effect of another deeper cause, investigating a "new" effect should result in a new finding which represents the deeper cause of an examined deviation. Deeper cause emphasizes the nature of the weaker cause and should easily prescribe appropriate action. If not, the process of "amplification" has to be repeated.

For Effect's Cause determination, it is essential to determine why the effect occurred which is done by use of different RCA tools (Ziedelis and Noel 2011). To simplify the approach, 5-Why's RCA tool could be used as good example. Thus, Effect's Cause element provides answer to question "what is the cause of this effect?" i.e. "Why does this effect exist?". In this illustration, the Amplifications Module should be used five times to get the positive passage through "Deepest Manageable Cause?" element.

Some critics emphasize the weakness of 5-Why's tool mainly because one "Why's" wrong answer could mislead investigator. To avoid this trap, one can use the Five-by-Five tool principle (Bill-Willson Net, March 2019, www.bill-willson.net/b73) in which five questions are defined to help in defining the right answer to each Why:

- What is the proof that this cause exists?
- What is the proof that this cause led to the stated effect?
- What is the proof that this cause actually contributed to the analysed problem?
- Is anything else needed for the stated effect to occur?
- Can anything else lead to the stated effect?

The other trap is that one effect could have more causes. Thus, they should be considered too. This would add parallel amplification loops and amplify additional hidden deficiency signals of Organisation.

6.5.3 Actions

In plain language the results of Condition/Deviation Report assessments are well defined causes of registered deviations and adverse conditions. They present vulnerabilities of the assessed Organisation and should be eliminated.

Regardless of deepness of retrieved causes, they should be well defined. This means that they should provide important information necessary to define effective corrective actions. If they are properly defined, definition of actions should be unambiguous.

E.g. production process in one NPP was stopped for some days because of failure of one safety control electronic circuit board. Further investigation found that one electrolytic capacitor failed because of aging. Extended/deeper investigation revealed that this model of circuit board is used in several other applications and that the manufacturer discontinued production, so it was no longer possible to replace the electronic control circuit boards, but this was not the case for the capacitor. Thus, an action plan was made to replace the same model capacitors in all circuit boards.

Actions defined this way should be reasonable and achievable on time, i.e. manageable by the Organisation. Some operating Organisations set requirements for corrective actions that have to be respected in a more popular way. E.g. in Organisations that use SMART approach (Specific, Measurable, Achievable, Relevant, Timely), typically the Corrective Action Program process enables the launch of corrective action when confirmation of all these requirements are achieved. Since it based on a defined problem's cause, it is obvious that the cause has to be well defined.

Furthermore, the Organisation should be motivated to perform even deeper investigation, even though this can bring Organisation to situation where it is unable to perform corrective action because it could be beyond the power of the Organisation or because of lack of resources (Roed–Larsen et al., 2005). Nevertheless, knowing more about the background of the problem could be an asset in Foresight in Safety. E.g. after defining actions to replace all affected capacitors in the safety electronic boards (because of ageing and inability of manufacturer to re-produce the same type of electronic board), extended investigation might reveal that the manufacturer is experiencing a survival risk due to a superior competitor. This might be important information for the Organisation because of difficult troubleshooting of possible other equipment failures with

equipment produced by the same manufacturer. Thus, this information is not crucial for immediate action plan, but might be critically important to possible future decisions and negotiations outside the Organisation.

6.6   Conclusion

To truly understand the problems in our operating Organisation, they should be visualised to the level of plain clearness. Tools and processes explained in this chapter can help to understand those problems deeply enough and enable visualisation of their weak and faulted roots. Only in this state of understanding, are we able to obtain good qualitative insight of our vulnerabilities. With the knowledge and experience of design and operations we will then get all necessary elements for smart decisions in fighting discovered weaknesses.

This chapter tried to explain how registered anomalies in operating Organisation should be processed to reveal deeper causes i.e. to visualise warning signals and provide a clear basis for definition of a corrective action plan. The four phases of visualisation process consist of Detection, Acquisition, Modulation and Amplification of warning signals. Naturally, output of this process becomes an input for determination and implementation of corrective actions process – typically called Corrective Action Program.

The presented work mechanism of fictional electronic device illustrates main elements of an efficient EWS treatment process in an operating Organisation. This can be developed in more details by adding additional electronic elements e.g. "Noise filter" or "Screener discriminator" etc. But the intention of this chapter is, at the first place, to give operating Organisations interested in continuous safety improvement an idea about an alternative approach in fighting the problems.

In the context of Foresight in Safety, revealing weak and hidden signals of decreased safety enables stakeholders to foresee potential safety consequences and initiate timely actions. Thus, visualisation of hidden and weak signals has an important role in predicting possible incidents and accidents. Nevertheless, deeper investigation of causes than needed for efficient action plan definition, can bring additional information to decision makers and enable them to use this extra knowledge in future decisions and negotiations.

6.7   References

Bach B., NENE conference 2007, Slovenia Corrective Action Program at the Krško NPP - Trending and Analysis of Minor Events

Bill-Willson Net, March 2019, www.bill-willson.net/b73

Brainyquote.com, March 2020, https://www.brainyquote.com/quotes/humphrey_bogart_108860

ESReDA, 2015, A Case study analysis on dynamic learning from accidents — The ESReDA Cube, a method and metaphor for exploring a learning space for safety

International Atomic Energy Agency, 1997, Procedures for self-assessment of operational safety, IAEA-TECDOC-954

International Atomic Energy Agency, 2005, Effective corrective actions to enhance operational safety of nuclear installations, IAEA-TECDOC-1458

International Atomic Energy Agency, 2006, Application for Management System for Facilities and Activities, IAEA GS-G-3.1

International Atomic Energy Agency, 2012, Low Level Event and Near Miss Process for Nuclear Power Plants: Best Practices, SAFETY REPORT SERIES No. 73

International Atomic Energy Agency, 2013, Periodic Safety Review for Nuclear Power Plants for protecting people and the environment, IAEA No. SSG-25

International Atomic Energy Agency, 2014, Precursor analyses — The use of deterministic and PSA based methods in the event investigation process at nuclear power plants, IAEA-TECDOC-1417

International Atomic Energy Agency, 2015, Root Cause Analysis Following an Event at a Nuclear Installation: Reference Manual, IAEA TECDOC No. 1756

International Atomic Energy Agency, 2018, Operating Experience Feedback for Nuclear Installations, IAEA Safety Guide SSG-50,

International Atomic Energy Agency, March 2020, https://nucleus.iaea.org

International Atomic Energy Agency, November 2002, Self-assessment of Safety Culture in nuclear installations - Highlights and good practices, IAEA-TECDOC-1321

International Atomic Energy Agency, November 2006, Fundamental Safety Principles, IAEA Safety Standard No SF-1

Reason J., 1998, Achieving a safe culture: An International Journal of Work, Health & Organisations, Theory and practice, Pages 293-306

Roed–Larsen S., Stoop J.A., Funnemark E., Bolt E., 2005, Shaping public safety investigations of accidents in Europe, ISBN: 82-5150304-3

Strucic M., 53rd ESReDA Seminar, Ispra, Italy, November 2017, Use of Event and Causal Factor Short Cart Reports to Assess and Simplify Accident Reports

Stručić M., European Commission 2016, European Clearinghouse topical report: Events related to design deficiency, PUBSY No.: 105232

Strucic M., Kavsek D., Novak J., Dudas M., 6th International Conference on Nuclear Option in Countries with Small Grid, Dubrovnik 2006, Self-Assessment at Krško Nuclear Power Plant

USA Nuclear Regulatory Commission, March 2020, https://www.nrc.gov/reading-rm/basic-ref/glossary/corrective-action-program.html

Ziedelis S., European Commission 2014, Organizational and Management-Related Aspects in Nuclear Event Analysis, Topical Operational Experience Report

Ziedelis S., Noel M., 2011, Application and Selection of Nuclear Event Investigation Methods, Tools and Techniques, Final Technical Report

Ziedelis S., Noel M., Technical Report European Commission 2011, Comparative Analysis of Nuclear Event Investigation Methods, Tools and Techniques, EUR 24757

7 Utilizing the ESReDA Cube to detect early warning signs

John Stoop, Kindunos, The Netherlands,
Tuuli Tulonen, Tukes, Finland,
Ana Lisa Vetere Arellano, Joint Research Centre, Italy,
Milos Ferjencik, University of Pardubice, The Czech Republic,
Sever Paul, AGIFER, Romania.

7.1 Executive summary

The ESReDA Cube model was developed by the ESReDA Project Group on Dynamic Learning and published in 2015. In this chapter we propose that the Cube model may be utilized to identify early warning signs of the system functioning, and solutions to improve the safety performance of the system in which the failures occur.

The purpose of this chapter is not to have foresight results, but to show how the Cube can be utilized to identify foresight potential. Applying the Cube makes it possible to explore links between otherwise isolated solutions, causal aspects and contributing factors. Here we are interested in the context, operating conditions and system states in which such factors, actors and aspects can serve as early warning signs. In doing so, we can use the Cube as a proactive, analytic tool to identify early signals of system safety performance during design regarding assumptions, simplifications and modelling, including or excluding specific safety aspects. Moreover, we can identify suppressed safety signals that get lost in daily operations, traded-off against economy, environment, lead time and production stress.

By such a coherent scan of all system dimensions, in practice the mapping will give us insight by e.g. helping to identify early warning signs and safety management system components that need to be followed in a systematic way in order to gain foresight and avoid future accidents and other undesirable events. Yet, it is not only the goal, it is the road towards it: The Cube model should not be looked at as merely a mapping of all the elements involved. In addition, the mapping of the elements generates a process that aims to provoke discussion of aspects (incl. insight, and foresight potential) that might otherwise be ignored or get lost in the shuffle. The 'rote faden' of our chapter is to structure the search for order in the complexity chaos, discriminating between safety interventions in either the event or the system, while accumulating understanding of the coherence and dynamics of what is going on.

7.2 Introduction

The ESReDA Cube is a model developed by the ESReDA Project Group on Dynamic Learning during 2010-2015 and published in 2015. Originally, the Cube model was developed to give structure to the various descriptions of what a 'system' could look like in accident investigations. Throughout the debates in the project group, several additional options for application became feasible regarding foresight and early warning signs. Modelling the system behind accidents could better identify successes and failures behind occurred accidents, identifying learning agents and change agents at various system levels. This resulted in a Cube 2.0 version.

The ESReDA Cube has been published on the ESReDA website where it is freely available (ESReDA 2015). As named, it is a cube, a three-dimensional model which can be used as a tool to demonstrate and communicate feasible, credible and plausible system changes and adaptations. The Cube is illustrated in Figures 1a and 1b, and in more detail in Figure 2. The three dimensions represent viewpoints to change potential and learning: 1) stakeholder affected: the organizational levels where learning occurs, 2) work organization: the constituent parts of operational environment where the learning is executed and 3) degree of renewal: an assessment of the impact such learning may have on the system's safety performance. The cross-sections of these viewpoints create 3x3x3 "building blocks" of learning and change, expressed as elements that can be utilized as a systematic way to create new understanding and solutions. The Cube is chosen to visualize issues of a complex nature, in analogy with the Rubik Cube. Careful manipulation of the 'building blocks' should demonstrate the ability to produce the primary system dimensions in a logical and formal manner in an apparently chaotic environment.

The Cube 2.0 version does not add another communication metaphor to existing ones, such as the Swiss Cheese, Domino Stones and Iceberg. The Cube aims at structuring the collected investigation information, interpretation of findings and

potential for change in terms of system dimensions, contributing variables, actors and actions. The Cube has two functions: first it helps the analyst to identify the position of learning agents, and second, it identifies change agents and change drivers in a complex modern socio-technical system. By combining the two functions, a relation is established between the problem definition (what went wrong) and solution diagnosis (what can be done about this). These activities may intervene in the event as well as the system. Since it is not likely that – due to a variety of non-linear relations between multiple contributing factors- the problem definition and solution diagnosis are to be found in the same cell of the Cube, interrelations between cells are identified. Consequently, the Cube may picture the complexity and dynamics of modern socio-technical systems. The actual accident scenario under scrutiny is considered the sequence of events, vectoring its path through the Cube by linking cells.

The Cube is supported by a Template (see ESReDA 2015) that serves as a user guideline for categorizing and ranking of contributing factors, actors and aspects into the Cube. While the Template provides an assessment of the results of the analysis phase of the investigations, the Cube provides a communication tool for further learning and needs for change.

Since the publication of the Cube it has been tested by both members of the ESReDA project group and external researchers. These analyses have led to knowledge that the Cube may be used in several ways that are beyond its original purpose, and to the further development of the Cube. Therefore, the Cube 2.0 presented in this chapter is the more matured second version of the Cube.

Figure 1a. The ESReDA Cube 2.0. A model with three dimensions and 27 individual cells that represent different possibilities to improve safety.

Figure 1b. The Cube may be sliced into planes. If sliced into horizontal planes, each plane represents a different organizational or societal level where safety may be improved. Slicing vertically (different aspects of work organization) or in-depth (levels of learning) is also possible, depending on the objectives of the analysis. We can discriminate between the questions: WHAT to change (aspects, factors), WHO is changing (actors, agents) and WHERE to change the system (components, configurations, concepts)

Figure 2. The "building blocks" of Cube 2.0.

The Cube 2.0 version differs from the original ESReDA Cube most notably by dealing with the former fourth operational aspect "context" separate of the Cube, for two reasons. First, the context adds case specific short-term operational conditions (such as deteriorating weather conditions, occupational stressors or aggravated system states) which may pull the event outside the intended and normal operating envelope. Second, it puts the event in a long-term historical development which may deviate from intended performance by design and foreseen operations. In short, the description of the event in the first phase is a specific representation of the occurrence. This occurrence may deviate from the designed and intended performance of the system under normal conditions in its intended operating envelope. The difference between 'as is' versus 'as ought to be' may provide transparency to what actually happened and requires further investigation into why this difference occurred.

Safety experts are hereby invited to test the Cube and the ESReDA project group is highly interested in feedback from various industrial domains, scientific disciplines and actors with different expertise and backgrounds.

7.3   Application of the Cube in the investigative process

Generally, the investigative procedures can be divided into three phases, each with a specific function. The aim of the first phase is accident reconstruction, the aim of the second phase is the analysis and interpretation of system deficiencies leading to the accident, while the third phase aims at adaptive intervention in the system itself. The investigation itself is an iterative process. In the investigation, and in exploring the usefulness of the Cube, we rely on the tools and techniques that are readily available from investigation practices across various sectors and domains.

Most prominent are:

- Timeline reconstruction of each potential contributing event
- Preservation of volatile data by digital information storage and analysis
- Interview techniques, covering actors, witnesses and bystanders
- Document analysis, such as maintenance and operational logs
- Historical surveys of past performance, accident reports and safety audits
- Inventory and repositories of already published recommendations.

A completely new domain to be explored is the use of Big Data analysis techniques to identify patterns and trends in data that has been collected for other reasons, but that can be useful to create a more complete picture of the system under scrutiny.

Timeline reconstruction is essential since we want to elaborate on the dynamics of complex systems. This is what is asked for in classical investigative questions: "What happened?" and "When?" However, it is even more useful for the investigation of accident when the above-mentioned techniques help reach the reconstruction of the accident scenario. In an ideal case, such a scenario describes the accident as a (possibly branched) sequence of events occurring under certain circumstances.

The second phase of the investigation is to determine the causal factors behind the accident. Here the Cube may be utilized to identify causes systematically and thus achieve a more complete overview of the 'big picture'.

In the third phase of the investigation process, the function of adaptation and change is dominant in phrasing: What can be done to prevent similar occurrences. The Cube serves as a repository of already existing solutions and provides a learning potential for change, varying from optimization and adaptation to innovation. The focus can be on the event, the system or both, depending on the scope and goals of the investigation.

From the system perspective, it is ideal to apply the Cube to the learnings and recommendations of each causal factor. Because it clarifies interrelations between factors, components, decisions and intentions, dynamic behaviour can be made more transparent. Consequently, multiple applications of the Cube are necessary. The number of applications is equal to the number of causal factors that make up the accident. Results of individual applications of the Cube can be surprisingly variable not only because the causal factors are different, but also because the application of the Cube reflects that causal factors i.e. deficiencies in the system's functioning may occur in the different processes implemented within the system.

7.4   Foresight potential of the ESReDA Cube

In order to have foresight potential, it is essential that foresight potential elements can be identified. Every event (accident, incident or near miss) is the end-product

of a breakdown of the design or operations of a system. Such breakdown can be traced back to certification, its safety management system (SMS) or governmental oversight. SMS elements could be of the following nature:

- human-related behavior and procedures
- organisational-related behavior and procedures
- system component (equipment) condition
- environment conditions, etc.

The challenge to obtain foresight is to be able to select key SMS elements that would be ideal candidates to be potential foresight elements that could be monitored and evaluated on a continuous basis.

Additionally, it is essential to stress the importance of multi-stakeholder involvement in the process to obtain foresight potential elements: operators, line managers and senior managers are key actors in the process. At the operational level, operators' knowledge is essential because they are the ones doing the work. They will better understand which behaviors and work routines failed and succeeded. They know their designated system and they would be able to identify where improvement could be made, and which could be candidate foresight potential elements. To complement this, the knowledge of the overall management perspective is very important because floor workers are limited in what they are able to see at the organisational level. They may detect the cause of the organisational stress that led to the event but will be limited in identifying the foresight potential elements at the organisational level. Thus, managers and operators need to be together on this fact-finding journey to be able to design a foresight program and a foresight seeking system ad hoc to the SMS of their organization. They should also be together in the design of the audits and corrective actions (e.g. new procedures) related to the foresight potential elements.

Examples of foresight potential elements that could be regularly monitored and audited by both operators and managers are:

- workflow systematism
- equipment abuse/misuse
- deviation in work procedures
- behavioral trends

- production pressures
- etc.

This chapter describes five application types of the ESReDA Cube. The objective was to verify whether some foresight notions could be obtained for each use case. Each application type is also accompanied by at least one example and discussion on the application's potential to identify early warning signs and foresight potential. Figure 3 below portrays, for each application type, when the ESReDA Cube analysis has been carried out.
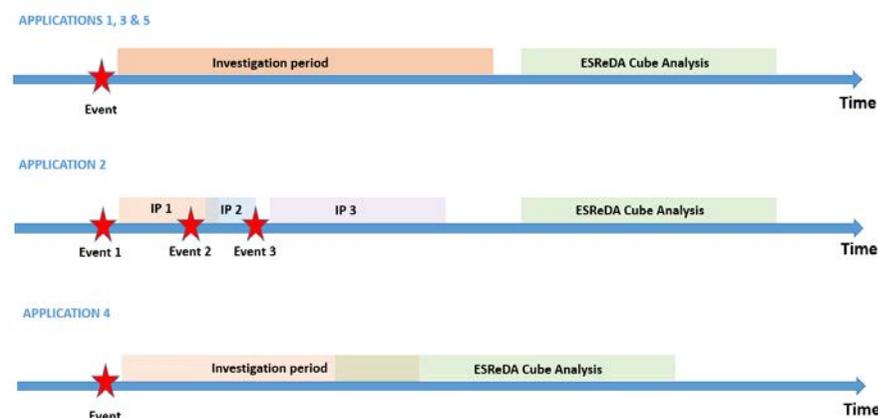
Figure 3. Temporal position of the ESReDA Cube analysis.

7.5 Application 1: Single case analysis

Single case analysis. The original purpose of the development of the ESReDA Cube was the need to construct a model that would help analyse all the pieces of a large-scale accident in a systematic way.

Example of single case analysis
One of the cases utilized in the original Cube analyses was an explosive fire that had occurred in Finland in 2003. Inserting the causes and recommendations onto the Cube identified e.g. the several culture-related changes that had been made in the company's safety management system after the accident, including better reporting of accidents and near misses, and sought-after changes in attitudes and safety culture. Examples of analyses are available in the original ESReDA Cube publication (ESReDA 2015). Also, three large-scale aviation accidents were analysed by Martens (2015).

7.6 Application 2: Theme analysis

Theme analysis is a method to get more from less. Analysing several similar "smaller" cases with the Cube will result in more information than can be obtained from one case alone. Examples of use could be e.g. incidents occurred with LPG (liquified petroleum gas) cylinders, building fires with similar consequences or immediate causes, and other incidents which constitute a need to make an emergency call to the Fire and Rescue Services, but do not usually necessitate the creation of a larger investigation body. The reports of these incidents usually have the common factor that one case itself does not contain enough information to draw wider conclusions that will prevent similar accidents in the future. Nevertheless, combining the information of all known such incidents may result in new findings concerning most common problem areas and e.g. what should be improved on the organizational level. Basically, the outcome of the analyses is a "trend" or knowledge of what the cases have in common. This can be used as

foresight to what might happen if such knowledge deficiencies can be resolved and new solution domains are opened for further elaboration.

Example of theme analysis
Instead of grouping causes, Karanikas et al. (2018) used an adaptation of the ESReDA Cube to group recommendations. They grouped the recommendations of 82 aviation accident investigation reports and the results of their analysis indicated that most recommendations are located near a particular corner of the Cube, but that there are significant differences across investigation authorities. Also, the severity of the accident influences the location of recommendations in the Cube. These results may be used to gain insight of how recommendations are constructed and to subsequently identify best practices for the formulation of effective regulations.

To identify the knowledge deficiencies and new solution, theme analyses could be made inside of an organization, or (probably most important) inside of a network (system). Also, the analysis could be made between different systems (aviation, railway, maritime).

7.7 Application 3: Analysis of investigation success

This type of analysis may be utilized both in-house and by outsiders. Learning from investigations has two purposes: first, to improve the skills of the investigators, which will mean better investigations in the future. Second, to analyse the completeness of the investigation and thus the investigation report. The ESReDA Cube will reveal if there are aspects missing from the report. Aspects which either were thought to be irrelevant by the investigators (in which case they still should have been addressed in the report to avoid elements that remain open and questions that remain unanswered to the reader) or they were actually missed by the investigators altogether, and should therefore be re-opened for scrutiny. A third possibility, which should be addressed in the final investigation report, are the boundaries of the investigation that are due to e.g. a predefined scope (e.g.
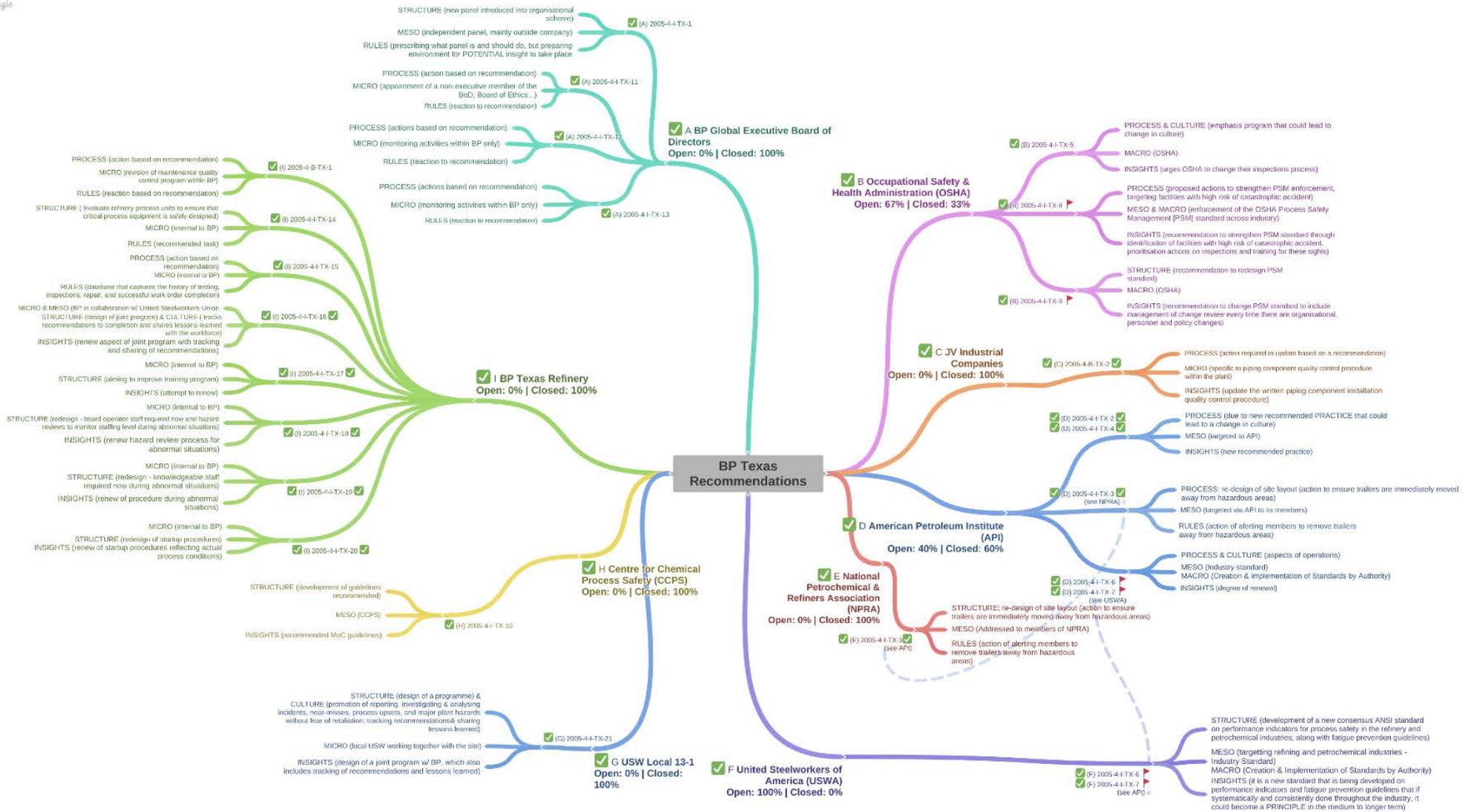
Figure 4. Mind map of recommendations according to groups requiring to implement them, along with ESReDA Cube textual classification (created using Coggle).

"we will not examine further than..", "we will not ask the managers" or "subcontractor's organizational culture issues were not examined") or stoprules during the investigation process, such as the limited time, funding or human resources of the investigative body.

Example of analysis of investigation success

The explosion at BP Texas in 2005 led to several accident investigations executed by different stakeholders. This resulted in the production of several investigation reports, most of them now available online (see e.g. list of references on the Wikipedia page of the accident). Although the incident's chain of events remains the same, in other aspects the investigation reports are significantly different and concentrate on different issues, whichever viewpoint deemed most important by the investigative group. Still, one of the key results of the investigations was that BP Texas had not monitored the process safety status of the site. Nowadays companies have the foresight (as a result of e.g. the investigation of this accident) to include several process safety indicators in their safety management system to gain insight of increase in accident risks via the monitoring of process aberrations.

The CSB accident investigation report concerning the BP Texas accident was analysed by Tulonen et al. in 2018. The analysis was continued further for this report. Before mapping the recommendations onto the Cube, an initial mapping using a mind map (see Figure 4), was carried out according to actors required to address the recommendations. The mindmap helped in the analysis of the recommendations and facilitated mapping onto the ESReDA Cube.

The mapping of the meso-scale recommendations led to some interesting results and related questions:

- There were no triple-loop learning recommendations.
- There were no culture-oriented recommendations.

With the ESReDA Cube each investigation report (like the individual reports concerning the BP Texas accident) may be analysed separately, to identify the priorities of each investigation group. The Cube will also identify the possible differences or contradictions in the results of the investigative groups, and aspects (empty cells in the Cube) that each group has possibly missed or omitted.

On the other hand, as investigations with differing objectives produce reports with differing focus, combining the results of all the investigations and inserting them into the cells of the ESReDA Cube will gather all the results into a systematic and easy-to-understand form that will help assess the incident as a whole. In addition, it will identify accident factors that need to be addressed without the risk of ignoring others.

Perhaps, when designing recommendations, safety boards should take care to also target recommendations that lead towards triple-loop learning and culture-building at all levels (meso in the example above, but also micro and macro). That should be an objective from the start. Interestingly enough, Karanikas et al. (2018) arrived at a similar conclusion: 80% of the recommendations in aviation were targeted toward single-loop learning, and 95% towards other than culture-oriented recommendations (see also summary of their results in Tulonen et al. 2019).

Concerning the example above, it would be interesting to do the exercise for different levels for this event, but also to do it for several other accidents. Would the outcome be similar? It would be interesting to see how many boards in the past have strived to make recommendations addressing triple-loop learning and culture-building in an explicit manner.

Additionally, what would be very interesting to check with this BP Texas accident is to verify if all recommendations have been implemented and consequently map the current follow-up situation onto the Cube. Such an exercise would trigger questions such as:

- What were the obstacles encountered at the meso level that made it difficult to implement a given recommendation?
- Which recommendations led to an improvement in the mapping classification onto the Cube?
- How can these results feed into improving safety-related aspects at the meso-level across the given industry?
- How can we design better recommendations that target a triple-learning loop learning level and a culture-building organisation level?

The use of the Cube serves as a consistency check and quality control on the investigation process management.

In this application, the aim was to validate whether the Cube can be used as a tool for foresight in safety based on knowledge gained from past accidents. When studying past accidents one is usually interested in knowing four main issues: causes, lessons learned, recommendations and follow-up (implementation, effectiveness).

The BP Texas accident causes, lessons learned, and recommendations have been identified as knowledge elements that were mapped onto the Cube cells. This process resulted in a systematic way of better understanding the accident from an operational perspective in terms of different stakeholder levels, operational aspects and degrees of learning potential. The Cube can assist to gain insight on foresight as the mapping process facilitates identification of questions that need answers, such as what do the empty cells mean. Why do some cells have many knowledge elements more than others? Which learning from these knowledge elements can be implemented for foresight in my organization? Answering questions such as these, with a visual tool such as the Cube, facilitates and structures multidisciplinary teams to discuss foresight issues in relation to a specific scope in a systematic manner, e.g. cell by cell.

7.8   Application 4: Investigation support

The ESReDA Cube cannot be classified as an accident investigation method, but as a tool.

Utilizing the Cube at the end of the investigation, or during different phases of the investigation process, can be very useful. After those involved in the accident have been interviewed and the collection of information seems to be at an end, the Cube will help identify which aspects of the accident have indeed been covered and which have not (Cube cells remain empty).

Thus utilization of the Cube will help formulate further questions that require answers in order to obtain a more comprehensive investigation and report.

Example of investigation support
This was tested at the end of an accident investigation concerning a chemical explosion that occurred in Finland in March 2018. (Tulonen et al. 2018). The identified causes of the accident were compared against the cells of the Cube, with the aim of provoking further discussion about potential accident causes.

At the end of the investigation the ESReDA Cube may be used to help construct the final chapter of the report, the recommendations on how to prevent similar accidents in the future.

7.9   Application 5: Recommendations check

Recommendations check helps to detect if there are differences between the results of the analysis, the recommendations and implementation.

Example of recommendations check
This was also tested at the end of an accident investigation concerning a chemical explosion that occurred in Finland in March 2018 (Tulonen et al. 2018). The cells of the Cube were compared to the draft investigation report to identify possible missed viewpoints.
This application was also illustrated in the previous examples in applications 2 and 4 concerning the analysis of the 82 aviation accidents and the analysis of the CSB recommendations concerning the explosion that occurred at BP Texas in 2005.

Ideally, cells in the Cube that should be occupied by root causes should be identical to points in the Cube that should be occupied according to recommendations resulting from the accident analysis and to the points in the Cube that are occupied by the measures which were implemented after the accident.

7.10 Conclusions

The Cube and its underlying template is an investigation tool available to analyse the lessons learned in order to explore solutions and recommendations. The structure of the Template and the Cube provide a checklist for completeness of analytical conclusions, facts, and findings. The results yield information about not only the direct causes of the accident(s), but also the underlying surroundings. This information may be used to formulate an overview and foresight on what should be done to prevent other accidents.

The Cube identifies an overview of shared lessons and solutions for specific types of events beyond a case by case basis of analysing of events. To provide an encompassing overview, single event investigations should be rich of information or combining information into a category of similar cases should be used. The Cube serves as a quality check for completeness of findings and conclusions of an investigation. Rather than counting the number of recommendations and solutions in a cell of the Cube, understanding of their relevance and potential for learning and change counts.

7.11 Discussion

Differences in investigation results may origin from vision, mission and legal mandates of an investigative author, resources, past performance and competences. Differences may also stem from the methods applied, procedures and limitations in qualifications and competences and proficiency in conducting investigations. In general, investigation authorities are highly qualified in the forensic first phases, sophisticated in the second phase of analysis of systemic deficiencies, but less experienced in effectively dealing with systemic safety deficiencies. The Cube 2.0 version might add to their intervention potential in enhancing safety at a systemic level.

Utilizing the ESReDA Cube is integrated into modern accident investigative procedures. The Cube is applied within the different phases of investigative procedure in direct relation to timeline development, creating scenarios (see chapter 5 on scenarios) and system perspective (see chapters 1 and 2).

Creating a timeline is a collaborative effort of all investigative parties and intends to lead to consensus on the sequence of events. Creating a timeline is based on the available data at the moment of the on-scene data collection. Since such data collection is never complete, missing information on critical steps and building blocks may create gaps in understanding the sequence of events and leaves opportunities for speculation, interpretation and continuation of the fact-finding process. A more encompassing data set can be created by combining similar events and re-opening of investigations in case of new evidence. Here the ESReDA Cube shows its strengths.

The company gains foresight from the results of the Cube analyses. The results give information about underlying problems and needs for learning and change. Through this information the company gains foresight to change its safety management system, e.g. to redefine what should be measured – what could be indicators of a rising accident risk.

The ESReDA Cube provokes discussion on learning beyond 'the obvious', beyond the general idea of preventing similar accidents by identifying and eliminating the direct causes of an occurred accident. The Cube invites parties to discuss potential for double- and triple-loop learning: to adapt and to innovate. These discussions will result in new ideas to increase safety, beyond the prevention of 'similar' accidents only by providing foresight to prevent any/all accidents by identifying underlying causal factors that are the root of the system safety as a whole, and needs to learn and change the system in more innovative ways.

Applying the Cube supports the transition from evidence based understanding of the event towards transition based intervention in the safety performance of the system. It visualizes the links between *insight*, -based on the investigative reconstruction-, towards *oversight* over the system operations and safety performance -by analytic interpretation-, leading to *foresight* over the consequences of adaptive interventions.

7.12 References

ESReDA. (2015). Case study analysis on dynamic learning from accidents: The ESReDA Cube, a method and metaphor for exploring a learning space for safety. https://esreda.org/wp-content/uploads/2016/03/ESReDA-dynamic-learning-case-studies-180315-1.pdf

Karanikas, N., Roelen, A. & Piric, S. (2018): Design, scope and focus of safety recommendations: results from aviation safety investigations, Policy and Practice in Health and Safety. https://doi.org/10.1080/14773996.2018.1539385

Martens, F.J.L.G. (2015). Evaluation of the ESReDA Cube Method for the Aviation Sector. First analysis of the methods applicability by applying it on 3 aviation cases. https://repository.tudelft.nl/islandora/object/uuid:b71d978d-7360-49e0-81bd-2af9b751a9e0/datastream/OBJ/download

Tulonen, T., Stoop, J., Vetere Arellano, A.L., Paul, S., Ferjencik, M, Peippo, M. & Teräsmaa, E. (2018). Reasonable recommendations. 55th ESReDA Seminar on Accident Investigation and Learning to Improve Safety Management in Complex Systems: Remaining Challenges. Bucharest, Romania, October 9-10, 2018.

8 Why and How to Employ Organizational Factors for Foresight in Safety?

Frank Verschueren, Federal Public Service Employment, Labour and Social Dialogue, Belgium.[23][24]
Yves Dien, Collectif Heuristique pour l'Analyse Organisationnelle de Sécurité (CHAOS), France.
Nicolas Dechy, Institut de Radioprotection et de Sûreté Nucléaire (IRSN), France.
John Kingston, Noordwijk Risk Initiative Foundation, The Netherlands.

8.1 Executive summary

Organisational factors, which are critical levers of the safety, reliability and resilience of operations and systems, are not considered enough. When in good order, these factors make it more likely that accidents will be prevented and have lower potentials to occur. However, when dysfunctional, these factors make accidents much more likely and serious because they impact many activities and equipment. In some organisations, top management and boards of directors may be unaware of the dysfunctionality of the organisations over which they preside. This unawareness may result from several causes. For example, not receiving or listening to bad news, not being proactive enough to seek out learning and improvement opportunities through event investigation, audit reports, organisational diagnosis in normal operation, or not anticipating future threats.

We provide some definitions and some guidelines for elaborating a framework of analysis that includes organisational factors and the dynamic status of the system state (improving or degrading safety). We illustrate these using some practices from inspection and regulatory assessment of safety management.

Our key messages are:

- organisational factors affect risk in the totality of the organisation, due to a multiplying factor;
- organisational factors have been and are still ignored or under-used;
- top management lack knowledge about the effects of organisational factors because of flawed investigations (hindsight problems), audits (insight problems) and perception of future threats (foresight problems);
- the lack of practical oversight of organisational factors and their interrelationships is an existing operational gap;
- a guiding framework with adequate questions could be developed to fill this gap.

8.2 Introduction: the organisational factors as an opportunity for safety I and II?

8.2.1 Purpose of the chapter

The purpose of the chapter is to provide some reasons and some guidelines for practitioners for investigating, auditing, diagnosing, inspecting and detecting whether organisational factors are leading to a more dysfunctional organisation that degrades safety, or enhancing the reliability and resilience of the organisation and improving safety.

More specifically, the goal is to encourage safety actors (managers, operators, investigators, auditors, and inspectors) to treat 'organisational factors' (OFs) as important variables that could, according to the type of OF, either limit or enable foresight about safety accidents and their prevention. For many safety analysts, those organisational factors have been and are still not enough used by operators and regulators to enhance risk prevention, though they offer a strategic multiplying factor that is worth investing in.

Safety actors' investigations and interpretations should address past and current weaknesses and strengths but should also foresee the future: the likely outcomes of existing trends, as well as forthcoming threats and opportunities. On this basis, improvement actions can be taken to increase safety margins and, perhaps, acknowledge the conditions that are essential to achieve current success.

The assumption is that organisational factors can be very powerful and durable conditions that lead to safety degradation, and which have contributed to many

---

major accidents. However, it also assumed that OFs are key levers to reach a more reliable and resilient organisation and improve safety.

In general, accidents do not happen 'like lightning striking out of a clear sky'. Before major accidents happen, it has been observed that there occur early warning signs (EWS), such as near misses which were not recognised or alerts by some actors which were not treated early enough during the "incubation period" (Turner, (1978). This empirical lesson from accidents provides an opportunity for preventing other accidents, but it depends on investing sufficiently action-oriented effort in investigations, audits, inspections, and situation understanding / interpretation, as well as the management and regulatory actions that achieve effective change.

Concerning strategy (i.e. how to improve foresight and prevention by addressing organisational factors), two paths can be followed by safety actors:

- On the one hand, there is a normative route that starts by comparing the conditions in an organisation to some ideal model (founded on theory, experience or research findings) about organisational safety, reliability and accidents and implementation of measures for accident prevention, then makes recommendations designed to make things better.
- On the other hand, there is a more dialectical route, for example: obtain data and offer interpretations; then discuss different interpretations; act on some things, but on other things ask new questions; collect more data, encourage stakeholders continue to contest meanings, and so forth.

Of course, a combination of the two paths is also possible.

This chapter will first define in general the organisational factors and show the logic that connects facts to organisational factors of accidents (as root causes but especially as underlying (latent) causes). It will also provide readers with some examples, directions and guidance to inquire and act on organisational factors, such as from the BP Texas City accident, organisational diagnosis of safety

management in the nuclear industry in France, and regulatory inspections of Seveso chemical plants in Belgium.

As a final element, the future principles for elaborating a framework are explained. This 'OF Framework for questioning Foresight in Safety' is a work in progress[25]: the present authors aim to develop a 'part two' to this 'part one'. This kind of framework could be used by safety actors (managers, operators, investigators, auditors, and inspectors) to develop their own "road map" to guide themselves during processes of investigation, diagnosis and inspection, specific to the features of the organisation investigated. For instance, it aims at giving further plausible directions and targets for organisational investigation once a dysfunctional organisational factor is detected and confirmed.

8.2.2 Organisational factors and failures of foresight: BP Texas City refinery accident

The Texas City BP refinery accident that occurred in 2005 is a reference case. It illustrates the key concepts: organisational factors[26], early warning signs and opportunities for Foresight in Safety (FiS) through organisational analysis of safety. It is a typical case where the accident confirms the prognosis "an accident waiting to happen" that was made by several actors (managers, workers, and health and safety engineers) and by different processes (internal audits, event investigations, and an external audit by a consultant) (Dechy et al., 2011).

On March 23, 2005, an explosion and fire at the BP refinery in Texas City led to 15 deaths and 180 injuries. The board member and CEO of the US Chemical Safety Board (CSB), Carolyn Merritt (2007) said: "The combination of cost cutting, production pressures, and a failure to invest caused a progressive deterioration of safety at the refinery." But the accident has its roots deeper in the past. The CSB report (2007, p. 20) found that "cost-cutting in the 1990s by Amoco and then BP left the BP Texas City refinery vulnerable to a catastrophe." The CSB (2007, p. 18) noted also that "The Texas City disaster was caused by organisational and safety deficiencies at all levels of the BP Corporation. Warning signs of a possible disaster were present for several years, but company officials did not intervene effectively

---

[25] 1/ our work is in Progress, because not fully tested; 2/ we only provide directions, guidelines on how to develop a framework for specific inspection, audit, investigation, because the examples provided do not cover all topics

[26] We chose this accident as paradigm, knowing that many other events could have been chosen (e.g. the nuclear power plant accident at Three Mile Island (1979), head on collision of trains at Ladbroke Grove (1999), disintegration of the space shuttle Columbia (2003, …). https://www.csb.gov/u-s-chemical-safety-board-concludes-organizational-and-safety-deficiencies-at-all-levels-of-the-bp-corporation-caused-march-2005-texas-city-disaster-that-killed-15-injured-180/

to prevent it." Merritt (2007) also said that "adhering to and enforcing federal regulations already on the books would likely have prevented this accident and its tragic consequences."

Indeed, the CSB investigation showed that some BP members had identified the major risks already in 2002. The new director of BP's South Houston Integrated Site observed in 2002 that the Texas City refinery infrastructure and equipment were "in complete decline" (CSB, 2007, p. 151). In consultation with senior managers based in London, the director ordered a study that looked at mechanical integrity, training, safety, and economic opportunities. The study, which was shared with London executives, concluded that mechanical integrity was one of the biggest problems (CSB, 2007, p. 151).

The BP Group Refining Vice-President suggested a follow-up inquiry asking in an e-mail (August 16th, 2002), "How has [South Houston] gotten into such a poor state?" This follow-up report, entitled "Texas City Refinery Retrospective Analysis," was issued later in 2002, and had the objective of determining why Texas City refinery performance had deteriorated. The analysis concluded that "the current integrity and reliability issues at TCR [Texas City Refinery] are clearly linked to the reduction in maintenance spending over the last decade." Capital spending was reduced 84 percent from 1992 to 2000 (CSB, 2007, p.153).

Several other studies, surveys, audits and also serious incidents alerted and signalled the severity of deficiencies, but the response of BP managers was "too little and too late" (CSB, 2007, p. 26) with the implementation of corrective action plans that were poor. CSB found that "at the end of 2004, the Texas City site had closed only 33 percent of its PSM [process safety management] incident investigation action items; the ISOM [isomerisation] unit closed 31 percent. Furthermore, CSB note that BP management made a presentation in November 2004 on the reality of safety, saying: "Texas City is not a safe place to work" (CSB, 2007, p. 172).

BP managers were not alone in holding these views. A safety culture assessment conducted by an external consulting company (Telos Group) alerted the managers in January 2005 to the critical and degraded state of the refinery. The Telos report identified many of the same problems later found by the CSB in retrospect after the March accident. The business unit leader who initiated the Telos survey was looking for "brutal facts" concerning "our management systems, our site leadership, our site cultures, and our behaviours for safety and integrity management" (CSB, 2007, p.168).

The CSB (2007, p. 169) stated that the Telos safety culture assessment findings included:

- Production pressures impact managers "where it appears as though they must compromise safety."
- "Production and budget compliance gets recognised and rewarded before anything else at Texas City."
- "The pressure for production, time pressure, and understaffing are the major causes of accidents at Texas City."
- "The quantity and quality of training at Texas City is inadequate…compromising other protection-critical competence."
- "Many [people] reported errors due to a lack of time for job analysis, lack of adequate staffing, a lack of supervisor staffing, or a lack of resident knowledge of the unit in the supervisory staff."
- Many employees also reported "feeling blamed when they had gotten hurt or they felt investigations were too quick to stop at operator error as the root cause." There was a "culture of casual compliance."
- Serious hazards in the operating units from a number of mechanical integrity issues: "There is an exceptional degree of fear of catastrophic incidents at Texas City."
- Leadership turnover and organisational transition: the creation and dismantling of the South Houston site "made management of protection very difficult."
- The strong safety commitment by the Business Unit Leader "is undermined by the lack of resources to address severe hazards that persist," and "for most people, there are many unsafe conditions that prove cost cutting and production are more important than protection. Poor equipment conditions are made worse in the view of many people by a lack of resources for inspection, auditing, training, and staffing for anything besides 'normal operating conditions.'"
- Texas City was at a "high risk" because of a widespread "check the box" mentality. This included going through the motions of checking boxes and inattention to the risk after the check-off. "Critical events,

(breaches, failures or breakdowns of a critical control measure) are
generally not attended to."

When the business unit leader received the results, he wrote (in an e-mail March 17, 2005) that "seeing the 'brutal facts' so clearly defined was hard to digest, including the concern around the conflict between production and safety. The evidence was strong and clear, and I accept my responsibility for the results" (CSB, 2007, p.171). But the same day he wrote a summary to all plant supervisors stating that "the site had gotten off to [a] good start in 2005 with safety performance that "may be the best ever," adding that Texas City had had "the best profitability ever in its history last year" with over $1 billion in profit—"more than any other refinery in the BP system" (CSB, 2007, p.171).

The downward trend of reduction in the numbers of occupational incidents was misinterpreted by some managers as a sign of improvement of industrial safety, while the number of losses of containment increased (from 399 to 607 per year from 2002 to 2004) and costly accidents occurred (e.g. 30 million $ in 2004). But at the same time, the 2005 Texas City HSSE Business Plan (presented March, the 15th) warned that the refinery likely would "kill someone in the next 12-18 months." This fear of a fatality was also expressed early 2005 by the HSE manager: "I truly believe that we are on the verge of something bigger happening," referring to a catastrophic incident (CSB, 2007, p. 173).

Thus, the lessons learned from this accident clearly show that signs of deteriorating safety had been detected by many actors, despite the differences in their approaches and methods (observations from operators and from managers, internal and external audits, safety culture survey, incident investigation) and were confirmed after the accident by the CSB investigation (Dechy et al., 2011). In general, 'advanced' industrial systems are resistant to errors and the accident is "hard to obtain" (Perrow, 1984). An 'incubation period' (Turner, 1978) is observed, implying an accumulation of EWSs. The systematic study of accidents (Llory, 1996) demonstrates that the deficiencies are sometimes severe, often visible to a certain number of actors that are able to make the adequate diagnosis or prognosis if they are given adequate resources.

8.2.3 Why do organisational factors have such potential to enhance or endanger safety?

The BP Texas City case is a trenchant example and costly reminder of the significance of organisational factors to accident prevention. Organisational factors are of strategic interest for accident prevention because they create the conditions in which safety efforts benefit from a multiplication factor on the positive side (safety II) and are the basis on which to counter negative effects on a larger scale (safety I).

This multiplication factor deals with the impact of local factors versus macroscopic factors on probabilities of errors and failures (as illustrated in the following scheme). This multiplication factor is the key strategic reason why it is worth employing and investing in organisational factor leveraging for safety enhancement.
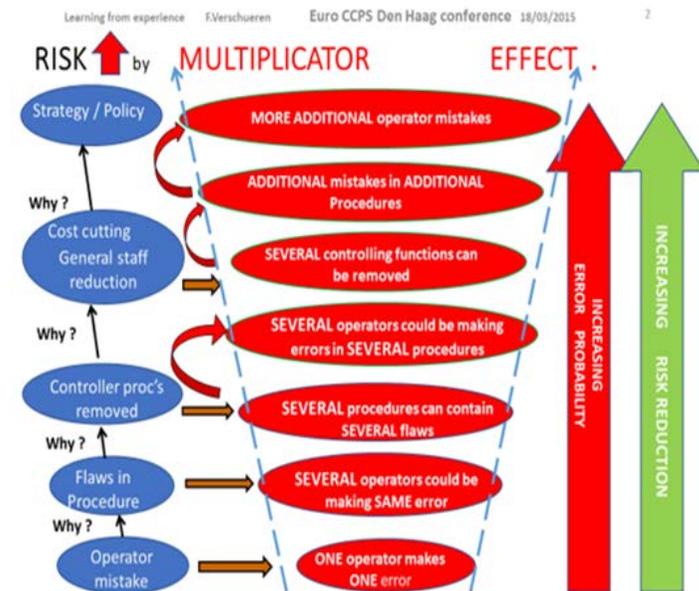
Figure 1. Organisational factors strategic interest for prevention-the "Multiplication Factor" (adapted from Verschueren, 2015)

Figure 1 shows how an investigative process (accident investigation, organisational diagnosis, or audit) could develop on different levels of the socio-technical system

(e.g. Rasmussen, 1997) and how measures to reduce risk could be implemented at all levels although not with the same impact.

It starts with one of the direct causes of an accident/incident (investigation) or a direct potential disturbance and weakness or deviation (audit). A generic example of a direct cause is 'operator makes a mistake'. If lesson learning reduces a complex event to a single human error and marries this to a corrective action 'fire the scapegoat of the organisational chain' or 'train the individual', then the 'multiplication factor' at this most local level is the smallest: 1 to 1: one operator fired or better trained. This at best stops the repetition of one mistake or some mistakes at this level. We are on a 'human error' level.

What if someone else was in a similar work situation: could an error occur or recur? If one searches for the possible causes of the operator error, the verification by auditor/investigator of the procedure used by the operator could lead one to find errors in the written procedure. So, the next level cause would be the 'faulty procedure'. Here the multiplication factor will be more than 1, especially if the task is performed by more than one person or when someone else takes the position: one erroneous procedure => several operators could make similar mistakes.

We are now at the 'human factors' level related to 'working situations'. Faulty procedures are not the only factors influencing working situations (e.g. human-machine interface, staffing, fatigue, etc.). Of course, an investigator or auditor shouldn't and usually will not stop here (beware of 'stop rules'[27]) but should further ask 'why'; for example, why this procedure contained an undetected error(s). The investigator might find that, in the past, there existed a person/function who checked every new written procedure before it was put into use and that at the time of investigation/audit this function no longer existed. If one procedure contains error(s), then it brings into question the whole process of designing and verifying procedures. In this situation, the multiplying factor would be even greater than those restricted to one working situation. This is because a flawed process of designing and checking procedures => possible more procedures contain errors => even more errors could be made by operators.

Once more the auditor or investigator should ask why this function/person was removed from the existing control chain if that is the reason that explains the

flawed organisational process. One of the possibilities he could find is that this function was removed during (and due to) a cost reduction campaign.

Depending on the size of the company/organisation (several services, several departments) it could be that during this cost-reduction campaign, several other control functions were removed. It should be clear that in that way the 'semi quantitative' relation will be enhanced again. Thus, due to cost cutting => removing several control functions => possible even more procedures or measures to manage risks over the whole organisation contain undetected errors => even more errors could be made by operators in all departments and all working situations.

The cost reduction in this example might have been decided in a company-wide overall strategy definition and policy review aimed at finding a new financial balance and, of course, this cost-cutting campaign may have other transversal effects to other organisational controls beyond just reducing the "procedure designing and checking activity". We are now on the "highest organisation" level.

The investigation would continue, of course, to question the rationales and the evidence for such decision-making processes. This was done by the US CSB for BP Texas City refinery from the business unit level in South Houston to the top of the BP Group at BP headquarters, the board of directors in London and even questioned the role of regulators, especially in the US (OSHA and EPA).

The essence of all this, is that with every new level [operator => procedure => checking function => cost reduction => strategy], the probability of having errors and the negative impact of a flawed process gets multiplied.

If we state this in a positive manner to enhance robustness and resilience: if the organisation responds with the related measure on a higher level, the gain in risk reduction will be larger: many more potential mistakes will be prevented (green arrow: 'increasing risk reduction'. In other words, the higher the level of organisational factors the auditor/investigator gets to and that a manager corrects, the more they open windows of opportunity to prevent many types of accidents (and not only a similar one).

8.2.4 History of accidents and the importance of organisational factors

ESReDA has campaigned since the 1990s for better analysis of events and investigation of accidents (ESReDA, 2003, 2005, 2008; Roed-Larsen et al., 2004; Dechy et al, 2012, Dien et al., 2012), for 'dynamic learning' to consider the issues of the follow-up of lessons from investigations, and for the removal of barriers to learning (ESReDA, 2015a, 2015b, 2015c). However, many event and accident reports were still not "well investigated and analysed" meaning that they did not correctly address organisational factors (Dien et al., 2012). This issue is revisited in subsection 8.3.3. We will define these organisational factors (also known as organisational influence factors) in table 2 of 8.2.5 and distinguish them from human and technical factors.

Researchers who reviewed many 'well' investigated and analysed accident reports (e.g. Cullen, 2000; CAIB, 2003; CSB, 2007;…) observed recurring root causes, similar patterns, "striking similarities", "echoes", "parallels" between accidents (Turner, 1978; Llory, 1996, 1999; Reason, 1997; Rasmussen, 1997; Dien et al., 2004, 2012; Hayes and Hopkins, 2015). These recurrences offer the opportunity to capitalise them into a 'Knowledge of accidents' (Dechy et al., 2010, 2016, 2018). This knowledge of accidents is then useful for guiding and for interpreting in organisational analysis of accidents (Llory and Dien, 2010a, 2010b; Llory and Montmayeul, 2010, 2018) and organisational diagnosis of safety (Rousseau and Largier, 2008; Dechy et al., 2011, 2016, 2018).

One outcome is related to the identification of a pattern of accident causation. An accident model has been observed and defined (Dien, 2006; ESReDA, 2009) with weak signals of safety degradation (Turner, 1978), latent failures (Reason, 1990, 1997) that go unrecognised during an incubation period (Turner, 1978). For Foresight in Safety (FiS), this accident model and definition have implications for risk prevention and specially to provide an opportunity to detect and act on early warning signs (EWS) before a severe or major accident happens.

A socio-technical system generates uncountable gigabytes of data every day; some information, and potentially some EWS, will get buried in the noise. Some of the EWSs are the symptoms of a safety degradation caused by a root cause of an accident waiting to happen.

To establish or enhance FiS, the goal of an investigation, audit, inspection, diagnosis that aims at preventing an accident is to capture those EWSs. Strategies to collect data, to generate and filter information, to recognise and interpret signs related to negative and positive organisational factors) will be proposed.

8.2.5 Definitions of Organisation, Technical, Human and Organisational Factors

As many writers and disciplines use the terms' 'organisation' and 'organisational factors', for a practitioner perspective, it is important at this stage to start to clarify our definition of these terms in the context of technical, human and organisational factors in high-risk industries. The document also suggests a classification scheme for organisational factors. For example, it locates management systems as a subset of organisational factors, and suggests governance 'functions' as another subset.

Definition of 'Organisation'

An organisation is an entity comprising multiple people that has a particular purpose and is more than a crowd of acting actors.[28] It can operate in the public sector (fulfilling public duties) or private sector (developing commercial activities) or in both. For the purpose of this chapter on organisational factors for FiS, we will look into the organisations as found in high-risk industries. In this perspective, an organisation is viewed as an entity being organised, reorganised and where the focus is to support processes, tasks, decisions and actions that enable sustainable performance and risk management.

An organisation may be therefore understood as the planned, coordinated and purposeful action to reach a common goal or construct a tangible product or service. Part of the organisation is governed by formal and management provisions, structures, systems, processes, rules, procedures, auditing, inspections to implement, enforce, enable the different activities of workers at all levels of the socio-technical system; part of the organisation is controlled by real practices of workers in the field, taking into account informal aspects, making sense of signals,

---

[28] *An organisation is more than a crowd. Because it has a particular purpose, an organisation imposes constraints on actors' behaviour. In contrast, each actor in a crowd of undifferentiated individuals has more freedom to act on his own accord, albeit within the wider dictates of society at large.*

individual and collective decision-making processes and the influence of power, social and cultural aspects as well.

Other authors refer to an organisation also as a socio-technical system (e.g. Rasmussen, 1997; Rasmussen and Svedung, 2000; Leveson, 2004) as it is a combination of social elements (individual and groups of people belonging to the organisation), technical elements (infrastructure, installations and individual apparatus), interacting and performing different activities to produce or operate safely. Additionally, an inter-organisational dimension is to be considered (Wilpert and Fahlbruch, 1998), or organisational network (Dien, 2006), implying to consider interactions with regulators, subcontractors, competitors, non-governmental organisations, citizens.

By coordinated and planned cooperation of all these elements the organisation can solve tasks that lie beyond the abilities of the single elements. This is the positive side of the mix of these socio–technical–organisational elements. The negative side is that because of the same mix most organisations tend to be complicated or even complex by nature.

Definition of 'Causal' and 'influence' factors: technical, human and organisational

In accordance with research on accident and system models (e.g. 'organisational accident' in Reason, 1997; 'socio-technical system' in Rasmussen, 1997) and accident investigation such as Management Oversight Risk Tree (Johnson, 1973), Tripod (e.g. Groeneweg et al., 2010) developed on basis of the Swiss Cheese model (Reason, 1990, 1997), Accimap (Rasmussen and Svedung, 2000), STAMP (Leveson, 2004), organisational analysis (Dien et al., 2004, 2012), we can distinguish three types of causal and influence factors influencing safety in an organisation: technical causal factors, human and organisational influence factors.

Technical factors can be considered as causal factors because they refer to a mechanistic and deterministic causality, while human and organisational factors are better considered as influence factors because they refer to complex relationships, with cause-consequence relationships transformed and belonging to a different paradigm (Morin, 1977), and that are more probabilistic (e.g. Dien, 2006; ESReDA, 2009; Vautier et al., 2018).

In many accident investigation methods (Sklet, 2004; Institute of Energy, 2008), it is common to distinguish direct or immediate causes, which are the last "stage" of

the event, meaning that they are both the visible phenomena and consequences of root causes. Direct causes are generally technical failure and / or human error, while root causes are related to underlying deficiencies in upper levels of the socio-technical system and latent effects (Reason, 1990) such as an inadequate maintenance policy.

Table 1. Definitions of technical, human and organisational factors.

Type of causal and influence factors	Definition
Technical, causal factors	Technical causal factors related to technical elements: processes used in the industrial organisation and technical components (equipment, apparatus and installations) used in these processes. In the immediate chronology of events, often, the failing of critical equipment can lead to the start of an incident sequence or can lead to the failure of a technical barrier, so an incident sequence is not interrupted but continues and escalates. In the remote chronology of events, the failures of equipment and barriers are influenced by underlying human and organisational factors levels (e.g. poor maintenance action because of inadequate competencies and resources for adequate maintenance).
Human, influence factors	Human influence factors are factors that influence and may determine the performance of an individual, such as fatigue in some working situations. They are related to all humans (operational people, as operators and planners; people in all supporting services, such as maintenance, design, research, logistics, and procurement; decision makers on all levels from front-line operators and front-line technicians, the supervisors, the managers, every senior manager up to the CEO and the Board of Directors. They can also be identified in the work situation and activity of each actors (does he/she have the resources and tools (e.g. man-machine interface, procedure) to properly achieve the required job?), which lead to human errors such as omissions and faulty human

Organisational influence Factors	Definition
	decisions which impact an individual action or (possible and mostly worse) the actions of many others.
Organisational, influence factors	They are the factors relating to the influence of the organisation. We propose to distinguish here three groups of organisational factors (Table 2.)

Table 2. Different groups of organisational influence factors.

Organisational influence Factors	Definition
Management system failures	Indeed, an organisation has several essential Management Systems to "manage" its activities, especially its key functions (production, safety...). Examples of Management Systems (MS) are Production, Safety, Quality, Project, Maintenance, Logistics, Procurement, Human Resources, Facility, and others such as risk analysis, learning from experience, management of change, emergency management... Each of these Management Systems can fail and produce "System Failures" which directly impact the equipment and barriers deficiencies, unsafe acts. When these Management system failures emerge, they can generate or contribute to other failures (technical or human factors). The criticality of these Management system failures depends on their direct or indirect impact on safety critical elements in the scenarios of major accidents. As production, quality or maintenance management systems tend to be more closely related to safety, their failures are more often critical.
"Organisational dysfunctionalities"[29] [30]	Organisational dysfunctionalities have a direct or indirect impact on almost every part of the organisation. They can directly cause safety management system failures. There is an internal part of the organisation (usually, at the 'top' of the organisation: Board, Senior Managers, decision makers) that has greater responsibility than other actors for the adequate internal governance of the total organisation. This top part of the organisation has the role and power to define internal governance functionalities as the strategy of the organisation (mission and vision, priorities, and strategic objectives), the policy of the organisation (setting of objectives and deployment in tasks in order to reach the strategic objectives) and the structure of the organisation (roles and functions with their authorities and responsibilities, power distribution, trade-off processes). But also, on top of formal dimensions of the organisation, historical, social and cultural factors may facilitate or hamper organisational performance. If this strategy, policy or structure is poorly defined, implemented or protected, then the organisation will be in a state of dysfunction and will not reach its purpose (as per definition) with inadequate decision-making processes and trade-offs.
Regulation dysfunctionalities	There is an external part of the operating organisation, the regulatory context (laws), and their enforcement by control authorities (e.g. inspection). The external directed governance of the system includes the operator and his relationship with the regulator but also internal (health and safety committees, trade-unions) and external stakeholders (NGOs, neighbours...).

---

[29] *In this work, the term "dysfunctional" is used for all organisational factors that lead to an impaired function, a failing to serve an adjustive purpose (here, the safety of an organization with all its constituents especially including people).*

[30] *Presentation by Frank Verschueren "Learning from Organisational Dysfunctionalities" at Energy Institute Conference on "Human Factors Application in Major Hazard Industries", (17 - 18 Oct 2017); London.*

8.3 How to employ organisational factors for Foresight in Safety? Some guidance and examples in investigation, auditing and inspecting

8.3.1 Hindsight, insight, foresight

We distinguish three temporal perspectives where organisational factors can be employed to the benefit of accident prevention: hindsight, insight and foresight.

In our view, the goal is to turn hindsight (past) or insight (present) into foresight (future).

In our definition, foresight has two main activities: interpreting the weaknesses/strengths and vulnerabilities/resiliencies and making a prognosis of their outcome or sufficiency (safety margins); and, eliminating deficiencies to reduce or eliminate risk factors of recurrences of accidents; creating new measures to prevent accidents.

In this chapter, to create some foresight in safety, we will see and study events in the past or present to proactively prevent accidents in the future and also give an example of foresight of future risks. Other chapters of this ESReDA deliverable address the three parts.

As mentioned in the introduction, concerning the strategy of how to improve foresight by addressing organisational factors, there seem to be two paths that safety actors can follow (the normative route or the more dialectical route), taking in consideration that a combination of the two paths is also feasible.

8.3.2 Past/Hindsight: Improving Investigation of Accidents and Incidents to gain Foresight

A first source of FiS, foresight in safety, is exploring the system or organisation in hindsight, meaning studying the past. It implies studying the accidents and events which happened in the past and deducing all the pertinent and likely recurring causes related to organisational factors.

It is important to develop approaches to address the real root causes and to detect the relevant organisational dysfunctionalities that contributed to an event. The lessons from past accidents and incidents can help to show how organisational

factors can improve or degrade safety within an organisation or in general for all high-risk organisations if the study is enlarged to all industries' accidents.

It can be valuable also to study past normal functioning to establish the former state of SMS performance, organisational and regulatory performance as well.

It may seem paradoxical that an approach of FiS advocates first looking at the past, but the idea is to find specific and generic factors to prevent similar events to recur in the future elsewhere as well.

Searching for underlying root causes to prevent next accidents

Investigating root causes related to underlying deficiencies in the depth and history of organisations is far from an easy task. It requires a multi-disciplinary approach aiming at questioning the different dimensions of the socio-technical system that may influence accident causation. The investigation commission into the space shuttle disintegration, the Columbia Accident Investigation Board put it this way: "Many accident investigations do not go far enough. They identify the technical cause of the accident, and then connect it to a variant of "operator error" – the line worker who forgot to insert the bolt, the engineer who miscalculated the stress, or the manager who made the wrong decision. But this is seldom the entire issue. When the determinations of the causal chain are limited to the technical flaw and individual failure, typically the actions taken to prevent a similar event in the future are also limited: fix the technical problem and replace or retrain the individual responsible. Putting these corrections in place leads to another mistake – the belief that the problem is solved. The Board did not want to make these errors" (CAIB, 2003 - p. 97).

Examples are numerous and can be easily found on the web, so we do not need to develop an example here. Several very good accident investigations (e.g. Cullen, 2000; CAIB, 2003; CSB, 2007) highlighted that some incidents and events before accidents were not properly investigated. In the British rail network, several signals were passed at danger (signal was at red) by train drivers, highlighting systemic vulnerabilities in a complex system. However, rather than consider the systemic effects of privatisation, specifically the fragmentation of the system, the passing of signals at danger were considered by railway management as wholly the responsibility of drivers (Cullen, 2000). The CAIB (2003) found that prior foam losses were not well analysed by NASA, especially why the foam losses occurred more frequently on the left side of the bipod. The CSB (2007) found that several

incident investigations by BP Texas City refinery failed to address root causes, especially an investigation of a 30-million-dollar accident in 2004.

Multi-layered approach or "Why does an investigation need to be considered on multiple levels?"

Reason (1990, 1997) developed an accident model that included the concept of "latent causes". These latent causes could be, for example, an undetected deficiency in an equipment design or the consequence of a poor maintenance policy, which opened the questioning towards the engineering and management failures (in decision making). Those failures of decision making were situated in the front-line team and its management, middle management or senior management. So, these alone made already three levels of an 'organisational accident' model (Reason, 1997).

The simplest methods for investigating accidents focus on one level (Frei et al., 2003; Sklet, 2004): the actors in the front line, typically, operators and technicians, such as train drivers. They produce direct causes (activities and decisions of these "front liners") with a very limited "penetration insight".

Many accidents have shown that by questioning upwards the role of hierarchical lines, one can detect how the decision making of higher management levels play a large role in degrading the working conditions of several frontline actors. These conditions are later involved in combination with other direct causes, triggering events, and are therefore influencing the causation of incidents and accidents as underlying causes. Only a thorough investigation of the accident will reveal the different connections between direct and underlying causes.

A thorough investigation of the accident should consider the multiple levels of the organisation, which interact (e.g. human resource management impacts the skills available in working situations) and with the environment (e.g. technological and political changes have an impact on: the skills needed to operate new technologies; the level of regulation, and; the acceptability of high-risk industries). Several researchers recommended this multi layered approach, and among them, Reason (1997), Rasmussen (1997) and with Svedung for Accimap (Rasmussen and Svedung, 2000). It was followed by others, such as Nancy Leveson's STAMP (2004). Several other examples exist (in Sklet, 2004; Dekker, 2006; Dien et al., 2012; Institute of Energy, 2008).

The layers of governance or supervision levels should be considered in those organisational 'cause–consequence' schemes that can be visualised in a top-down perspective in accordance with the organisational accident view (Swiss cheese model of Reason, 1997). A complementary but consistent view integrates the bottom-up flows of information. An important issue is to consider the system behaviour as a product of interactions, with consideration of systemic effects as a whole (e.g. Vautier et al., 2018).

It remains a challenge to go up to the top (i.e. senior management and the board of directors) who make the decisions on strategy, policy, structure, and "steer" the whole organisation. As they make very important decisions on resources and budgets, they can either limit or enable discretion at lower levels. Examples of policies of cost cutting are numerous (e.g. Texas City) and can paralyse the management of integrity maintenance and the whole departments related to safety.

For that reason, it is rare that internal people from lower organisational levels doing investigations or auditing do ask questions in their investigation all to the top of their own organisation to show how strategy, policy or structure were dysfunctional or not. There is self-censorship, stop rules, pressures from management and taboo subjects (e.g. Dien et al., 2012).

8.3.3 Present/Insight: Improving Auditing and Diagnosis to improve Foresight in Safety

A second source for gaining FiS is exploring the system or organisation in insight, meaning studying the present. It requires that the company governance, provisions and practices related to safety are investigated and assessed to detect dysfunctionalities and assess the quality of safety management.

Organisational performance can be assessed in a range of dimensions, production, quality, safety, environment, social… For our safety purpose, it is important to develop approaches that provide insight into on-going performance, through the knowledge of influencing factors on daily performance, that can be related to company's governance and safety management. When we use the term 'real performance, in practice', our meaning goes beyond effort to formalise safety management processes, provisions, procedures and is related to the concepts of activity and work, that are more than applying procedures. These approaches require as well, collecting formal and informal data, interpreting information, signs

and infer EWS, symptoms or indicators of weaknesses or strengths related to best practices, reliability and resilience organisational factors.

Analogies of Auditing with Investigating

Conducting an organisational diagnosis in hindsight differs from conducting it on the basis of real-time data. Some analysts consider that the two configurations are radically different. Indeed, analysis of events is often criticised for its hindsight bias (e.g. Reason, 1990; Vaughan, 1996). Knowing the end of the story brings an effect of wisdom (Reason, 1990) to the investigator (however investigating root causes is not easy!) while the actors in the system did not benefit from this knowledge prior to the accident. Hindsight bias can be "harmful" if the aim of the investigation is to find (only/mainly) one or more people to blame. Especially, EWS would be easy to detect in a retrospective approach, while actors seeking insight in real time have inherent difficulties to extract EWS from daily noise. On that point, Vaughan (1996) considered that some weak signals could not be understood before the accident because they were normalised in NASA culture.

It is partly true, but we disagree with the generalisation to all cases. The Texas City accident just recalled in this chapter is a contrary example and shows that many EWS were, indeed, recognised before the major accident by operators, managers, audits and event investigations.

In fact, there are similarities in the two configurations of organisation's diagnosis (hindsight and insight) (Dechy et al., 2011). Background knowledge in human and organisational factors used by analysts (investigator, auditors, inspectors) are mainly common requirements for both. The interviews with people might be biased (people who refuse to speak-up or misinform investigators) in the two configurations, but not on a dichotomous basis.

In both cases, we can find events that can reveal symptoms of organisational weaknesses prior to the accident or the diagnosis in normal operation (e.g. Texas City refinery accident). In both cases, making an expert judgment upon the complex causalities of influence factors remains uneasy, though evidence and proofs of (un-)reliability might differ.

Auditing and Inspecting: Looking for underlying deficiencies in Safety Management

An example from Belgium Competent Authority for Seveso plant regulation[31]

The following example relates to a normative approach which is especially efficient for compliance-oriented approach, but also because there is consensus on the expectations of what are good practices for the management of safety. For instance, it requires that safety management systems are implemented (a regulatory requirement of the Seveso Directive). Also, safety actors know that good event investigations should address root causes. But, is it the case?

A study by the Belgium Competent Authority found that a sizeable minority of companies have in their incident investigations a 'blind spot' to organisational factors. In a representative sample of Seveso companies, the study found that 36% carried out investigations that poorly identified the underlying organisational causes of events, and 27% carried out investigations using procedures that were very poor at identifying underlying organisational causes and organisational factors. [32]

All companies under the Major Hazards regulation in Belgium are obliged to have a complete and well-functioning Major Accidents Prevention Plan (MAPP). The core element of this MAPP is a safety management system to prevent major accidents. One of the components of a safety management system is the investigation of accidents and incidents; this is audited by Seveso inspection agencies in their role as enforcers of the Seveso regulation in Europe.

In 2013, the Belgian Seveso inspection agencies studied specific regulatory audits to get a more detailed view on the strengths and weaknesses of companies' accident and incident investigation systems. The data for the study were drawn from seventy audits that had been carried out by 14 inspectors over four years.

---

[31] Presentation "Inspection of Investigations of Accidents and Incidents" by Frank Verschueren at MJV Seminar on Learning of Incidents (11 - 13 September 2013 Gothenburg, Sweden)

[32] These findings are consistent with the practices declared fifteen years ago (limited use of investigation procedures, very limited investigation of root causes, very limited involvement of experts in

investigation, learning, human and organisational factors (ESReDA, 2003; Roed-Larsen et al., 2004, Dechy et al., 2012) though we can observe some improvements.

The seventy audited companies are a representative sample of the total 375 companies under Seveso regulation in Belgium.[33]

Each company was audited in the same way using a specific inspection instrument. This instrument contains 53 questions arranged in eleven question blocks. The general topic of incident/accident investigation is spread across a number of these blocks, such as those focused on the reporting system, investigation & analysis, and remedial actions. Each of these blocks is subdivided into smaller subtopics, each with a set of questions for the auditor. Each question block has a set of criteria based on the expectations of the inspection agencies and have been discussed in advance with the relevant industrial bodies.

In recent years, the Belgian Competent Authority has made efficiency a priority in the design of its inspection instruments. Inspection agencies have decreasing resources, often less time, and an increasing number of companies to inspect. Therefore, every question asked in an audit instrument must count in the sense of demonstrable relevance. All audit questions in the inspection instrument are of the closed type: the answer is 'yes' or 'no' and reflects the presence or absence of certain objects in the company's system. For the purposes of this special study, these objects or items comprise the elements of a company's investigation system.

The questions are focused on objects or items established as essential and necessary for an effective and efficient investigation system. Each question addresses an enforceable requirement, justified with reference to:

- Legal compliance;
- Official Standards;
- Codes of good practice[34]; and,
- Accepted (and necessary) risk analysis measures.

Furthermore, the questions must be capable of producing answers that can be verified by the company's documents, standard operating procedures, investigation reports, or by interviews. The necessity and verifiability of the questions are critical qualities for regulatory inspections. Because each item is established as necessary, its absence from a company's system can be considered as a deficiency and registered as a shortcoming. This also allows the inspection agencies to enforce improvements.

The seventy audits (each asking 53 questions) produced a total of 537 shortcomings. To see patterns of weaknesses in the companies' investigation systems, the results from the sample of seventy companies were expressed as frequencies. For each audit question, the maximum possible frequency is 70, meaning that 100% of the companies had this shortcoming in their system. The higher the frequency of observed shortcomings per question, the greater the significance of this item as a weakness in the investigation systems of the Belgian Seveso companies in general.

When ranking the frequency of all shortcomings, the observed result showed that some of the questions with the highest frequencies were directly related to organisational factors:

- "Were the underlying organisational causes identified?" was the question associated with the largest number of shortcomings (highest frequency). This shortcoming was registered for 25 companies out of the total of 70, or 36% of the audits.
- Another question high in the ranking (top 5) and germane to this chapter was: "Does the general instruction specify an investigation technique that is explicitly focused on not only investigating the immediate causes but also the underlying organisational causes?" This shortcoming was registered for 19 companies, or 27% of the audits.

The upshot of this finding is that a sizeable proportion of companies seek to explain accidents and incidents without examining the organisational conditions that may be undermining how they manage the major risks created by their operations.

An example from Institut de Radioprotection et de sûreté nucléaire (IRSN)

As mentioned previously and in the previous example, one strategy to enhance prevention of accidents and their foresight relies on a normative strategy which references organisational factors. Another one relies on a more dialectical route within organisational diagnosis. For instance, and referring to the previous

---

[33] *In Belgium, two-thirds of Seveso companies belong to the following sectors: Oil and Gas; Chemical, Petrochemical, and Pharmaceutical manufacturing, and; Distribution and Warehousing of dangerous goods. The remaining third are dispersed across several smaller sectors.*

[34] *Codes of good practices are practices that are considered by one or more industrial sector(s) as practices who should be used as they have shown by multiple experiences to have a proven reduction in risk.*

example, when the company does not investigate root causes, the judgment is automatic. But for companies who do so, how can we rank the most and less performing and how can we judge if their practices and outcomes are good enough or bad? How can we make a judgment of the quality of analysis, the depth of organisational analysis and the relevance of the organisational factors addressed? This kind of approach is necessary as well to improve safety management but requires more data collection and collective expert judgment. The following example aims at illustrating the approach and some challenges.

Institut de Radioprotection et de Sûreté Nucléaire (IRSN) is the technical support organisation to the French nuclear safety authority (ASN). IRSN experts are in charge to conduct safety assessment on engineering provisions but also in-depth organisational diagnosis on safety management effectiveness of French nuclear power plants all operated by Electricité de France.

For our example here, the main scope of the safety assessment conducted was on maintenance activities performed during the 50 outages per year for the 58 nuclear reactors in France. Between 3,000 to 15,000 maintenance activities are performed per outage, involving several hundreds of workers over a period of one to six months. Most workers are subcontractors. So, a first challenge is related to the scale and complexity of the system: a nuclear reactor fleet of 58 reactors operated in 19 plants involving around 30,000 employees (including central engineering divisions) and 20,000 subcontractors employed by 400 companies.

A second challenge is related to the definition of the scale, scope and focus of an "open" audit or an organisational diagnosis. This can become especially challenging if the approach combines formal and informal data collection, interpretation of evidence of vulnerabilities or reliability/resilience, and debates about the necessity to implement preventive measures.

A multidisciplinary team of ten IRSN experts in human and organisational factors, safety and radiation protection engineering, conducted the safety assessment. Its goal was to assess the risk management efficiency in the 'daily' 'normal' functioning. In other words, it focuses on real practices, not on paper, nor it is rule or compliance based. It relied on an in-depth investigation over 2.5 years, 150 interviews, and 70 days spent observing working situations during three outages on three NPPs. Data collected is more or less subjective and therefore an objectification process aims at establishing evidence, facts and findings. It also relied on an in-depth review of documents running into thousands of pages from

several hundred of documents (e.g. safety procedures of the nuclear operator, reportable events analysis, inspections findings).

Six months were necessary to prepare the diagnosis, its scope and framework of analysis and the strategy for data collection. The preliminary analysis implied reviewing procedures to understand the safety management policy, structure, provisions implemented by the operator in a complex system. A determining factor to escape 'cognitive capture' was to benchmark across other strategies and provisions implemented by foreign nuclear operators. It required identifying the key organisational factors to be investigated. Investigating all organisational factors is not possible in one diagnosis for such a complex system, so the idea is to justify the selection of the most relevant organisational factors, based on major safety issues, such as the ones raised by organisational changes, or the vulnerabilities found in event analyses, known former vulnerabilities and new provisions dedicated to improve safety management. Five key organisational factors were selected (Dechy et al., 2016, 2018): organisational changes due to a new program of multiple changes; human resource management (in quantity and quality in a context of a wave of retirements); decision-making challenges within a complex organisation with multiple interfaces between people and processes of which subcontracting was a particular topic; and learning as efforts to improve the processes were undertaken. A transversal perspective, related to the historic dimension with the picture of a previous organisational diagnosis conducted five years before by IRSN (Rousseau, 2008) helped to address if safety management was improving or not.

All the key organisational factors selected were related to a 'pathogenic organisational factor' (Dien et al., 2004, 2012) though this was not the main selection criterion. Indeed, several other background knowledge references were used from human and social sciences, and good practices seen abroad. This 'knowledge of accidents' (Dechy et al., 2016, 2018) that contains pathogenic organisational factors helped to raise assumptions and support interpretations. IRSN experts were able to recognise echoes of accidents: a programme of multiple changes to improve performances (production, safety,…) echoed the 'torrents of change' at NASA before the Columbia accident; 'inversion of burden of proof' deviations at NASA that contributed to both Challenger and Columbia space shuttles accidents; and also a drift pattern of erosion of defence-in-depth, echoing the Swiss Cheese model (Reason, 1997). Collective expert judgment was produced to consider if the organisational weaknesses were serious and would need to

implement strengthened prevention measures, or if the safety management provisions in place were robust enough based on evidence of their efficiency. IRSN made fifteen recommendations, whose relevance and efficiency were assessed and challenged within a contradictory debate with the nuclear operator and thirty experts from the advisory committees[35] to the nuclear safety authority, before being translated by the safety authority in a new regulation to be enforced.

8.3.4 Foresight of Future Risks for Proactive Management of Risks as used in an organisational diagnosis

After hindsight and insight, the third temporal perspective where organisational factors could (and should) be employed to enhance proactive risk management is related to the future. For this chapter on employing organisational factors for FiS (Foresight in Safety) the act of gaining FiS is translated as getting the knowledge about how accidents in future could happen due to organisational causes or could become more likely due to new threats. In this perspective, the organisational dysfunctionalities to investigate, recognise and assess are not only the past ones nor only the current ones, but some that could occur in the near, mid-term and longer-term future. In other words, the goal is to implement an approach like risk analysis but related to plausible future threats to organisational safety. As a consequence, it requires developing all measures to counter those organisational causes and threats and to reinforce or to seize opportunities to implement new reliability/resilience factors to prevent these accidents from occurring, reoccurring or to decrease their likelihood by adding safety margins.

An example from Institut de Radioprotection et de sûreté nucléaire (IRSN)

A practical experience we can refer to is related to the organisational diagnosis performed by IRSN and described in previous pages. As mentioned earlier, the historical perspective of organisational analysis (Dien et al., 2006) was useful to investigate if in the current situation, one could notice safety improvements or degradations compared to the past diagnosis performed a few years previously (Rousseau et al., 2008). The historical perspective integrated also the future of the organisation towards potential forthcoming threats.

Indeed, the rationale was that IRSN experts had to consider if the dysfunctionalities found so far and the countermeasures to increase risk

management robustness were enough to cope with new threats forthcoming in the next few years. The main threats that were recognised at that time in 2013 were the lasting effects of the wave of retirements, ageing of the equipment especially because the nuclear operator was aiming for operating the nuclear reactors over forty years of operation (which was their design assumptions), which implied to increase up to 50% the workload in some maintenance domains and refurbish some critical equipment.

However, in 2012, IRSN experts observed a vicious circle (delays in outages that shorten time and resources available for learning post outages and the preparation of next outage; in such a case, it would generate new delays in outages). This drift was considered as a clear EWS by IRSN experts and they recommended to the nuclear operator some measures to counter it. This kind of safety degradation that is theorised (Dien, 2006) within accident models (such as the 'incubation period' (Turner, 1978), "latent failures" (Reason, 1990), and EWS that are not recognised or treated (Vaughan, 1996; Llory, 1996)) is not familiar to nuclear operators who are culturally educated with the so-called mantra of "continuous improvement" as a natural outcome of quality approaches and changes, which in itself is a fallacy (Dechy et al., 2011; Rousseau et al., 2016). The company for instance decided to reduce the maintenance workload in order to better manage the maintenance activities during outages of the next year and reduce therefore the vicious circle.

The company was also advised to reduce the frequency of changes so as to enable their implementation and ownership by an overloaded workforce that was coincidentally stressed by a wave of retirements and a heavier workload in maintenance work. This should enable better consideration of the impact of changes and especially the risks related to interactions of changes which remained under-investigated so far. In short, the company was invited to reconsider the overall strategy of changes which it finally did by delaying some changes and by giving more subsidiarity to local nuclear power plants than to central engineering and management departments of the nuclear fleet.

Last, the company was invited to consider the concept of organisational resilience and the need to diagnose and reinforce its resilience to potential new troubles.

---

[35] *Advisory committees (Groupes permanent d'experts), http://www.french-nuclear-safety.fr/ASN/Technical-support/The-Advisory-Committees*

8.4 Key elements of an OF Framework for guiding and questioning Foresight in Safety

The last suggestions in the previous example fit with our proposal to employ the positive sides of organisational factors to lever management of safety for the future.

After explaining the reason why organisational factors are a key lever for prevention and specifically for foresight in safety and providing examples in the way they were employed in inspection and auditing contexts conducting organisational diagnosis, this part addresses the practical challenges of employing organisational factors to prevent accidents. It outlines some guidelines for enquiring into the organisational factors of FiS. It also uses lessons learned from the nuclear and chemical industries that can be applied to other industrial sectors.

8.4.1 Background and foundations for elaborating a framework

This subsection aims to develop additional definitions about organisational factors in relation to safety, either positive and negative, by illustrating some of them in multiple literature sources, and their combined outcome on system states.

Diagnosing the dynamic state of functioning with opposite forces?

As explained (figure 1), quality and efficiency in organisational functioning has a great impact on safety, whether its outcome is positive or negative. Our diagnostic challenge is to anticipate risks and enhance safety.

If we look at the impact on safety, the functioning of an organisation can be placed on a continuum, moving from time to time and oscillating between different states. Of course, an organisation has different parts, and these may differ, but we consider that the functioning of the whole is dependent on the weakest part. At least, in a simple manner, we can identify three specific organisational situations that lead to three different safety states: **dysfunctional**, **normal**, and **resilient**.

The impact on safety of a **dysfunctional organisation** ranges from negative to very negative. Chronically dysfunctional organisations are sometimes called 'pathological' (Reason 1990, 1997; Westrum, 1992; Dien et al., 2012). In these organisations, the degradation of safety is severe enough to be detected with relative ease by several actors or processes (audit, investigation, as in Texas City 2005 accident). However, many EWSs and alerts are not treated accordingly. Part of the culture within the dysfunctional organisation does not want to know and

discourages 'bad news'. These are the organisations who 'shoot their messengers', punish whistleblowers, blame individuals for failures and discourage new ideas (Westrum, 1992). In such organisations, the likelihood of a system accident grows as negative organisational factors accumulate. Dynamically, the effect may be seen as a 'system drift' (e.g. Dekker, 2011) accompanied by normalisation of deviance (Vaughan, 1996). In the longer term, the likelihood of accidents may become very high as the system becomes critically vulnerable, and its ability for adaptive change becomes embrittled (Woods, 2009). In such an environment, an event can trigger cascading effects because several lines of defences are already weak or lack safety margins. Foresight is low and even reactive measures are lacking. For the present authors, the Texas City accident is a case that demonstrates the effects on safety in a severely dysfunctional, pathological organisation.

A **normally functioning organisation** has adequate safety in normal conditions. This would be the minimum expectation of a responsible, law-abiding management. Its approach to safety is characterised by adequate preventive and protective measures; a reactive and proactive attitude towards near misses (e.g. what if?); regular auditing, and; looking for root causes when investigating and inspecting. It is already doing more than treating safety in a bureaucratic manner (Dekker, 2014) which would be limited to listening to messengers if they arrive, and responding with local repairs only (Westrum, 1992). It does not mean there would be no incidents or local accidents, but their impact would be limited, as some safety margins and barriers would block their escalation into a major accident. The likelihood of a system accident remains low, and although some limited drift may occur, it can be recovered in time if action is focussed. Unlike the dysfunctional organisation with its eyes closed, foresight in the normally functioning organisation is practiced with conventional tools. Overall, we could say that the normally functioning organisation is a robust system: it can withstand deviations and anomalies to degree, especially if these stay within the design basis.

A **resilient organisation**, in contrast, can withstand and even repel unforeseen events and disruptions and still stay safe. It is highly reliable (e.g. Roberts, 1990) and resilient in the sense set out by Hollnagel et al. (2006). The resilient organisation is highly proactive (sometimes called generative, Westrum 1992) about tackling residual risks. This proactive behaviour is characterised by challenging and reinforcing their defence-in-depth, conducting stress-tests on beyond design basis events (e.g. in the nuclear industry after Fukushima), learning from their own events and opportunistically from others, and challenging basic

assumptions and questioning the status quo. Bad news is welcome in the resilient organisation, in fact it searches for divergent opinions (messengers are trained and rewarded, Westrum, 1992). Its approach to foresight goes beyond the use of conventional tools. It involves outsiders, sponsors 'red teams', investigates root causes—not just of near misses but even of EWSs. The resilient organisation conducts organisational reforms and invests in additional safety margins without regulatory requests.

Table 3. Three organisational situations leading to three different safety states.

Organisational situations	Dysfunctional	Normal (even robust)	Resilient
State	Unsafe to very unsafe	Safe within design basis in normal conditions	Safe, even when under stress beyond design basis
Trend	Degrading safety Increasing vulnerability Organisational dysfunctionalities can become pathological	First target is maintaining safety. There is a positive balance between positive and negative forces with the safety margins that remain	Safety is maintained by adding safety margins
Descriptors	Signs of safety degradation are recognised by several actors within or outside the organisation. Alerts are not treated adequately. Messengers are "shot". Blame culture. Local repairs, only.	Reactive and proactive safety management system functions are performed with energy in more than a merely bureaucratic way. Addresses root causes.	No self-satisfaction; challenges assumptions and status quo (stress tests their defence in-depth). Uses unconventional methods to see and think differently (e.g. 'red teams'). EWSs are treated at the "right" level and may lead to organisational reforms. Adds safety margins without regulatory requests.

8.4.2 Which are the relevant organisational factors to investigate to enhance safety and foresight in safety?

Investigating, for prevention purposes, the potential or actual effects of organisational factors on the system requires a general mapping of the relevant organisational factors that could be addressed.

Review of Lists of Organisational Factors

Since Turner (1978) and Reason (1990, 1997) broke the ground, a lot has been said and written about organisational accidents, reliability, resilience and safety. The literature contains several lists of organisational factors that are claimed to be relevant. Each of these lists has its own logic and arises from its author's theoretical tradition (safety, psychology, sociology, management sciences or economics). In this paper, we call them organisational factors (OFs) as defined in part 8.2.5. However, other authors have used terms such as pathogenic and resilient organisational factors, dysfunctional factors, latent causes, and so forth.

We reviewed about 30 of those lists, but there are more. This work is still in progress, but readers are invited to regularly update their lists with insights from accidents and new researches. The lists we reviewed came from different sources, researchers and safety analyst but also safety organisations including: the US Center for Chemical Process Safety, who published a book on the subject; the Energy Institute; inspection agencies, such as the Health and Safety Executive in United Kingdom; and many individual authors writing about safety, or about organisational management.

The review found that the lists exist in isolation and do not, except in a very few cases, refer to each other or have common links. We see this as an important missing characteristic; it is one of the reasons why we wish to develop a guiding framework.

Moreover, many lists included either positive OFs, negative OFs, or both, without distinction. This can sometimes mislead readers. We took from this that OFs should be labelled or stated in a way that makes it absolutely clear whether it purports to have a positive or a negative impact on the safety of an organisation. In future work, we want to fulfil this requirement by providing an assessment question for each OF in the framework.

In the next two subsections, we provide illustrative lists of purely negative OFs, and purely positive OFs.

An Illustrative list of negative organisational factors

Here are three examples of entries on a list of organisational factors that are purely negative in their effects on safety. Those below are presented as illustrations; many other negative OFs could be included.

- **Production pressures.** These result in behaviours and injunctions aimed at overriding or voluntarily ignoring certain dimensions of safety in order to favour short-term technical or economic performance. Production pressures arise when the production culture—a set of knowledge, know-how, etc. contributing to technical or economic results—is no longer counterbalanced by the safety culture. Often in a competitive environment, the strategy and priorities set-up by top management initiate or reinforce those production pressures. A first difficulty of detection comes from the confusion between the culture of production and the pressures of production, that is to say that the pressures of production can be assimilated within a dimension of the culture of production.

- **Weakness of operational feedback.** The feedback (lesson learning) process comprises: the detection of malfunctions, the collection of these data, the analysis and synthesis of the causes of these malfunctions, the definition of corrective measures, the implementation of these measures, the evaluation of the effectiveness of the measurements, and the memorisation of the treatment. The implementation of corrective measures aims to prevent the occurrence of new incidents and accidents. This process is iterative and dynamic, and in this sense the feedback is 'alive'. The lesson becomes 'unlearned' when the organisation has difficulty in recalling the experience. This difficulty appears when the feedback process is either weak or not at all supported in the organisation, or when the associated resources are insufficient, or when a step is (systematically) absent or deficient.

- **Weakness of control bodies.** Control bodies are the entities responsible for verifying compliance with duties for safety. These duties, owed by the operator of the socio-technical system at risk, arise from various obligations: legal, regulatory, contractual, procedural, social, moral, and so forth. The control bodies reflect these different classes of obligation. They include those attached to the installation (local safety departments, for example), those at a "corporate" level of the company in charge of the installation, and those outside the company (safety authorities for example). The possible weaknesses of control bodies refer to the weaknesses of their interventions and actions, meaning that they do not play the role of counterweight and counter-powers as they are supposed to.

Several other negative organisational factors have been identified by investigators and researchers. Among the most important recalled here, we could add the lack of re-examination of design assumptions, flaws in human resource management and the organisational complexity including subcontracting (e.g. Dien et al., 2013). More examples are e.g. the POFs, called "pathogenic organisational factors", as described in Pierlot et al. (2006).

An Illustrative of a list of positive organisational factors

In general, positive OFs are those that either maintain or improve the level of reliability of the system or its robustness and resilience, and therefore have a positive impact on its safety performance. Note, however, that compiling a full list of positive OFs may be complicated. For example, the benefits to safety that are associated with positive OFs might in some cases be indirect or may depend on circumstances that change. Furthermore, some organisational provisions can improve both production performance and safety but can oppose as well. There are also conceptual differences between reliability and safety (see Nancy Leveson, 1995 and 2004; Llory and Dien, 2006).

The items below are presented as illustrations; many other positive OFs could be included. These particular items arise from studies of highly reliable organisations (Weick and Sutcliffe, 2001).

- **Efficient treatment of malfunctioning.** A major negative event is never a standalone situation. It does not occur by chance. It is (almost) always preceded by 'little' events which are all early warning signs—symptoms of the deterioration of the safety level. So, if every event is detected, analysed in terms of its generic aspects (i.e. considering what could have been worse) and, if the corrective measures designed cover also generic aspects, then the organisation increases the likelihood of avoiding a more serious event. Too often, 'little' events are treated as 'here and now', meaning the only corrective measures defined are those that will only avoid reoccurrence of this specific event. To treat an event as unique and wholly exceptional is to deny its significance. It is fair to say that an

organisation with 'a lot' of small events treated well is safer than an organisation with no events. As well as stimulating the search for improvements, feedback keeps an organisation watchful for danger. One could say that when this OF is present, people in the organisation are mindful of failure rather than blinded by success.

- **Real operations oriented.** Every activity and organisation, and especially those with large safety risks, are governed by rules, procedures and regulations. In general, these are attempts to define and describe boundaries of operations. Unfortunately, makers of rules cannot foresee all the possibilities of real life, which contains unexpected events and unforeseen situations that operators must cope with on the front line. Furthermore, these situations or events can arise in any of several domains of activity: operations, maintenance, or training. Because of this complexity, a feature of maintaining an appropriate level of safety is a great reliance on those closest to the process. 'Great reliance' does not mean 'blindly' relying on everything done by front line workers. It means that regulators, designers and managers must pay attention to everything done beyond procedures and to check how right (or wrong) it was.

- **Deference to expertise.** Some situations, mainly hazardous situations (e.g. crises, incidents, and accidents) demand that decisions are made quickly. In those cases, especially if complex, the real-time processes of decision making cannot be based on the organisational hierarchy. Rather, decisions are made by the people locally in charge of operations, based on their knowledge and skills. An organisation needs to have prepared for this change in how decisions are made. Amongst other things, hierarchical leaders must be ready to allow these knowledgeable, skilled people to speak freely. Moreover, by virtue of their knowledge and skills, these people may also be able to improve decision-making in everyday, non-emergency, circumstances. Deference to expertise is the tendency to delegate decision making to those who have the most expertise, irrespective of their position or hierarchical status.

- **Open minded to debate.** Steep hierarchies in organisations often lead to bureaucratic management (Dekker, 2014). This situation favours the emergence of a single, not to say over simple, official 'view of the world'. Yet, organisations are generally heterogeneous entities, and not monolithic wholes (Dien, 2014). The usual situation, especially on issues of process safety, is for the coexistence of several opinions and views of a situation. Since safety is not only a matter of rule compliance, but is also a matter of debate, every opinion must be expressed, irrespective of the hierarchical position it comes from. Diverse and dissenting voices must be taken into account, although not necessarily agreed with. They must be listened-to without a defensive attitude. The ability to give room to debates (about safety) and welcome 'bad news' is a positive organisational factor for safety. This is notably true in crisis situations, where as well as pre-planned emergency actions, some time will be spent sharing information and interpretations—sometimes through sensemaking confrontations—to inform decisions about what needs to be done.

- **Reluctance to simplify**. Industrial facilities are usually manifest as complex systems. In order to be able to handle the whole process, organisations are tempted to simplify interactions between some subsystems and to exclude some others from serious study. An example of this is modelling, which even when detailed still represents a simplification of an even more complex reality. By putting aside what they consider to be outliers, organisations take risks. Treating outliers in this way creates blind spots over the corresponding zones of the process, so creating the scope for unanticipated and unwanted situations to occur. So, simplification creates a wrong picture of the real situation. The 'devil' as the saying has it, 'is in the details'. To put it another way, 'situational awareness' demands a questioning attitude, one which avoids easy, simplistic explanations and shortcuts in assumptions. Rather than those that simplify at every turn, it is organisations willing to grapple with the complexity of their processes that stay able to avoid major surprises.

These positive factors, it is contended, act in combination. The more of these and other positive factors are present within an organisation, the better safety is positively ensured with safety margins.

Other positive organisational factors are those which allow facilities to remain resilient. As proposed by the resilience engineering school of thought (Hollnagel et al., 2006), the four resilience features are: the ability to respond, the ability to monitor, the ability to learn, and the ability to anticipate. Hollnagel (2009) proposed a matrix which provides a measurement of the resilience level of an organisation according to the score it obtains for each ability.

8.5 Bridging the operational gap: from current 'part one' to future 'part two'

8.5.1 Synthesis of 'part one': A road map towards an OF's framework for guiding and questioning Foresight in Safety

By looking at the literature on major accidents and several accident theories and at our own investigating experiences (in nuclear and chemical industries) we have shown (section 8.2) that it has been known for more than twenty years that OFs affect risk prevention heavily. However, they remain underused by industry and regulators.

In addition to the general definitions of technical and human influence factors that are widely accepted in the safety community, we have proposed three levels of organisational factors: management system failures, organisational dysfunctionalities, and regulation dysfunctionalities.

Exploring our past experiences together with our ideas about hindsight, insight and foresight permitted us (section 8.3) to link different activities (such as investigations, audits, and organisational diagnoses) to different temporal phases (past/hindsight: investigating accidents and incidents; present/insight: auditing and diagnosis; future/foresight: proactive management). The practical examples in this section gave us some ideas about how to improve foresight by combining these different 'sights'.

After research of the literature, we were able to define the three organisational situations leading to three safety states (dysfunctional, normal, and resilient defined in table 3 in section 8.4).

To get a view on which organisational factors are relevant, we reviewed several lists of OFs. Although in the safety literature, there exist several lists of OFs[36], they are scattered and many of them have a limited scope. Moreover, a global and coherent view is lacking and support for a thorough coherent organisational diagnosis remains limited.

The result is that organisational factors remain vague, opaque and, relative to their importance, barely visible as latent causes of accidents or levers for risk prevention. Therefore, there is a need for a framework that enables practitioners and researchers to more readily use these OF constructs in the search for weaknesses/threats and strengths/opportunities in organisations. The use of the framework would be not only backward looking (hindsight) and present (insight) but also for the future (foresight).

8.5.2 Future work for 'part two'

This section describes the work that we plan to do in the future.[37] We identify two tools to be developed.

As result of the above (section 8.5.1), the first step in our plan is to develop a framework that practitioners can use as a tool to help them find significant organisational safety weaknesses and strengths.

The value of the framework to the practitioner will be to guide a systematic search for relevant OFs. We see it as assisting, not replacing, practitioners' existing fieldwork processes. For example, if a practitioner finds that an OF in the framework to be relevant in a particular audit or investigation, they would gather more data using their existing skills and practices to evaluate the relevance and potency of that OF on the organisation's safety.

The framework will permit people acting as investigator (of accidents), auditor, organisational analyst to look from the starting OF to the surrounding, (neighbouring or adjacent) organisational factors so those factors can be investigated and assessed in turn for their impact on safety. Consequently, the "spreading" of the negative or positive impact inside the organisational framework can be made more obvious. In this way, new foresight is created.

Although one must be careful not to overcomplicate the framework—usability is crucial— the possible connections and plausible links between organisational factors could be mapped and will be part of the interpretative framework. For example, too many organisational changes in a short notice might be linked to production pressures or a misperception of the effects of changes, for instance on roles and responsibilities.

On the other side when practitioners find plenty of evidence for a particular OF, they will be very tempted to close too soon the analysis. This is contrary to the

---

[36] *We studied 26 lists from authors of different competences and skills (safety, engineering, sociology, psychology and management)*

[37] *Any volunteering is welcome! Contact frank.verschueren@werk.belgie.be or nicolas.dechy@irsn.fr*

principle of thorough inspection and investigation. Our aim is that the future framework will support the questioning and prevent the premature stop of an analysis.

The second step in 'part two' will be a set of "assessment questions". Ideally, each component of the framework will have questions to help the practitioner assess the quality of impact (positive or negative) and its level (weak or strong impact). However, although safety issues evolve in general ways, (e.g. ageing, digital transformation, etc.) they manifest uniquely in every specific organisation. Mindful of this interplay between general and specific, when applying them, practitioners will always need to adapt the framework and the set of questions.

Our approach implies to look for the effects, observable outcomes of the combination of OFs in some specific normal functioning situations of the system, activities or events inside the organisation.

In contrast with the classic basic audit which delivers an instantaneous picture of the present situation and performance, we see an approach that is more extended over time, and dynamic in what it focuses on and the temporal perspective taken.

If time is a limiting factor, this kind of organisational diagnosis (audit, inspection, investigation) can be performed on specific topics rather than the whole management of safety. To avoid staying at the surface of the organisation and the 'speeches of the front stage', we recommend a tighter focus: selecting a few sub-topics to be questioned, such as those revealed by previous audits or investigations, or areas that the organisation is changing.

By assessing organisational factors, which we see as characteristics of the organisation that are critical for safety, we can identify weaknesses. Some weaknesses will require urgent remedy, but others allow a more gradual approach to improvement. Similarly, by referring to positive factors, we can detect areas for consolidation.

To summarise, we propose these as relevant objects for study and intervention:

- Historical vulnerabilities
  => to be found especially in the past

- EWS, symptoms of dysfunctionality, drift and changes
  => to be searched for in the past and in the present situation;

- New threats to consider and opportunities for improvement
  => to be looked for in future potential situations.

As we aim to support practitioners to find these objects and to assist them in verifying and proving that they identified the correct former vulnerabilities, the right present or past EWS and the relevant future threats (we cannot imagine now), we propose development of:

1. A framework to help practitioners detect and characterise if the set of organisational factors (the factors that are essential characteristics of the safe functioning of the organisation) are dysfunctional, normal or resilient and are producing observable effects (symptoms and EWS);

2. A set of questions that practitioners can ask as part of their exploration of organisational factors in a given investigation, audit or assessment;

3. An assessment method to foresee the future effects—whether positive, negative or neutral — of organisational factors on safety in a given organisational setting. This will be a kind of SWOT analysis (Strengths, Weaknesses, Opportunities and Threats);

4. A protocol for collective and debated judgment of the overall safety state: stable, improving or degrading.

All this will be further developed and finalised in our future part two, where in addition we would like to see development of the following capacities:

- Guiding Hindsight, Insight but specially Foresight in Safety by the ("guided") search of other plausible EWS starting from a detected and confirmed EWS;

- Enhancing Insight and Oversight on the safety performance of an organisation;

- Questioning Foresight in Safety by assessing the impact of present and future decisions and behaviours of all levels, but in particular those of the Board and senior management.

8.6　Acknowledgement

8.7　References

Columbia Accident Investigation Board (2003), Report Volume 1, National Aeronautics and Space Administration, the Government Printing Office.

CSB - U.S. Chemical Safety and Hazard Investigation Board (2007), "Investigation Report, Refinery Explosion and Fire, BP – Texas City, Texas, March 23, 2005", Report N°2005-04-I-TX.

Cullen, W. D. [Lord] (2000), The Ladbroke Grove Rail Inquiry, Part 1 & Part 2 Reports, HSE Books, Her Majesty's Stationery Office, Norwich, 2000. [Part 2: 2001].

Dechy, N., Dien, Y., Llory M. (2010), Pour une culture des accidents au service de la sécurité industrielle, Congrès lm17 de l'IMdR, La Rochelle, 5-7 Octobre

Dechy N., Rousseau J.-M., Llory M. (2011), Are organizational audits of safety that different from organizational investigation of accidents, proceedings of the ESREL 2011 conference in Troyes, France, "Advances in Safety, Reliability and Risk Management – edited Bérenguer, Grall, & Guedes Soares, Taylor & Francis Group, London.

Dechy, N., Dien, Y., Funnemark, E., Roed-Larsen, S., Stoop, J., Valvisto, T. & Vetere Arellano, A.-L., on behalf of ESReDA Accident Investigation Working Group. (2012). Results and lessons learned from the ESReDA's Accident Investigation Working Group, Safety Science journal, 50.

Dechy N., Rousseau J.-M., Dien Y., Llory M. Montmayeul R., (2016) Learning lessons from TMI to Fukushima and other industrial accidents: keys for assessing safety management practices, proceedings of IAEA Conference 30 Years of Safety Culture

Dechy, N., Rousseau, J.-M., Largier, A., Tillement, S., Hayes, J., Journé B., (2018), Using the 'Knowledge of Accidents' in Normal Operation: a Case Study on the Framing of Organisational Diagnosis of Safety Management, Proceedings of the 55th ESReDA seminar on - Accident Investigation and Learning to Improve Safety Management in Complex Systems: Remaining Challenges, Romania, 9 – 10 October, AGIFER

Dien, Y. (2006), Les facteurs organisationnels des accidents industriels, In : L. Magne & D. Vasseur (Coordonnateurs), Risques industriels – Complexité, incertitude et décision : une approche interdisciplinaire, pp. 133-174, Éditons TED & DOC, Lavoisier

Dien, Y. (2014), "Les signaux faibles à l'aune des lanceurs d'alerte", Congrès Lambda Mu19, Dijon

Dien, Y. et Llory, M. (2004), Facteurs organisationnels des accidents industriels : revue bibliographique, Rapport EDF R&D HT-52/04/003/A

Dien, Y., Llory, M. & Montmayeul, R. (2004). Organisational accidents investigation: methodology and lessons learned, Journal of Hazardous Materials, 111 (1-3), pp 147-153.

Dien, Y. Dechy, N., Guillaume, E. (2012) "Accident investigation: From searching direct causes to finding in-depth causes – Problem of analysis or/and of analyst?" Safety Science, Volume 50, Issue 6, pp 1398-1407

Dien Y., Dechy N., Stoop J. (2012) Perspective regarding industrial event investigations, Safety Science journal 50 (issue 6): pages 1377-1379.

Dekker S. (2006), The field guide to understanding Human Error, Ashgate

Dekker S. (2011) " Drift into Failure: From Hunting Broken Components to Understanding Complex Systems", Farnham: Ashgate Publishing Co.

Dekker S. (2014)"The bureaucratization of safety" Safety Science 70:348–357

Energy Institute (2008), Guidance on investigating and analysing human and organisational factors aspects of incidents and accidents, London

ESReDA (2003), Eds, Valvisto T., Harms-Ringdahl L., Kirchsteiger C., Røed-Larsen S., Accident Investigation Practices: Results from a European Inquiry. DNV safety series.

ESReDA (2005), editors Roed-Larsen, S. R., Stoop, J., Funnemark, E., *Shaping public safety investigation of accidents in Europe*, DNV publishing, ISBN 82 5150304 3

ESReDA (2009), Guideline for Safety Investigation of Accidents, ESReDA Working Group on Accident Investigation, (2009), available at www.esreda.org

ESReDA (2015) Case study analysis on dynamic learning from accidents - The ESReDA Cube, a method and metaphor for exploring a learning space for safety, Van der Vorm and al., available at www.esreda.org

ESReDA (2015), Barriers to learn, edited by Marsden E. and the ESReDA Project Group on Dynamic Learning. available at www.esreda.org.

ESReDA (2015), Eds., ESReDA project group on dynamic learning as the follow-up from accident investigation "Guidelines for preparing a training toolkit on event investigation and dynamic learning", available www.esreda.org

Groeneweg, J., Van Schaardenburgh-Verhoeve, K.N.R., Corver, S.C., Lancioni, G.E., 2010. Widening the scope of accident investigations. In: Society of Petroleum Engineers Conference on Health, Safety and Environment in Oil and Gas Exploration and Production held in Rio de Janeiro, Brazil, 12–14 April 2010. SPE, pp. 127–157.

Hayes J., Hopkins A., (2014) Nightmare pipeline failures, Fantasy planning, black swans and integrity management, Edited by Wolters Kluwer, CCH,

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

Hollnagel, E. (2009) RAG – Resilience Analysis Grid, Technical Document prepared by the Industrial Safety Chair, https://erikhollnagel.com/onewebmedia/RAG_introduction.pdf, retrieved on August 3, 2020.

Hopkins, A. (2003), Lessons from Longford. The Esso Gas Plant Explosion, 7th edition, CCH Australia Limited.

Hopkins, A., Maslen, S., (2015), Risky Rewards: How Company Bonuses Affect Safety, Edited by CRC Press, Taylor and Francis.

Johnson, W.G. (1973), The Management Oversight and Risk Tree, US Atomic Energy Commission SAN821-2. US Government Printing Office.

Leveson, Nancy G. (1995) Safeware, Boston: Addison-Wesley Publishers, 1995.

Leveson, Nancy G. (2004) A new accident model for engineering safer systems, Safety Science, 42(2004) 237-270, Pergamon Press (Elsevier), 2003

Leveson N.G. (2002), A new accident model for engineering safer systems, MIT working paper.

Leveson, N. (2011), "Applying Systems Thinking to Analyze and Learn from Events" Safety Science 49(1): pp 55-6

Llory, M., "Accidents industriels : le coût du silence, Opérateurs privés de parole et cadres introuvables", Éditions L'Harmattan, 1996.

Llory, M., "L'accident de la centrale nucléaire de Three Mile Island", Éditions L'Harmattan, 1999.

Llory, M. & Dien, Y. (2006), Les systèmes sociotechniques à risques : Une nécessaire distinction entre fiabilité et sécurité, Partie 1 : Performances n°30, septembre – octobre, pp. 20-26, Partie 2 Performances n°31, novembre – décembre, pp.9-13, Partie 3 Performances n°32, janvier – février, pp. 20-26.

Llory, M., Dien, Y. (2010), L'analyse organisationnelle de la sûreté et de la sécurité des systèmes complexes à risques, Les Techniques de l'Ingénieur, TI AG 1577.

Llory, M.; Montmayeul, R. (2010) "L'accident et l'organisation", Éditions Préventique.

Llory, M. & Montmayeul, R. (2018). Comprendre l'accident, la pratique de l'analyse organisationnelle de sécurité, Editions L'Harmattan.

Merritt, C. W. (2007). Testimony of Carolyn W. Merritt, U.S. CSB, before the U.S. House of Representatives, Committee on Energy and Commerce, Subcommittee on Investigations and Oversight, May 16, 2007

Morin, E. (1977), La méthode, Editions Seuil.

Perrow, C. (1984) *Normal accidents, living with high risk-technologie*s, Princeton University Press.

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

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

Reason, J. (1990). Human Error, Cambridge University Press.

Reason, J. (1990) "The Contribution of Latent Human Failures to the Breakdown of Complex Systems". Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. pp1): 475–484

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

Roed-Larsen S., Valvisto T., Harms-Ringdahl L., Kirchsteiger C. (2004), Accident investigation practices in Europe — main responses from a recent study of accidents in industry and transport, Journal of Hazardous Materials, Volume 111, Issues 1–3, 26 July 2004, Pages 7–12

Rousseau, J.-M. (2008). Safety Management in a competitiveness context, EUROSAFE Forum Proceedings, 3-4 novembre, Paris, France,

Rousseau J-M., Largier A., (2008). Conduire un diagnostic organisationnel par la recherche de facteurs pathogènes, Techniques de l'Ingénieur, AG 1576

Sklet, S. (2003), Comparison of Some Selected Methods for Accident Investigation, Proc. of 24th ESReDA Seminar, Petten – May 12-13, 2003.

Turner, B., & Pidgeon, N., Man-Made Disasters, Second edition, Butterworth Heinemann [1rst edition: Turner, B. (1978), Wykeham Publications]. 1997

Vaughan, D. (1996), "The Challenger Launch Decision. Risky Technology, Culture, and Deviance at NASA", The Chicago University Press, Chicago

Vautier, J.-F., Dechy N., Coye de Brunélis, T., Hernandez, G., Launay, R., Moreno Alarcon, D., P. (2018), Benefits of systems thinking for a human and organisational factors approach to safety management, Journal of Environment System and Decision, 38:353-366

Verschueren F.[38] (2013), Presentation "Inspection of Investigations of Accidents and Incidents" at MJV Seminar on Learning of Incidents (11 – 13th of September 2013 Gothenburg, Sweden)

Verschueren F. (2015) Presentation "Learning from Experience" at Euro CCPS Den Haag Conference (18th of March 2015, Den Haag, Netherlands)

Verschueren F. (2015) Presentation "Workshop on Generic Causes or Organisational Factors" at TapRoot Global Summit (4th of June 2015, Las Vegas, USA)

Verschueren F. (2017) Presentation "Learning from Organisational Dysfunctionalities" at Energy Institute Conference on "Human Factors Application in Major Hazard Industries", (17 – 18th Oct 2017, London, United Kingdom).

Weick, K. & Sutcliffe, K. (2001), Managing the Unexpected – Assuring High Performance in an Age of Complexity, Jossey-Bass Publishers.

Westrum, R., (1992), Cultures with requisite imagination, in Wise, J., Hopkin, D., Stager, P. (eds), Verification and validation of complex systems: Human Factors Issues, Berlin, Springer Verlag, pp 401-16

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, Elsevier Science Ltd, pp 235-248.

Woods D. (2009) Escaping failures of foresight, Safety Science 47 (2009) 498–501

---

[38] All presentations are available on request: mail to frank.verschueren@werk.belgie.be

9   Safety foresight in asset management

Paulo Maia, Energias de Portugal (EDP) – Gestão da Produção de Energia, S.A., Portugal,
John Kingston, Nordwijk Risk Initiative Foundation (NRI Foundation), The Netherlands.

Executive summary

In recent years, *asset management* has been applied consistently as a structured discipline to several areas of economic activity, including infrastructure, industry, banking and insurance. Banking and insurance are mainly related to the financial sector. Infrastructure refers to energy networks (electricity, gas, district heating), water and sewage, roads, rail network and telecommunications, and the relevant industrial sectors include, amongst others, power generation (renewables, fossil and nuclear), chemical and petrochemical, and pulp and paper. However, to present this subject in a reasonable level of detail, only industrial assets will be considered here, with a special attention to the power generation industry.

There are many reasons why asset management has recently become an essential part of management activities and management science. Several examples can be cited, such as the ageing of industrial asset systems and its increasing integration; more stringent quality, safety and environmental requirements for the industry, imposed by regulators; greater awareness of risks among workers, managers and stakeholders; globalisation and fierce market competition; and pressure on asset managers for higher profitability and return on assets. Frequently, several of these features generate a combined effect, making it difficult to identify the specific contribution of each one to asset management.

In the literature, several definitions of asset management can be found. However, the definition included in the ISO standard for asset management released a few years ago (ISO 55001:2014), will be used here as a reference, although its meaning is quite broad. The standard defines asset management as a coordinated activity of an organisation to realise value from assets. In turn, 'asset' is defined as an item, thing or entity that has potential or actual value to an organisation. However, in the specific context of this paper, the term *asset management* has its main application to the industry, that is, asset management focussed on physical assets.

The main objective of this paper is to identify areas of asset management that can be used in safety foresight, to enable the detection process for systems/equipment deterioration signals or anomalies before a serious accident can occur. To accomplish this objective, it is necessary to limit the period of the lifecycle of an industrial asset to its intended use, that is, to the operational/production stage. However, this does not preclude taking necessary safety measures during design, or later in the construction, commissioning or even in the decommissioning stages.

The relevant activity during the operational/production stage is Operation and Maintenance (O&M), which is accomplished both by personnel and systems/equipment through a set of processes following established operating procedures. This is the reference period that will be considered in detail in this chapter, it being simultaneously the longest and the most relevant in the life cycle of an industrial asset.

In industry, although *internal agents*, including managers and technical staff, put asset management into practice every day, the role of *external agents* cannot be neglected. They influence the way industrial assets are managed, too. External agents include: regulators, certification bodies, insurance companies, technologists (high-tech manufacturers), O&M specialist companies or industry institutes, users' groups and sector associations. They help companies by identifying non-conformities, weaknesses, gaps and process deviations in the way asset systems are being managed. When agents find these indications, they act: issuing instructions and alerts, or making recommendations aimed at correction. And occasionally, recommendations issued by one external agent can even induce a synergistic effect on related issues in the domain of other agents. As the contribution of these external agents to foresight in safety is not usually mentioned in the literature, it will be the main topic addressed in this chapter. To better illustrate the subject, practical examples taking from the conventional power generation industry are given. In this context, nuclear power generation is ruled out of this analysis, as oversight of regulators is quite intense, even dominant in respect to other external agents, and major accidents occur very infrequently. As a result, it is quite difficult to find areas for improvement in nuclear power generation that would have broad application to other industry sectors.

9.1 Introduction

"I am prepared for the worst, but hope for the best."
Benjamin Disraeli (1804-1881)

The term 'asset management' has become quite popular in recent years and is currently used to address management issues in several areas of economic activity, including infrastructure, industry, banking and insurance. Although asset management activities have been performed since society started to use capital assets, recent changes in our way of living and business environment have required the adoption of a more structured management approach. The efforts made to obtain this more structured approach culminated with the publication in 2014 of the first edition of the ISO standard on asset management Standard 55001. The standard defines a set of requirements that once established, implemented and maintained, will ensure the fair asset management performance of an organisation, responding to the requirements and expectations of interested parties and ensuring value creation and maintenance.

Assets can be physical, financial, human or 'intangible'. However, to reduce the broad spectrum of assets to be addressed in this chapter, only physical and human assets from the industry sector will be considered. Physical assets include systems and equipment, the environment and the associated production processes. Special attention will be devoted to the conventional power generation industry, from which examples will be provided to further illustrate the statements and, hopefully, enabling replication to other industries, whenever applicable. Nuclear power generation is ruled out of this analysis, for several reasons, including the specificity of this industry and the intense oversight of regulators, mainly concerned with nuclear safety. As a result, it is quite difficult to find areas for improvement in nuclear power generation that would have broad application to other industry sectors.

The main objectives of this chapter are to (i) identify tools, practices and agents (internal and external) that can contribute to safety foresight in asset management within the power generation sector, and (ii) to describe how this can be achieved.

9.2 Systems and Equipment

In the power generation industry, as in other industries, physical assets can be divided into several categories and levels. Complementary to this, a coding system is usually adopted for asset management purposes, including operation and maintenance. At the top level, power plants are divided into units, of similar or different design, technology, installed capacity, etc. When a power plant is to be constructed within a specific investment project, typically between two and four units, similar units are considered. The division into units provides adequate operational flexibility to satisfy power grid needs and minimise the risk that all units might fail simultaneously. However, in recent years, as an effect of the European liberalised electricity market and the opportunity to use extra installed capacity, some reservoir hydro power plants have been subject to repowering projects. In these cases, the tendency is to construct units with a higher installed capacity, thereby taking advantage of the benefits of a higher electricity price in peak hours and, simultaneously, of the technological advances made in relation to the original units, which sometimes were built decades before.

The second level of physical assets are the systems and equipment; the third, the components; and fourth, the parts.

Having in place an efficient coding system is a key-element for operation and maintenance activities. It enables the asset owner to better manage systems and equipment failures, stocks of critical spare parts, and communication with manufacturers.

Reference to all these asset levels will be made, where appropriate.

9.2.1 Operation

9.2.1.1 Performance Requirements

Systems and equipment performance requirements are established through a set of standards, manufacturer operational instructions and emergency procedures that should be followed by the industrial asset personnel, including managers, supervisors and operators.

Any advanced industrial process is managed by a command and control system, usually known by the acronym SCADA (Supervisory Control and Data Acquisition). This system includes safety instrumentation and process control systems that are

run automatically, under operator supervision. In recent years, major advances have been made in those systems. Due to this, command and control system obsolescence cannot be ignored, as this system is the 'heart' of the plant. Based on experience, a command and control system becomes obsolete approximately in 20 years. This is critical, especially for hydro power plants, as these power plants may reach 100 years, without major improvements in the remaining systems, besides regular maintenance overhauls.

Operators should be submitted to a specific training programme in accordance with their own functional requirements, comprising theoretical courses and on-the-job training. A simulator aided operator training through a dynamic operator training system can improve and speed up this process significantly, enabling the trainee to repeat actions that were not performed correctly until an acceptable performance level is reached. Depending on the functions allocated to each operator, in some cases, training can last one year, before being able to run a unit autonomously. Refresher training sessions should also be delivery periodically, to check if appropriate actions are taken, when an immediate response is required. Although human error cannot be eliminated, training is one key aspect to lower the operational risk, especially during emergency situations, when human factors are at stake and a swift and appropriate response from plant operators is intended and expected.

9.2.1.2   Safety Requirements

9.2.1.2.1   Procedures

Safety alerts are included in the operating instructions set out by systems and equipment manufacturers (technologists). The Emergency Safety Plan (ESP) should also be readily available to all personnel, preferably both on paper and electronically, in the internal information technology network.

The ESP provides details about the actions required in the event of emergency situations. Usually, foreseeable loss scenarios are included and selected as the basis for periodic safety drills. The main purpose of drills is to test the preparedness of the industrial asset personnel to react to a specific emergency and limit the damage. Drill results highlight areas for improvement and allow corrective actions to be scheduled and included in the safety drill report [see also chapter 5].

Other relevant safety procedures that might be available are referred in 9.3.2.3.2.

9.2.1.2.2   Proactive Controls

Proactive controls include alarms, emergency or unplanned shutdowns (trips), proactive safety performance indicators, and event analysis. Under certain circumstances, they can be interpreted as early warning signs (EWS) of system malfunction (instantaneous) or as system safety deterioration (over time), when a benchmark or reference parameter indicating a normal operation situation can be established [see also chapters 6, 7 and 10].

In operational safety, an alarm activation means that an anomaly has been detected in the system or equipment and an urgent action is required to eliminate the cause. Alarms are installed in systems and equipment to allow actions to be taken well in advance of a more serious event. These include actions on a process variable, machinery malfunction, fire outbreak, etc.

In terms of process control, an alarm is an indication to the operator that is initiated by a process variable or measurement that has passed a predefined limit considered to be an undesirable or an unsafe value. Poorly functioning alarm systems, or lack of training operating a system or equipment under emergency situations, worsen the seriousness of upsets, incidents and major industrial accidents.

Alarm activation can be attributed to two main sources: operator error or equipment malfunction. False alarms may also occur, either due to a non-calibrated or faulty sensors. To overcome this situation, where critical parameters on critical systems and equipment are concerned, the '2oo3' (2 out of 3) voting system principle should be applied. Under these circumstances, this principle will issue a shutdown command if at least two modules (that is, modules of critical parameter sensors) issue a shutdown command. This voting system will fail to perform its intended function on demand if two failures occur together. In addition, both failures will have had to be undetected by the system's internal diagnosis; or one failure must be 'dangerous undetected' and the other failure has to be 'dangerous detected'. When two dangerous and detected failures occur, it is assumed that the system responds in a safe way, and a system trip will occur. This can be considered the means of last resort for the system to prevent a potential serious failure.

In fact, trips are the last resort available to halt the system operation, to avoid further deterioration and potential widespread damage to equipment or harm to personnel and the environment. This action can be triggered by human

intervention (the operator) or automatically by the system, when predefined operational parameter values are reached. The costs of trips, which happen more frequently than accidents, can be quite significant. Besides process interruption, that can only be resumed after the possible cause of the unplanned shutdown is well identified and corrected. Also, although a more serious failure has been prevented, an unplanned shutdown reduces slightly the useful life of equipment. When an abnormal event is stopped before causing any damage and the operational parameter returns to its normal zone, this is considered as a process near-miss.

To take full advantage of the information provided by such indications of system disturbances or malfunctioning, alarm and trip analyses (also called event analysis) are highly recommended.

Many industrial companies record these alarm occurrences in distributed control systems (DCS) and emergency shutdown (ESD) databases. Operators, supervisors and managers seek guidance from these databases, by recording key indicators and paying special attention when alarm flooding occurs (Oktem, *et al.,* 2013).

Asset managers are becoming increasingly aware that these databases are rich in information related to near-misses. In recent years, researchers have been developing key performance indicators or metrics, associated with potential trips and accidents, leading indicators and failure probabilities of each of the safety systems. When conducted at frequent and regular intervals, analyses that are associated with these performance indicators are usually referred to as dynamic risk analyses or simply near-miss analyses.

9.2.1.2.3   Reactive Controls

Reactive controls are actuated when an unwanted event materialised and an immediate response is needed to limit the damage to a system, equipment, property or even harm to people and environment. Contrary to the proactive controls, where no or limited damage is expected, reactive controls are the last resort. An asset manager implements reactive controls to prevent an incident escalating to the degree where widespread damage and great financial loss would result. Typical examples on the conventional power generation industry are automatic or manual fire protection systems, dam spillways, fire brigade action, and also drills (derived from accident scenarios) conducted to assess the emergency preparedness of power plant personnel and reactive safety

performance indicators. The main objective here is to react, in the shortest time possible, to suppress the source of danger with maximum effect. Even if limited damage has occurred due to the swift action of first responders, the objective is to learn lessons and so prevent similar accidents. Although safety foresight was not effective to prevent this specific accident, as it was not able to assess the risk and treat it adequately, if appropriate corrective measures will be adopted, similar accidents will be prevented. However, if corrective actions were to be implemented in the system, the conditions heralding its occurrence will be detected in an earlier stage, enabling precautionary actions to be taken.

9.2.1.2.4   Safety Performance Indicators
In the literature, there is no unified approach concerning terminology and definition of Safety Performance Indicators (SPI).

An SPI can be defined as a basic parameter, described qualitatively or quantitatively, that is perceived as having potential meaning or a relationship to plant safety (Davies, *et al.,* 2006), (Hale, 2009), (Van Binnebeck, 2002).

A robust performance indicator should comprise the following features:

- Relevant: to provide useful information in due time for decision taking;
- Reliable: it provides the same value when used by different people;
- Measurable: able to be measured in an objective and clear manner;
- Feasible: cost-effective to collect;
- Comparable: it should allow comparisons over time;
- Resistant: resistive to manipulation, misuse and misunderstanding;
- Clear: easy to define, report, and evaluate;
- Specific: in relation to what is to be measured;
- Sensitive: reacts to what is being measured;
- Significant: in terms of sample size (number of events);
- Auditable: likely to be auditable.

Some precautions should be taken to prevent manipulation and misuse of performance indicators, especially when using performance indicators in bonus pay systems.
The metrics chosen to establish a quantitative or qualitative SPI system that can be compared to a reference (benchmark) is the main difficulty to be overcome. If an in-house SPI system is used, reference values or qualifications (benchmark)

should be defined by the asset manager, taking into account the objectives set by the management board. However, in the case of an external SPI system, the benchmark should be defined by an independent organisation (e.g. a user group or industry sector association) to enable comparisons between peers.

In addition to the selection criteria, several types of indicators have been proposed.

According to the OECD, indicators can be divided into two types:

- 'activities indicators'; and
- 'outcome indicators'.

Activities indicators are designed to identify whether organisations are taking actions believed to lower operational risks. In contrast, *outcome indicators* are designed to measure whether such actions are, in fact, leading to a lower likelihood of an accident or reducing the potential impact on human health or the environment, should an accident happen (OECD, 2005).

On the other hand, indicators can be divided into two groups, according to their use (Dahlgren, *et al.*, 2001):

- leading (or proactive); and
- lagging (or reactive) indicators.

Leading indicators are useful as a precursor of safety degradation, allowing early management reaction. Lagging indicators are commonly used to drive plant performance, for monitoring, and for benchmarking against similar plants.

Finally, two dimensions of safety indicators can be considered (Hopkins, 2009):

- personal safety indicators versus process safety indicators; and
- lead versus lag indicators.

According to Hopkins, although the distinction between personal and process safety indicators is relatively clear, the distinction between lead and lag indicators, while frequently referred to, is rather more problematic. However, if one is interested in knowing how well safety is being managed, the distinction between lead and lag indicators becomes largely irrelevant, as both will provide relevant information to assess and monitor safety performance.

In summary, besides the great number of definitions that can be found in the literature, what should be kept in mind is the fact that if properly selected, SPIs are useful for:

- evaluating/measuring and comparing safety performance over time for a given asset or group of assets, over a cross-section of assets at a given time, etc.; and
- informing decisions about the safety performance improvement of an industrial asset.

Finally, the selection process of any SPI should consider the effort spent on data collection, treatment and reporting against the usefulness of the information provided, especially in terms of risk-mitigating actions or safety improvement measures that it can generate.

In summary, although performance indicator systems have been established by sector associations to make results comparable between peers, no standardised SPI system has been established so far. Performance indicator systems may provide information regarding the process and equipment safety, but the main objective of such systems is monitoring operational performance and not operational safety.

9.2.2 Maintenance

Maintenance can be managed by different methods according to the type of system or equipment requiring intervention. The simplest method is corrective or run-to-failure maintenance, based on the principle break/repair or replace, which nowadays is seldom applicable to systems or equipment, as usually basic preventive tasks are performed in all of them, such as lubrication, calibration or visual inspection.

Another method is preventive maintenance, where tasks are based on regular time intervals or running hours. This takes into account the specific mean-time-to-failure (MTTF) statistic for each type of equipment, usually available from the manufacturer or in the specialist technical literature. The main disadvantage of this method is that MTTF is an average value that is not kept constant for all similar equipment, which means that either unnecessary maintenance interventions or catastrophic failures can happen. In the first case, labour and material are wasted. However, the second case implies that the run-to-failure method was applied, which is even more costly.

Predictive maintenance is based on a regular monitoring of several operational parameters of the equipment, and process systems will provide the data required to maximise the interval between repairs and minimise the number and cost of unplanned unavailability due to failures. Predictive maintenance is a condition-driven preventive maintenance method. Instead of being based on industrial or equipment average-life statistics (i.e. MTTF) to schedule maintenance activities, it uses direct monitoring of equipment condition, process efficiency and other parameters to estimate the MTTF or loss of efficiency for each system or equipment.

9.2.2.1   Type of Components

Not all equipment and systems have the same importance for the asset manager or the production process. A few of them are so relevant, that a serious failure can produce significant financial losses, resulting from widespread damage and process interruption over a large time period, from several months to one year.

The identification and selection of critical components is generally carried out by dividing the industrial asset components into *critical* and *influence* (non-critical).

A *critical component* can be defined as having the following features:

- its failure can cause an extended forced outage, or
- its failure can endanger the safety of the asset, the environment or personnel, and
- has long lead times and high costs for repair or replacement.

And an *influence component* is characterised by:

- failure results in significant degradation of asset performance but does not cause forced outage, or
- failure does not endanger safety of asset personnel or cause widespread secondary damage, and
- failure susceptibility is known due to 'asset specific' experience.

Examples of critical components in the conventional power generation industry, namely fossil-fired power plants, include live steam piping, turbine and electrical generator rotors and step-up power transformers.

Special attention should be directed towards critical components because the maximum probable losses due to their failures are by far more serious than those

from 'influence' component failures. Critical components are the ones where the application of an asset management methodology can reap the most benefits, by avoiding losses due to major accidents.

Bearing in mind this observation, critical components should have clearly defined requirements that are substantiated by records. It is much more advantageous and useful to maintain the strict requirements on a smaller number of items than it is to assign the same criteria to every piece of equipment in the facility (Newslow, 2001). This distinctive feature allows event analyses that can reveal anomalies in systems and the opportunity to enact the necessary preventive measures in good time.

9.2.2.2   Predictive Maintenance

Predictive maintenance is designed to help determine the condition of in-service systems and equipment, and to estimate when maintenance should be performed. This approach enables cost savings over routine or time-based preventive maintenance, because activities are performed only when justified. So, it is regarded as condition-based maintenance, as is carried out in accordance with estimations of the degradation state of a component.

The main advantages of predictive maintenance are to allow convenient scheduling of corrective maintenance, and to prevent unexpected system and equipment failures. In this way, safety foresight is applied in predictive maintenance to define the optimal maintenance period.

In the conventional power generation industry, a critical system, specifically the turbine-generator set, dictates the interval between two power plant maintenance overhauls, as per the manufacturer's instructions. These overhauls are based on running hours and number of starts and allow all the remaining predictive maintenance actions to be adjusted to the turbine-generator overhaul.

9.3   Process Control

According to ISO 9001:2015 Quality Management Systems standard, a *process* is a set of activities that are interrelated or that interact with one another. Processes use resources to transform inputs into outputs and are interconnected, because the output from one process often becomes the input for another process. An effective process control enables a product to be delivered, or service provided to

clients according to the procedures in force at the organisation and hence achieving the required quality standards.

9.3.1 Certification (Quality, OHS, Environmental Management)

In addition to the internal process control in force at the organisation, the quality management system certification provides a universal assessment level. This enables a competent and independent entity—the certification body—to periodically assess the adherence of an organisation's processes to the quality management system principles set out in the ISO 9001 standard. The same is applicable to the occupational health and safety (OHS) management system standard (ISO 45001) as well as to the environmental management system standard (ISO 18001). The certification acts as a guarantee that the organisation is following the requirements established in the relevant standards, and that control mechanisms are set in place to enable the early detection of process degradation. The main advantages of certification claimed by certification bodies include, amongst others, better management control and improved internal communication. Both advantages may have a positive impact on safety foresight.

In addition, industrial companies have physical assets and staff associated with the manufacturing or production process. Usually, in these cases, audits to award or renew the certification include on-site visits to the plants, where areas for improvement can be recommended, including aspects related to process safety. When this happens, safety foresight has been applied.

9.3.2 External Entities

External entities play an important role in establishing rules and controlling organisation activities by reference to directives, regulations, standards, specifications, etc.

Regulators, certification bodies, insurance companies, technologists (high-tech manufacturers), O&M specialist companies, users' groups, and sector associations are a few examples of entities that can help an organisation to be aware of process safety deterioration. In the main, they do this by identifying weak points and recommending preventive and corrective actions to put the process back into conformity with the principles of the relevant reference document or technical specifications.

9.3.2.1 Regulators

Regulators, also known as regulatory agencies, regulatory authorities or regulatory bodies, are public authorities or government agencies responsible for exercising autonomous authority over some areas of economic activity. This includes rulemaking, enforcing rules and regulations, and imposing supervision or oversight for the benefit of the public at large. Some independent regulatory agencies perform investigations or audits, and others may levy fines on the relevant parties and order certain measures to be implemented [see also Chapter 13].

In power generation but also in other industries, there are generally two areas where regulators act to exert their authority: occupational health and safety (OHS) and protection of the environment. The first is a multidisciplinary field concerned with the safety, health, and welfare of people at work. The goals of occupational safety and health programs include providing a safe and healthy work environment. The second deals mainly with environmental pollution control, including air quality, water quality, waste management and contaminant clean up.

Other related areas supervised by regulators include the use of large amounts of dangerous chemicals in industry (under the Seveso Directive), dam structural safety, and emergency safety valves for pressurised equipment used in several process industries that make use of steam in their production processes. In the case of that last example, accredited laboratories calibrate and certify emergency safety valves periodically, according to the regulations set out by the competent authority. If these requirements are not met, the asset owner will have its operating licence cancelled.

So, when major accidents happen, these can pose a significant threat to people and the environment, cause huge economic losses and disrupt sustainable growth; hence the value of external regulation.

9.3.2.2 Certification Bodies

The main roles of certification bodies are to award and issue certificates of compliance with the relevant management system. An auditor is usually involved in the company's certification process (first certification or renewal), by conducting the audit on behalf of the certification body and then by reporting their findings back to it. A consultant can also participate in the process, providing specialist advice to ensure that the management systems will meet the certification requirements.

In the case of OHS, the main advantages of certification to the organisation are to:

- increase employees' awareness and motivation towards safety;
- ensure that official and legal OHS requirements are met;
- prevent accidents;
- reduce downtime and production stoppages;
- reduce the cost of insurance policies;
- improve the organisation's image as a safe and reliable business in the eyes of its clients, suppliers, authorities and investors.

The relevant fact here is that both internal and external audits enable the detection and correction of non-conformities that otherwise would contribute to the occurrence of serious accidents.

9.3.2.3   Insurance Companies

In developed market economies, insurance plays an important role in controlling the risks carried by industrial assets. In simple terms, a company agrees a contract with an insurer to transfer a certain portion of its risk for a determined sum of money, called a *premium*. The contract is called an insurance policy.

Usually, the insurance of large industrial assets is structured so that relatively minor losses are borne solely by the asset holder (Barnard, 2006). This implies that only losses above a certain amount, called the *deductible*, are incurred by the insurance company. On the other hand, the usual practice in the insurance market is also to limit the amount of losses covered by the insurance policy. Both limits, the deductible and the loss coverage of an insurance policy, have a financial impact on the premium paid by the asset holder.

In general, the risk retention and risk transfer strategy depends on the level of risk that a given company is willing to accept and on the premium offered by the insurance market. When the insurance market is *hard* or risk averse—especially following large losses after natural disasters that have generated widespread damage—premiums rise and companies tend to retain a higher level of risk, to keep a similar insurance premium value. When the market is *hard*, two options are available: to increase the insurance deductible or to lower the insurance limit. Increasing the deductible usually reduces the premium more effectively than would lowering the insurance limit. This is because by agreeing a larger deductible, which is the first layer of the insurance policy, the organisation accepts to bear a

higher portion of the loss from each accident, up to the deductible limit. Only when the loss exceeds the deductible limit, the insurance policy will cover the remaining portion of the loss.

Usually, high frequency risks result in low severity losses, so these are the risks that will be retained in the company, because usually they fall under the deductible limit. On the contrary, low frequency risks result in high severity accidents, so these are the risks that are typically transferred to the insurance market, as their materialisation could affect heavily the economic activity of the company or even jeopardise its existence as such.

For large industrial companies, there is also the possibility of self-insurance through an internal reinsurance company called a '*captive*', which enables them to retain more risk at an intermediate level of frequency and severity. The mechanism of risk retention and risk transfer is represented schematically in Figure 1.

Figure 1: Risk Retention and Risk Transfer Mechanism (source: Outreville, 1998, p177)

In terms of asset management, the most relevant insurance policy is called *Property Damage.* This type of policy may comprise two parts: *Material*

Damage/Machinery Breakdown (MD/MB), covering losses related to physical assets, and *Loss of Profits* (LP), *Business Interruption* (BI) or *Time Element* (TE), which cover losses derived from the interruption of the company's supply chain. In this context, supply chain means power generation, transmission and distribution, gas supply, pulp and paper production or any other manufacturing process. Only the MD/MB insurance contract can be awarded separately.

However, only sudden, unexpected and random property damage claims are eligible for payment by the insurance companies. This means that equipment failures resulting from wear and tear are ruled out from the insurance contract.

9.3.2.3.1 Audits

In the case of a disaster, the large economic value of industrial assets can impart great financial losses both to asset holders and the insurance companies. Having in mind this possible outcome, insurance companies carry out regular site visits, whose frequency is directly proportional to the total asset insured value (value at risk). Site visits are conducted to check if operational safety is being kept at an acceptable level, according to applicable codes and standards, internal procedures, controls and best practices. Following each site visit, a report can be delivered to the asset holder, mentioning areas for improvement, called recommendations.

9.3.2.3.2 Recommendations

Recommendations are intended to improve the risk and safety of an industrial asset as seen from the perspective of the insurance company. These are based on the global knowledge and experience of the insurance company (statistical data of similar incidents), international standards and insurance company technical datasheets or O&M instructions from critical equipment manufacturers (technologists).

It is well known that insurance companies rely on statistics of previous losses for premium calculation purposes. So, it is not hard to believe that some of the recommendations proposed result from causes related to similar events that have occurred elsewhere, that could not be anticipated and have resulted in a claim. In effect, the goal of an insurance company is to set up a robust asset management system of 'zero accidents'. Although the asset holder pursues the same objective, recommendations involving considerable financial investment should undergo a technical-economic analysis prior to taking the final decision regarding its

completion. As financial resources to invest in completion the recommendations are totally incurred by the asset holder, a detailed technical-economic analysis allows to prioritise the ones where major benefits can be achieved at the lowest possible cost.

Typically, the recommendations issued by insurance companies after field visits can be divided into three categories:

- Procedures;
- Inspection and Testing; and
- Systems and Equipment.

This division of insurance recommendations into categories is important mainly for asset management purposes. For example, if most pending recommendations fall into the 'procedures' category, it reveals that the company (the asset holder) has in place an operational safety system with a lot of scope for improvement. In such a case, the insurance company may even request a safety improvement plan, to assure that the company will reach a certain operational safety standard in the shortest time possible. On the other hand, if most of pending recommendations are in the 'systems and equipment' category, it shows a mature operational safety system in place. Under these conditions, and having in mind that operational safety is a continuous improvement process, insurance companies may prioritise those recommendations where, if followed, most benefits can be expected. Although both the insurance company and the asset holder pursue the same objective, which is to reduce operational risk to a satisfactory level, the investment necessary to complete recommendations falls completely onto the asset holder side. Under these circumstances, both will benefit from reducing the probability or consequence (or both simultaneously) of a serious failure, but the asset holder incurs all the costs. In the hypothetical scenario where all the recommendations of this category are implemented, the financial coverage provided by the insurance would decrease so drastically that it would become residual, possibly only useful for 'Act of God' events.

Recommendations relating to procedures ask for new ones to be written or existing ones to be improved, in either case to reach the standards established by the insurance company. However, in a broad sense, procedures have two dimensions. The first is the written procedure itself ('paperwork'), where all the instructions and warnings are laid down. The second is the strict fulfilment of the procedure.

When a procedure deals with inspection and testing instructions, the relevant part is the instructions and then the recommendation falls under the category of 'Inspection and Testing'. When the procedure, as paperwork, needs to be improved or updated, the recommendation is from the 'Procedures' category.

Procedures of the insurance companies address mainly operational safety and fire prevention. Examples include the no smoking policy, 'automatic fire protection systems impairment communication' to the insurance company (for assessing the need of reinsurance), hot work for maintenance works involving flame or heat generation (e.g. oxy-cutting and welding) and contingency plans for critical equipment, typically, generator step-up transformer spare units, for reducing the loss of profits related to the period of unavailability caused by the failure, which can last 12 months or more.

Usually, procedures are not costly in themselves and are easy to set up. However, they may be challenging in some situations because they can present practical difficulties. For example, to check that fire protection systems remain fully operational, emergency procedures require that the systems are regularly discharged. However, as the practical application of procedures are considered under the category of 'Inspection and Testing' recommendations, this will be referred to later.

Usually, procedures are of reduced cost and easy to setup, but may be challenging in some situations, because they can present practical difficulties or be quite expensive to carry out.

Inspection and testing' covers all periodic maintenance actions that are required to keep all safety systems fully operational, and permanently ready for actuation. Generally, recommendations made under this category can be fulfilled at moderate cost. However, in some cases, testing can increase operational risks or become quite expensive to carry out. For example, the overspeed test of a turbine-generator set under actual conditions calls for a 10% increase in rotation above nominal speed. This test creates the risk of serious widespread damage. An example of a costly test is checking the tightness of a turbine enclosure. This requires a full-scale discharge test of its fire protection system: emptying a rack of CO_2 gas containers is a quite expensive undertaking. In such cases it is important to reach an agreement with the insurance company to perform the test under electronic simulated conditions or to use an alternative gas for tightness checking of the turbine enclosure respectively.

Finally, systems and equipment installation, replacement, refurbishment or extension entail considerable investment that requires a technical-economic analysis to inform the final decision. Examples include automatic or manual fire protection systems and process safety control devices (e.g. a synchro check relay in the command line of the circuit breaker for synchro in manual mode).

Recommendations issued by insurance companies are one of the most effective tools to continuously improve industrial risk and operational safety. Safety alerts are also used to warn insured asset holders, when property damage has occurred elsewhere in similar equipment. This will enable asset managers to question the technologist regarding the failure risk in its own equipment. In this way, safety foresight is used by insurance companies to prevent accidents in similar equipment elsewhere.

The global knowledge and experience of insurance companies cannot be disregarded as a powerful tool in the prevention of serious industrial accidents. When these accidents happen, insurance companies are called-on to pay the claims, which represent the major part of the property damages incurred. In addition, experts (loss adjusters) are called to perform a thorough accident analysis aiming at determining the root cause and contributing factors of the accident. Once these are determined, insurance companies check if identical conditions are present in similar systems and equipment elsewhere, and if there is a match, recommend preventive actions. In this way, insurance companies use foresight in safety to prevent similar accidents from happening in other locations.

From the side of the asset holder, even if recommendations are not mandatory, and if it is not possible to establish a direct mathematical relation between the insurance premium paid and pending recommendations, it is important for asset managers to make their own judgment about investment priorities in terms of asset risk and safety improvement.

9.3.2.4   Technologists

Technologists, in this context, are specialist firms that have the required knowledge and experience to develop a critical asset, characterised by a highly complex technology manufacturing process, only available to a restricted number of companies. In the power generation industry, gas turbine manufacturers, also called original equipment manufacturers (OEM) for combined-cycle gas turbine (CCGT) power plants is a good example. As only a few exist worldwide, a limited

number of options are available to power generation companies. In addition, the increase in demand for CCGT power plants by power generation companies has raised the competition amongst the gas turbine manufacturers, pushing them to innovate, aiming at reaching higher efficiency rates. Several versions of the same equipment model were released, as manufacturers were introducing improvements constantly, some due to equipment malfunctions or failures in the previous versions. These actions increased the risk of failure, as innovative solutions were released without enough time to mature.

The maintenance of CCGT power plants, particularly gas turbines, is a complex discipline, especially with respect to the integrity of hot gas path components, which is the part of the gas turbine where temperatures can reach as much as 1,200°C. To help asset managers in O&M matters, long term maintenance agreements (LTMA) are offered by turbine manufacturers as a guarantee of specialist technical support, for a period of 15 years or more. Besides maintenance technical support, these agreements may comprise daily event analysis, implying the delivery of all plant operational data to the manufacturer. If process degradation signs are detected, the manufacturer will contact the asset manager for more information about how the equipment is being used or maintained. Finally, when agreed by both parties, LTMA contracts may also include penalties, which can be applied if a pre-set standard quality of service level is not met by the manufacturer.

Once again, the global knowledge and experience of these technologists play a very important role. They are aware of all equipment malfunctions and failures happening globally. When serious failures can jeopardise other similar units elsewhere, a technical information letter is issued and sent to all asset owners of the same equipment version. These technical letters include recommendations about how to operate or when specific parts should be replaced. The rationale is to prevent failures in other similar equipment elsewhere. In this sense, foresight in safety is being put into practice by the cooperative action of the technologist.

9.3.2.5   O&M Specialist Companies

These specialist companies and institutes are usually contracted to carry out specialised industrial tasks using advanced technological means and highly qualified human resources, such as those with expertise on power plant command and control systems and on critical process equipment. In the power generation industry, critical equipment comprises turbines, electric generators and step-up power transformers. These are responsible for the major failures in the power generation industry and the highest value claims paid by the insurance companies. O&M specialist companies have an in-depth knowledge and global experience in the field, which can be very useful when providing O&M specialist services to asset holders. Global experience brings awareness of the major risks and failures involved. An external view by qualified entities is of utmost importance to improve industry processes, procedures and practices.

9.3.2.6   Users' Groups and Sector Associations

Users' groups can be thought of as clubs focused on the use of a specific technology. The groups are usually associated with a company that is a developer or technologist. Although these are external interest groups, the participation of each asset holder allows peers to share relevant information about processes and equipment.

A good example in the power generation industry are the CGGT users' groups that are affiliated with each gas turbine manufacturer. On the CCGT manufacturing technology, the gas turbine is the most critical equipment. Challenged by the electricity market, the power generation industry demanded higher process efficiency rates, fostering a strong competition among gas turbine manufacturers. The rapid evolution of gas turbine technology turned it into a non-mature technology. Failure rates started to increase and company profit losses worsened. Sometimes, failure root-cause analysis carried out by the manufacturers took too long or could not provide satisfactory technical answers to the questions asked by asset managers. Users' groups were the solution found by asset managers to exchange technical information regarding problems encountered in this type of technology. Information about corrective actions that were effective for a specific failure could be shared and applied to similar equipment operated by another user, well in advance or when appropriate. The main objective would be reducing this specific failure probability and improving overall operational safety for all the other users.

Through regular group meetings, formal presentations and attendee-driven discussion sessions focusing on the design, erection, operation, and maintenance of the integrated plant; asset managers are aware of problems and solutions that could be useful to their specific case. In this way, in taking the appropriate

measures, safety foresight is put into practice, as similar potential equipment failures are anticipated and prevented from occurring. This is even more relevant where an LTMA agreement between the asset manager and the technologist is not awarded.

Sector associations act at a higher level than users' groups and deal with a wider range of issues. In the electricity industry, two examples are VGB and Eurelectric.

VGB is the technical association of energy plant operators. Members are companies that operate worldwide facilities for the generation of power, heat and cooling as well as for energy storage and sector coupling. As an independent technical competence centre and network, VGB supports its members in their operational business as well as in the implementation of innovations and strategic challenges. One of the main goals is to strengthen and safeguard a high standard in operational and plant safety as well as health and safety at the workplace. In addition to technical issues, VGB is also actively involved in the political and social debate on technical issues, on behalf of operators. Eurelectric is the sector association representing the common interests of the electricity industry at a European level, plus its affiliates and associates on several other continents. It encompasses all major issues affecting the sector, from generation and markets to distribution networks and customer issues.

As VGB deals with a more specific set of technical issues than Eurelectric, mainly related to the energy plant operators, it is easier to identify potential issues where safety foresight may be applied. VGB provides its members with an international network, a platform for the exchange and transfer of technical know-how, as well as access to qualified expert knowledge via, for example, operational and availability databases for benchmarks. These technical means can be used for a wide range of technical purposes. For safety foresight purposes, the most relevant is the benchmarking tool provided by VGB through power plant performance indicators. However, the indicators available are mainly related to plant performance in terms of availability of power supply to the electrical grid and not to operational safety performance. In this sense, the information obtained through users' groups is more relevant to the daily life of power plant asset managers. This information acquaints asset managers with technical problems that may seriously affect process and equipment safety, allowing them take adequate actions for its prevention.

9.4   Conclusions

Nowadays, industrial assets are managed according to internal procedures, controls, standards and best practices. However, these are also informed by the influence or oversight of external entities, which operate through legal and contractual obligations, or by the asset holder's own initiative. Legal obligations are mandatory, so they cannot be considered an option if the company aims at staying in the market. Contractual obligations can derive from negotiated agreements that require pre-conditions to be set out. If obligations result from the asset holder's own initiative, they are based on the trade-off between benefits derived from the commitments assumed and the incurred costs. In this case, although almost every benefit can be monetised, some of them are difficult to quantify, like the company's image or reputation.

Internal safety requirements like procedures and controls are a very important tool to prevent accidents.

Industrial companies have a set of tools available to monitor process and equipment safety, including proactive and reactive controls, event analysis and performance indicators.

Certification of quality, OHS and environmental management systems are a guarantee that the organisation follows the requirements established in the relevant standards and that control mechanisms are in place to enable the early detection of process degradation. The main advantages of certification claimed by certification bodies include, amongst others, better management control and improved internal communication. Both advantages may have a positive impact on safety foresight.

External entities play an important role in establishing rules and controlling organisation activities by reference to directives, regulations, standards, specifications, etc.

Regulators, certification bodies, insurance companies, technologists (high-tech manufacturers), O&M specialist companies, users' groups and sector associations are examples of external entities that through foresight in safety can contribute to improve the operational risk and safety of industrial assets at different levels. This is particularly true of insurance companies, who are usually the first external entities to be aware, are in close contact with the accidents through loss adjusters, and who have access to confidential, detailed information regarding the causes of

the losses. Inside information of the causes of loss is relevant to similar systems and equipment under operation elsewhere. Under these circumstances, recommendations issued by the insurance company towards improving risk and safety on similar units can be considered as a safety foresight measure.

Finally, asset managers are responsible for the process and equipment safety of the company. They should be aware of all the internal and external tools and entities available to prevent major accidents. By being aware of and using these tools adequately, asset managers can play a relevant role in reducing the probability of major accidents.

9.5   References

Barnard, I. (2006), "Asset Management – An Insurance Perspective", 1st WCEAM - Proceedings of 2006 World Congress on Engineering Asset Management (J. Mathew, J. Kennedy, L. Ma, A. Tan, D. Anderson Editors), 11–14 July 2006.

Chakraborty, S.; Flodin, Y;. Grint, G.; Habermacher, H.; Hallman, A.; Isasia, R.; Karsa, Z.; Khatib-Rahbar, M.; Koeberlein, K.; Matahri, N.; Melendez, E.; Moravcik, I.; Preston, J.; Prohaska, G.; Schwaeger, C.; Tkac, M.; Verduras, E. (2003), "Risk-based Safety Performance Indicators for Nuclear Power Plants", Paper # M01-6, Transactions of the 17th International Conference on Structural Mechanics in Reactor Technology (SMiRT 17) Prague, Czech Republic, August 17 –22, 2003.

Davies, J.; Finlay, M.; McLenaghen, T.; Wilson, D. (2006), "Key risk Indicators – Their Role in Operational Risk Management and Measurement".

Dahlgren, K.; Lederman, L.; Palomo, J.; Szikszai, T. (2001), "Safety Performance Indicators", Topical Issue Paper No. 5, pp. 2, International Conference on Nuclear Safety, Vienna.

Hale, A. (2009), "Why Safety Performance Indicators?", Safety Science 47, pp. 480.

Hopkins, A. (2009), "Thinking about Process Safety Indicators", Safety Science 47 – Special Issue on Process Safety Indicators, pp. 460-465.

Newslow, D.L. (2001), "The ISO 9000 Quality System: Applications in Food and Technology", John Willey & Sons.

OECD (2005), "Guidance on Safety Performance Indicators" (Interim Publication), Series on Chemical Accidents, No. 11, 2003, pp. 8, rev. 2005.

Oktem, U.G.; Seider, W.D.; Soroush, M.; Pariyani, A. (2003), "Improve Process Safety with Near-Miss Analysis", American Institute of Chemical Engineers (AIChE), May 2013.

Outreville J.-F. (1998), Retention, Self-Insurance, Captive Insurance Companies, chapter 10, In book: Theory and Practice of Insurance, Kluwer Academic Publishers, (pp.179-196), accessed on Researchgate, 23rd November 2020

Van Binnebeck, J.J. (2002), "Results of the WGIP Baltimore Workshop Sessions related to PIS", Specialist Meeting on Safety Performance Indicators, Madrid, Spain, 15-17 Oct. 2000, Safety Performance Indicators – Workshop Proceedings, ref. NEA/CSNI/R(2002)2, May 2002.

Van der Lei, T.; Herder, P.; Wijnia Y. (2012), "Asset Management – The State of The Art in Europe from a Life Cycle Perspective", Springer.

10   Big data analytics and early warning signs

Eric Marsden, Fondation pour une culture de sécurité industrielle (FonCSI), France,
Nicolas Dechy, Institut de radioprotection et de sûreté nucléaire (IRSN), France,
Ana Lisa Vetere Arellano, European Commission Joint Research Centre, Ispra, Italy.

10.1   Executive summary

The analysis of "big data" has generated a large amount of interest in industry over the past decade. Promoters can point to a number of benefits and success stories in the retail and entertainment areas, as well as in optimization of industrial processes, and more recently in overcoming security (Amanullah et al., 2015), safety (Huang et al., 2019) and reliability (Ham & Park, 2020) challenges.

To what extent can increased collection and analysis of data help to detect early warning signs of system failure, and predict the occurrence of the infrequent events that are relevant to the management of industrial safety and major accident hazards? What should we monitor? What are the challenges and risks of inadequate use of big data?

10.2   Key messages

Big data analytics has a significant potential to improve the detection of early warning signs of system failure, when compared with the use of standard statistical tools or of unassisted human monitoring of system data. It also provides opportunities in health usage monitoring of equipment (e-Health).

Industry 4.0 has fostered the burgeoning of digital twins, offering enterprises the capability of real-time situation awareness thanks to digital technologies (ubiquitous sensors generating big data, artificial intelligence, machine learning, predictive software analytics, etc.). These dynamic digital simulation models of physical assets allow continuous data, information and knowledge (DIK) acquisition related to system performance and failure. Thus, enterprises are able to rapidly run diagnostics, intervene to correct problems before they become critical and detect new improvement opportunities.

Effective use of predictive analytics faces a number of obstacles, some induced by the black-box-effect of algorithms and tools. Their use requires specialised skills of the analysts who build the data collection and analysis tools, but also from the users, who will require training though they will not become data analysts.

Algorithms can improve human judgment, but will never replace it completely. Humans, within organisational processes, will remain important at every step of a big data analytics process: framing individual and organisational attention to the data that important to analyse, putting in place and checking the data collection process, interpreting the importance of outliers found in data, and validating the causal nature of correlations identified.

Automatic decision-making based on algorithms should be very carefully controlled to check for potential biases in available data and its treatment.

The big data paradigm emphasizes the importance of data *quantity*, but analysts should start by checking the *quality* of data and its relevance to the analysis undertaken. This issue is particularly significant when addressing the human and organizational dimensions of system operation and searching for the underlying causes of events. Moreover, the big data paradigm focuses on official reported data, disregarding tacit and informal data, which is sometimes critical to understanding the complexity of social and organisational issues that affect safety.

Data and information are constructed by worldviews, tools such as sensors and human perception; different meanings can be associated to the same data or word by different people and groups of professionals.

Big data analytics raises questions regarding privacy, information security, ethics, which should be handled by a well-designed data governance process.

10.3   Context

Statistical analysis of reliability data has a long and illustrious history, including Dr Snow's analysis of the 1854 cholera epidemic in London, which identified the public water fountain at the epicentre of the outbreak. Research over the past 50 years has also shown that even simple mathematical models, such as linear regression models, provide better predictions and forecasts than human experts

in a range of situations[39], by avoiding a range of cognitive biases that affect human information processing.

Over the past decade, a new strand of work on "big data analytics", which applies statistical techniques and more recent machine learning techniques to larger sets of data, has gained much attention. The emergence of this concept has been enabled by a number of coinciding factors:

- Affordable computer systems that are able to store and process very large volumes of data.
- Increased use of "smart sensor" systems that can provide real-time data on various performance measures of equipment (temperatures, pressures, vibration levels, etc.). This "internet of things" trend has been enabled by technological advances in integration levels for microelectronics, by reduced power consumption and improved battery technology and by the development of low-power wireless communication technologies. Other sources of data include social media platforms, e-commerce and smartphones with geolocalisation features.
- Increased industrial use of sophisticated machine learning techniques[40] such as neural networks that allow classification, anomaly detection, and optimization.
- Development of natural language processing (NLP) tools that allow automated treatment of large volumes of unstructured text. These tools are often based on statistical learning rather than on language models developed by humans. Companies often possess large corpuses of unstructured text, including historical information which has not been manually classified when companies moved from paper-based to computer-based storage of reports and data logs. Until recently, unstructured text could not be used in analytical tools without significant manual effort to classify the data. New techniques can extract structured information from these data sources and allow their combination

with numerical data from other sources. NLP also enables automatic classification or clustering of documents according to their level of similarity.

Table 1: Comparison between structured and unstructured (with natural language processing) data analysis (Dechy and Blatter, 2018)

	Prior assumptions to verify	No prior assumptions to verify
Structured data analysis	Validation of the correlations identified by experts	Cross-referencing databases to identify targets to be analysed
		Cluster analysis to discriminate
Analysis of textual data extracted from event reports	Automatic semantic categorisation of incidents	Failure analysis to identify the features of the systems considered: type, manufacturer, composition,…

The term "big data" is generally used to refer to new generations of data collection and analysis systems, which were initially characterized by "three Vs":

- a large **volume** of data[41], that typically cannot be stored on a single personal computer;
- high **velocity**: data sources that generate large streams of events that cannot feasibly be stored, but must be filtered and analysed in real time;
- significant **variety**: different data formats, often unstructured or multimedia, which are difficult to store in traditional relational databases.

Four other V's have later been added to this motto (Khan et al., 2014):

- **veracity:** large volume and flows of data are automatically collected, but they may be erroneous or their accuracy becomes harder to check;

---

[39]*The field known as "clinical vs. statistical prediction" was developed by psychologist Paul Meehl, who reported on 20 studies that compared the predictions of well-informed human experts with those of simple predictive algorithms. The studies ranged from predicting how well a schizophrenic patient would respond to electroshock therapy to how likely a student was to succeed at college. Meehl's study found that in each of the 20 cases, human experts were outperformed by simple algorithms based on observed data such as past test scores and records of past treatment. Subsequent research has decisively*

confirmed Meehl's findings: more than 200 studies have compared expert and algorithmic prediction, with statistical algorithms nearly always outperforming unaided human judgment.
[40]*Machine learning is based on algorithms that can "learn" (infer relationships) from data without relying on rules-based programming.*
[41]*As an illustration, a typical offshore oil rig can generate two TB of data per day, and an A350 aircraft includes more than 6000 sensors (temperature, pressure, operating speed, stress, humidity, vibration, fuel flow rate, etc.) that generate more than 2.5TB of data per day.*

- **variability:** the rate of change of the structure of the data, which depends on the stability of the context in which data is extracted (for example, the same word used by the same person in different contexts and tones of voice may signify different meanings);
- **value:** the net added value for users, which is the difference between the gross benefit and the cost of collection and analysis;
- **visualisation**: the quality and relevance of data visualisation is a key to reveal the significance of data analysis and control biases of representation.

10.4   Motives and benefits of big data analytics

Some analysts, start-ups and industrial promotors, full of optimism concerning the potential of these technologies and techniques to improve industrial production, refer to a fourth industrial revolution, or "Industry 4.0", in which continuous streams of real-time data from sensors within the production line and upstream supply chain can be analysed using artificial intelligence techniques to allow increased product customization, performance optimisation, more flexible and adaptive production and anticipatory detection of critical events.

Leveraging these sources of data to improve decision-making (an activity called *predictive analytics*) requires a combination of skills in new computer technologies, statistical analysis, machine learning and data visualisation (an intersection called *data science*). Their application for reliability and safety purposes is more recent (10 years ago for software from start-ups).

Traditional collection and analysis of safety data (operational experience feedback, event reporting, generation of safety performance indicators) have long been used for safety management, to identify anomalies and to check that interventions result in a system improvement. Reliability engineers have long used trend analysis on critical system measures such as operating temperatures or pressures to identify deviations from design levels and from normal operating conditions. When thresholds are passed, alarms inform plant operators; when extreme levels are passed an emergency shutdown is triggered. Big data analytics, when applied to safety issues, is a complement to these traditional methods that analyses more complex correlations or interactions between multiple variables to help identify more subtle anomalies, that can have performance or safety implications. These more sophisticated analysis techniques are also better able to account for slow changes in plant performance as it ages than the static thresholds used in traditional trend analysis. They may be able to produce relevant insights when the plant is operating outside standard conditions (for example during startups and shutdowns), which is rarer for traditional statistical analysis methods.

Traditional statistical analysis of data requires the analyst to formulate a hypothesis, then collect relevant data and undertake an analysis to check whether the data supports the hypothesis. The specific promise of new "unsupervised learning" models is that a computer might identify 'patterns', 'features' or 'a model in the data' that predict specific outcomes (such as mechanical breakdowns or technical failures) in an "automatic" manner, without benefiting (or indeed suffering[42]) from the preconceived assumptions of the safety analyst.

As discussed in section 10.6, this promise does not eliminate the need for an experienced analyst to examine the features identified by the algorithms, verify the assumptions and assess whether they are of relevance to operational performance or safety in general and in a specific context of use[43].

10.5   Safety and security applications of predictive analytics

The most prominent applications of predictive analytics techniques have been in e-commerce and entertainment, with recommendation engines ("if you liked that, you might like this") and targeted advertising ("if you searched for this, you might buy this") being the most commonly developed features. The techniques also have applications in the safety domain, related to the detection and analysis of early warning signals. In particular, the algorithms, processing facilities and approaches can be applied during:

- *The detection phase of an experience feedback/reliability analysis process:* they can help to detect anomalies, unusual trends, typical configurations and emerging patterns of behaviour that may affect only a subset of equipment or a population of system users. This can be used to detect the early warning signs mentioned in chapter 6 (Strucic, 2020) dedicated to the visibility of early warning signs.
- *The analysis phase:* they can help analysts to dig deeper and test their hypotheses. The "slice and dice" data processing facilities that are associated with the development of a big data infrastructure can be used by analysts to extract all events that match specific search criteria, to filter issues and to check for trends or anomalies in these events, in a much more convenient manner than when data was fragmented across multiple departments and storage systems.
- *The prediction phase:* they can help safety experts anticipate the performance of system changes before they are implemented, through more sophisticated and precise system models.

Data sources that can be used for these safety applications of predictive analytics techniques include:

- Data generated by equipment and sensors, with very high volumes and high frequency of use in modern technological systems where many components have been instrumented to generate monitorable outputs.
- Operational data from management systems, such as the number of inspections undertaken, number of flights flown, number of customers.
- Text written by humans, such as the descriptions included in incident reports and inspection reports, the content of emails.

In the security domain, predictive analytics is being used in the following ways:

- *Maintenance of military vehicles*: predictive maintenance allows the military to reduce malfunction and failure of vehicles in operation, thanks to real-time data collection and analysis from sensors and telematics[44].
- *Prediction of soldier effectiveness*: in a virtual near-reality environment, soldiers can be monitored to predict how they will react with the help of biosensors that collect real-time data that can be analysed by predictive analytics and machine learning algorithms[6].
- *Tracking readiness of equipment*: in order to better manage military training and defence operations, real-time access to intelligence related to the degradation state and location of equipment is essential in order to make better informed decisions, e.g. if a tank needs to be moved from one military base to another[45].

Big data analytics techniques allow a move from a static analysis of the factors of system performance to a dynamic and continuous approach, allowing more customisation to the specific characteristics of the system.

In the next paragraphs, we describe a number of applications of predictive analytics to safety management.

10.5.1 Detecting new safety and security threats

Collecting and treating massive quantities of data may allow the early detection of anomalous situations which may represent new component failure modes or threats to system safety. Big data analysis techniques allow high-dimensional data to be analysed using data mining techniques, searching continuously for new correlations or new outliers between multiple streams of data, such as those provided by various sensors (temperature, pressure, flow rate, displacement, rotational speed, stress, vibration, concentrations, geographic or spatiotemporal location, etc.). This work can help analysts to identify and define new early warning signs, surprises and potentially problematic assumptions. After investigation to assess their relevance, these new features can be added to the system monitoring framework.

These new technological promises are particularly relevant concerning technical data, but are also to some extent applicable to data concerning human and organisational factors of safety, such as data extracted from reports written in natural language. Correlations and outliers identified also require a cautious investigation approach.

---

[44] *https://emerj.com/ai-sector-overviews/predictive-analytics-in-the-military-current-applications/*

[45] *https://defensesystems.com/articles/2018/06/19/comment-dod-analytics.aspx*

An example application of this approach in information system security is given by intrusion-detection systems, which monitor network data and machine usage patterns for signs of new security (malicious use) threats.

Big data for detecting hidden correlations and Early Warning Signs – lessons from an IMdR project

The literature review undertaken for this project[46] suggests that weak signals are not intrinsically "weak" (Guillaume, 2011; Jouniaux et al, 2014). Rather, the notion of weak signal is an extrinsic property of an observation; it requires links between the observation and other information sources to be established, like a pattern in a puzzle. The interpretation process needed to establish a weak signal involves 3 steps: (1) *detection*: identification of a link between one observation and a scenario that impacts risk, (2) *relevance*: qualification of the link between the scenario of impact to risk and the risk modelling, (3) *amplification*: confrontation of the weak signal with safety objectives and means to deal with it. Some strong signals can be weakened as they lead to no changes or actions. A number of accident scenarios (the Concorde crash in 2000, Three Mile Island in 1979, Air Moorea crash in 2006, Paddington rail crash in 1999) were revisited with these principles. A big data case study was tested on a database of several tens of thousands of incidents with sixty fields of data to describe incidents. Pre-treatment of data using principal component analysis enabled to reduce to five the number of relevant parameters to search for correlations. The algorithm based on random forests (a classification algorithm based on decision trees) was able to confirm dominant parameters but also identified a number of correlations that surprised experts. The experts were unable to understand the underlying causality relationship or why some specific system state emerged, but the big data treatment provided a new line of investigation or assumption to be verified.

10.5.2 Monitoring effectiveness of safety barriers

Big data allows organisations to measure and monitor the effectiveness of individual barriers, using data on operations (for example sensor data for physical barriers, and semi-structured text data such as incident reports for organisational barriers). This type of integrity analysis is not fundamentally different in nature from earlier work by reliability engineers and safety managers, though the use of larger quantities of data and more sophisticated statistical analysis techniques can improve the effectiveness of the monitoring.

Monitoring unsafe behaviours in Chinese underground mines

In order to effectively predict and decrease the number of "unsafe behaviours[47]" in Chinese underground coal mines, safety specialists analysed unsafe behaviour data of 2220 coal miners between 2013-2015 (Qiao et al., 2018).

Thanks to data mining techniques (association-rule and decision tree), the analysis of unsafe behaviours in underground coal mines helped to identify which unsafe behaviours needed to be better addressed to decrease frequency of accidents occurring. The study concluded that the factors that influence the frequency of unsafe behaviours were training (less training, more frequent unsafe behaviour), attendance (less attendance, more frequent unsafe behaviour), experience (less experience, more frequent unsafe behaviour) and age (very young and more elderly, more frequent unsafe behaviour).

10.5.3 Safety investigation

The data collected by equipment, if it is stored, can help safety investigators to understand the sequence of events that preceded the accident and to determine whether similar conditions have occurred in the past.

---

[46] *Project P12-1 (2013) - Institut pour la Maîtrise des Risques, a French NGO, www.imdr.eu*

[47]. *This notion of "unsafe behaviour" is related to behavioural approaches of safety which have been criticized by workplace analysis of working conditions and real activities challenges within organisational constraints*

British Airways flight 38

A Boeing 777 crash-landed at Heathrow airport in 2008 due to a loss of thrust from both Rolls-Royce engines upon landing. The flight had taken a polar route, which led to the formation of ice crystals in the fuel. Upon landing, the temperature increase led to a slush of crystals flooding the fuel-oil heat exchanger and restricting the flow of fuel. The initial phases of the investigation found it difficult to identify the cause of the loss of engine thrust. "Data mining" techniques[48] were used to attempt to identify whether the flight had any specific features that differed from 175000 other Boeing 777 flights and which might explain the problem. The flight was found to be unique in combining low fuel flow during cruising and high fuel flow with low temperatures during approach[49]. A fix to the fuel-oil heat exchanger was implemented on all aircraft using these engines.

10.5.4   Condition-based maintenance

Traditional maintenance plans are either corrective (equipment is replaced once it fails) or time-based (maintenance is planned according to predefined component lifetimes based on statistical treatment of successes and failures of several component). The data collected from machinery and smart sensors embedded in a plant allows the implementation of another category of maintenance plan, condition-based maintenance, where replacements are planned depending on the degree of wear or corrosion of the specific pieces of equipment. Use of the approach improves the predictive ability of maintenance workers and allows them to optimize the availability of the component and the logistics of spare parts management.

A few examples illustrate applications of this approach:

- Aircraft engine manufacturers now collect large amounts of data from multiple sensors embedded in their engines[50], which is transmitted to ground-based engineering centres. The data allows them to detect problems requiring maintenance even before the aircraft lands (a process called "engine health management"). The complexity and importance of this data analysis leads to new business models where engine manufacturers retain ownership of engines and bill airlines per hour of engine operation, rather than selling engines outright.

- Power plant operators can be warned in advance of changes in the operating conditions of a unit that are not sufficient to trigger standard monitoring alarms (because they don't exceed predefined thresholds), but do indeed, through the presence of a correlation between unusually high and unusually low readings for example, point to serious problems that could have a safety impact.

- Railway operating companies receive real-time data from their rolling stock indicating the state of braking systems, batteries, compressors, doors, cooling equipment and toilets. The French national railway operator estimates that these tele-diagnostics technologies have allowed them to reduce maintenance costs by 20% and to improve availability.

- Predictive failure analysis in computer systems allows failure of system components to be anticipated by recording and analysing internal diagnostic indicators that are continuously produced by components such as storage drives, processors and fans. The Self-Monitoring, Analysis and Reporting Technology (SMART) mechanism available in most hard drives is an example of this process that is available even in consumer equipment.

EU Horizon 2020 SafeClouds.eu project

SafeClouds.eu is a project funded under the EU H2020 programme addressing "SOCIETAL CHALLENGES - Smart, Green And Integrated Transport". Participants were aviation stakeholders, airlines, IT infrastructure experts, universities, safety agencies and air navigation service providers. The project investigated the use of artificial intelligence (AI) techniques, such as deep learning and artificial neural networks, to analyse the precursors of safety events. According to the project coordinator, Paula Lopez-Catala, *"Understanding the precursors and potential risks that may lead to a safety incident is critical to complementing the traditional methods of monitoring safety, reviewing accidents and incidents and extracting lessons learned"*. The techniques and algorithms were customised and tested to be effective in

---

[48]*Data mining describes the exploratory process of finding patterns and knowledge within data. Predictive analytics then attempts to leverage that knowledge to make predictions about the future (attempting to forecast, anticipate, or infer).*

[49]*See UK Air Accidents Investigation Branch aircraft accident report 1/2020, available online.*
[50]*An aircraft engine in the Pratt & Whitney 1000G family (used in the A320Neo) includes more than 5000 sensors, generating data at a rate of 10GB/s.*

every safety scenario identified, including unstable approaches to terrain warning, mid-air losses of separation, and runway safety.

10.5.5   Structural health monitoring

The integrity of mechanical structures can be monitored by collecting and analysing large numbers of measurements over time (temperature, pressure, vibration, strain, electrical conductivity, mass flow rates). Sophisticated condition monitoring and anomaly detection systems, monitoring real-time flows of data from smart sensors, can enable the safe life extension of aging industrial facilities. By analysing multiple sources of data, these systems can reduce the false alarm rate, which is a significant barrier to the implementation of simpler anomaly detection systems.

Lessons from an IMdR project on Health Usage Monitoring Systems

This project[51] identified four functions in implementing HUMS – Health Usage Monitoring Systems: (1) acquire and treat data from equipment, (2) diagnose the state of equipment by analysing flaws and failures observed, (3) establish a prognosis of the equipment state (4) aid decision-making based on current and foreseen evolution of state. Behind the technical vision of HUMS, the data format and software languages and interfaces, attention to organisational and human factors that influence the design of HUMS as well as their use, in operation, maintenance and logistics is a key. To calculate remaining useful life, some approaches rely on physical modelling of equipment (model-driven), some are data-driven, and some experience driven (based on expert judgment with inference to cognitive ontologies), while others combine the three. The prognosis should be established with regard to: its time horizon, the application domain, the level of decision, the perimeter, the freshness of data, the dynamics of the phenomenon, the level of detail and input data available. The project led to a practical guide that helps identify the questions to address with lessons from transportation (aviation and rail), the military and energy production sectors.

10.5.6   Fraud detection

Banks have long been using big data analytics to analyse large, unstructured data sets of transaction information and communication records to identify anomalous behaviour (internal fraud, credit card fraud, money laundering, etc.). Similar anomaly-detection techniques are used in the early stages of pharmaceutical research and drug development, with data mining techniques attempting to identify correlations between consumption of certain substances and health effects.

10.6   Challenges and risks to the effective use of big data analytics

In the following, we discuss a number of specific new challenges and recurring issues that safety analysts face in attempting to use predictive analytics to improve the detection and analysis of early warning signals. We also discuss ned risks generated by the use of big data techniques.

10.6.1   Level of confidence in predictions

Despite the high levels of interest and increasingly widespread implementation seen today, big data analytics are not a magical solution to all risk management problems. For example, a large quantity of information, both historical and real-time, is available on earthquakes, yet their prediction is extremely difficult (Silver 2012); the results of the 2016 and 2020 elections in the USA suggest that foresight concerning complex social systems is very difficult to achieve.

Failure in foresight on the Snorre Alpha oil rig

Over the past two decades, many companies in high-hazard environments have started to use safety climate surveys. These quantitative data analysis methods can be seen as an attempt to apply statistical methods to measure and analyse certain organisational factors of safety. As an illustration, staff operating the Snorre Alpha offshore oil extraction platform in the North Sea were subjected to a standard safety climate questionnaire in 2003, and analysis of the data indicated no points of concern. Only a few months later, the rig suffered a blow-out, an incident with potentially very severe

---

[51] *Project reference P15-2 (2017) undertaken by the Institut pour la Maîtrise des Risques, a French NGO,* www.imdr.eu

consequences. The resulting investigation identified a number of serious concerns with the way tradeoffs between production and safety were managed on the platform and whether the organisational culture encouraged staff to raise concerns (Antonsen 2009). This example suggests that the value of safety climate survey data in predicting safety performance is very low (worse, it may even encourage managers to develop a false confidence in their site's safety culture), though they may on occasion help understand certain organisational weaknesses[52].

10.6.2   Data silos and the challenge of data interoperability

For technical, historical, political and practical reasons, data is often generated and stored in "silos", or activity-specific information-processing systems which do not inter-operate. Companies operate a range of systems and equipment, each one specialized for a specific technical function and domain of expertise and for a different purpose (reliability, safety, purchasing); each produces data in a specific format and underlying data model. Establishing bridges between these data sources, i.e. making them interoperable, to enable cross-referencing and analysis of correlations, or moving to a unified "data lake" architecture (see Figure 1), is often a significant technical challenge to the effective implementation of big data programmes. It may also constitute a political challenge, because the ownership of data is a source of organisational power.

This challenge is compounded for companies that rely on numerous suppliers and contractors to design and assemble products, since relevant data is owned by large numbers of companies within the supply chain and network of partner organisations.
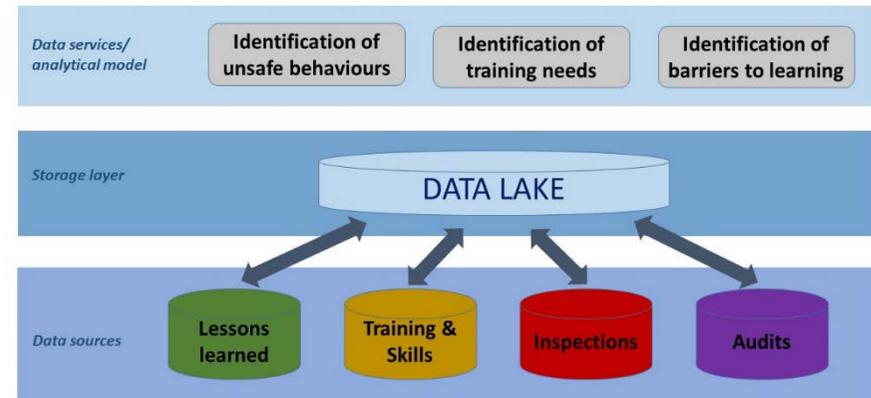
Figure 1. Example of data lake architecture.

Examples of interoperability challenges that companies could face when investing in a big data programme include (Scheerlinck et al., 2018):

- poor data quality;
- data protection considerations (confidentiality, privacy regulations);
- data sources with different data licences;
- requirement differences between data producer and user;
- difficulties in data integration, in particular when the number of data sources is high;
- rapid data integration difficulties linked to increased demand for near real-time analytics to generate insights in a timely manner;
- difficulties related to interfacing mechanisms between systems, e.g. incompatible communication protocols and data formats.

Further challenges include social dimensions such as vocabularies, different meanings and contextual backgrounds that co-exist with different sub-cultures (internally and externally).

---

[52] *For instance, the Baker panel report, is an independent audit of the five BP refineries in the USA; it has published after the accident at the BP Texas City refinery the 23rd March 2005; it has used questionnaires*

to help understand, ex post, a number of organisational weaknesses affecting safety management on the five BP refineries in the United States. http://sunnyday.mit.edu/Baker-panel-report.pdf

EASA's Data4safety initiative for the European aviation sector

The Data4Safety partnership initiated by the European regulator, the European Union Aviation Safety Agency (EASA) in 2017 aims to collect and analyse data from many organisations involved in aviation safety in the EU, including airlines, airlines, air traffic management service providers, national aviation authorities, aircraft and engine manufacturers, weather agencies. The project will allow centralized analysis of data that is currently fragmented across a large number of organisations.

10.6.3   Lack of expertise

Effective use of data to improve organisational performance and safety requires a combination of skills in statistical analysis, machine learning, programming, the use of new storage technologies and data retrieval and query technologies. Analysts must also understand the system and its safety barriers to appreciate potential safety impacts; they must also possess communication and storytelling skills to present results in a form which is understandable by decision-makers. Although the situation is progressively improving, there still is a significant lack of specialists with such skills (Espinosa et al., 2019). Such skills are sought after by companies in many industries that are building up multi-domain and integrated data science teams.

Given this range of competencies, which are rarely held by a single individual, the collective process to design and operate a big data system requires cooperation between domains of expertise and territories of responsibilities.

To address this skill shortage, some companies are also resorting to training their own staff and changing their traditional *modus operandi*. Implementing big data and automated natural language processing techniques are not just a technological change with new tools for users. It will lead to change of practices and organisations across the different interacting disciplines (Rousseau et al,

2018). Some companies are reluctant to adopt a big data solution due to the significant investment required.

To build confidence in the predictions and algorithms, it will be necessary that domain experts and data scientists[53], work together on the data samples to develop the performance indicators, and make several tests.

10.6.4   Black boxes

Some of the new classes of algorithms used to analyse big data, such as neural networks, originate in the field of artificial intelligence. These algorithms implement a form of "machine learning", being trained on large quantities of input data to optimize some specific output quantity, such as the ability to recognize faces in images, to group observations into clusters, to identify anomalies in a stream of data. The resulting models are "black boxes", since —unlike classical statistical models — the analyst is not able to inspect the model to understand why one event is being classified in a specific way or why one point in a time series has been highlighted as anomalous. For example, if a neural network misclassifies an image, it is very difficult to determine which specific features in the image led to the neural network's error[54].

The opaque nature of these models has several drawbacks, when compared with simpler statistical models such as regression models:

- When a potential early warning signal is detected (for example by a neural network trained for anomaly detection), typical tools do not provide an explanation for why the situations is judged to be anomalous, providing little assistance to analysts who must attempt a diagnosis.
- They do not help analysts to build and test mental models of system operation, nor to make "rule of thumb" checks on model plausibility.
- Regulators have no way of checking the model's internal validity or underlying assumptions.
- The legal system cannot inspect the logic underlying the model's predictions, in case of an accident.

---

[53] *IMdR project n°17-4 (2019) concerned "big data and reliability".*

[54] *Consider for example a problem that affected a machine learning system developed to identify skin cancer in photographs of skin lesions. The system was trained on photographs of skin that were labelled by dermatologists, some malignant (affected by cancer) and some benign. Unfortunately,*

dermatologists often include a ruler in a photograph of a skin, to provide a reference of scale, and rulers were often present in the photos labelled "malignant" and absent in the photos labelled "benign". The machine learning system therefore learned to detect rulers [Esteva et al 2017]. See also https://www.technologyreview.com/s/601860/if-a-driverless-car-goes-bad-we-may-never-know-why/ for a description of why this may lead to problems understanding the behaviour of autopilots in vehicles.

- Model predictions (related to the effects of implementing a new safety barrier, for example) may obtain less buy-in from decision-makers, because they do not help identify a plausible cause-effect mechanism and develop intuition concerning system operations.

Research into "explainable AI" (Hagras 2018) attempts to resolve these challenges related to transparency, lack of bias and fairness in the application of artificial intelligence techniques to decision-making in the public and private sphere.

10.6.5 Man-machine interface

In large organisations, the end users of big data analytics are generally not those who implemented the machine learning algorithms or the technological infrastructure that collects, stores and processes data. The data treatment process appears to them as a "black box". The man-machine interface should be carefully designed to help users in their activities and decision-making, to avoid a "master-slave" relationship between the technical infrastructure, the algorithms and the users. The system should be designed with affordances that help users understand the data treatment process and the underlying assumptions. Additionally, automatic steps enabled by algorithms and software should be carefully designed or limited to enable the user to decide between steps (Blatter and Dechy, 2018). The user should always have a questioning mindset and not blindly accept the outputs of the machine learning process. System designers should also be aware of the ironies of automation (Bainbridge, 1983). Users are not involved in the design and only act when problems arise, or when tasks cannot be automated, and must be handled in unexpected situations without understanding the algorithms.

10.6.6 Incomplete data – formal versus informal, tacit, and quality

Data is (quite obviously!) central to big data analytics, but some key questions concerning quality and relevance of data are sometimes overlooked by analysts in their haste to apply cutting-edge technologies and analysis techniques. Indeed, big data generates dreams of an ideal world in which visits to the shop floor and exposure to the sharp end of operations are no longer necessary, since all relevant information will be pulled into their dashboards. Some important characteristics of sociotechnical system operation such as perceptions, ideas, intentions, beliefs,

and non-verbal interactions will remain hard to cover using automated data collection mechanisms.

A fundamental issue to raise is the over-focus on data reported. What about data that is not reported? When some data is collected, what about its relationship to the context in which it was obtained? A lot of context is lost when formalising data and this context must be reintroduced or 'compensated for' (Koornneef and Hale, 2004) by the user in their activities and decision-making processes. There are fundamental elements that are not collected: tacit data, informal information, overlooked items. Not collecting them hampers human and organisational factors analysis (especially for root cause analysis). These factors can also dramatically impact quantitative safety analysis. The frequency of occurrence of some conditions and events can be underestimated.

Even the data reported and collected raises questions as well. The data which is collected on system operation is determined and constructed by the worldviews of the people who decide which elements are important to monitor; a narrow worldview may limit the analyses that can be undertaken[55]. This social constructivist perspective of reality should be acknowledged. It implies that even similar data or words, may have different meanings for different people and professional groups. Therefore, critical doubt will remain needed as a complement and check on data codification and automatic decisions based on algorithms.

Data elements are collected, filtered, validated, enriched, analysed and interpreted by multiple people at different phases of the data collection, processing, and decision-making process. The separation between all these people (different professions and objectives) can lead to deviations in the interpretation of the meaning of the data, which can be a source of risk. This phenomenon is particularly relevant concerning data obtained from automated textual analysis, because it is known that different professions or different sites of a same organisation may give different meanings to the same word. In addition, what to report as feedback and as data to collect is affected by political issues within organisations, managerial decisions, and the level of front-line confidence in the reporting system. It could lead individuals and work groups to withhold or under-

---

[55] *This is an instance of the "What You Look For Is What You Find" or WYLFIWYF problem described by Erik Hollnagel concerning incident investigation.*

report certain events to protect their professional reputation or avoid unwanted intrusions.

Safety analysis is often based on event reports. These reports contain not only textual data, but information and even knowledge from field experts and analysts. Contextual and historical factors, which might be critical to interpreting the textual content of a report, are difficult or impossible to handle using big data and NLP techniques (Rousseau et al, 2018).

Finally, the big data paradigm over-emphasizes the importance of data quantity, and can lead to a "shift from too few to too much data[56]" (Lannoy, 2018). It is known that event recording, especially for near-misses, is not as complete as for serious events, and is sometimes very poor with records registered that are only a few lines for an event in a database. Data quality and the level of coverage of events of interest inevitably impact the level of insight that can be generated concerning safety issues. Efforts for big analytics should be accompanied by renewed investment in the data collection process, not only about its quality, but also about its relevance in particular concerning the human and organisational dimensions of sociotechnical system operation (Dechy and Blatter, 2018; Rousseau et al, 2018).

10.6.7   Data analysis
The streams of data collected by big data infrastructures allow more dynamic analyses than in the past. Data analysis can become more specific and customised with digital twins.

The promises of these new techniques that can find patterns and models in the data may have side effects. Indeed, this added value should not lead to lack of prior analysis, knowledge modelling, formalising heuristics and expert judgment. Real-time thinking for designers and users will require them to have some knowledge readily available; they should not wait for models to emerge from data.

Implementing digitalisation and big data analytics can only work if there is a strong analytical program, with ontological efforts, to clarify rules and models for coding data, language and sense-making issues especially when preparing the machine learning (Rousseau et al, 2018; Dechy and Blatter, 2018).

In other words, big data analytics should not lead to 'small thinking' (Alloing and Monet, 2016). Rousseau et al. (2018) recalls that the digitalisation and big data issues are not fundamentally new with regard to the questions already raised in the 1980s and 1990s during the first AI and expert system wave.

10.6.8   Machine learning biases
The training data used to build machine learning models may lead to embedded biases which are illegal but difficult to identify. For example, if members of a particular ethnic group tend to have lower than average incomes, they may also have higher rates of incarceration.

Consider an insurance company which builds a machine learning model to estimate credit default risk by feeding input data from existing clients into a large neural network. This neural network may associate specific first names, which are highly correlated with the low-income ethnic group, with higher credit risks, embedding a bias within its decision-support tool that may produce legal problems[57].

10.6.9   Data governance and ethics
We live in a world in which each individual generates 1.7 megabytes of data each second (Petrov, 2020). This rapid data generation brings both opportunities and risks (AIHLEG, 2019). On the one hand, the big data analytics market is estimated to reach $103 billion by 2023 (Petrov, 2020).

On the other hand, there are many evolving risks in our digital landscape such as privacy and security risks (UNDG, 2017; Micheli et al., 2018). The predictive ability of machine learning models may lead to intrusions into people's privacy.

Against this background, data governance is of utmost importance to ensure data is effectively managed and used in an ethical manner. A number of principles for ethical use of big data analytics have been proposed to limit some of these threats (Schwartz, 2010; AIHLEG, 2019).

---

[56] *https://www.imdr.eu/offres/file_inline_src/818/818_pj_260419_164033.pdf*

[57] *See for example http://www.marketwatch.com/story/big-data-can-lead-to-big-legal-problems-for-companies-2016-06-01 and [O'Neil 2016].*

An annoyed customer walked into a 'Target' store in Minneapolis to complain about the store sending coupons relating to pregnancy products to his high school daughter. A few weeks later, the same customer apologized to the store manager: a discussion with his daughter revealed that she was in fact pregnant [Duhigg 2012]. It is worth noting that an individual's "data footprint" today in 2020 is more than 100 times larger than at the time in 2012.

10.6.10 Invalid conclusions

Appropriate use of machine learning techniques requires high levels of skills in causal reasoning, and subtle mistakes are easily made. Analysis of any large volume of data will very often identify a number of correlations between different variables. Some of these correlations will turn out to be spurious "flukes", and others will be due to the presence of hidden underlying variables, meaning that there is no causal mechanism which could motivate a safety intervention.

Medical research shows[58] that American men aged between 45 and 82 who skip breakfast have a 27% higher risk of coronary heart disease than other age categories over the 16-year followup period. This does not necessarily imply that eating breakfast reduces heart disease risk; the research also found that people who skip meals may have less healthy lifestyles than average.

In general, it is necessary to implement some form of experiment to check that changing the "predictor" variable does indeed lead to a change in the observed outcome variables. For obvious ethical reasons, this may be difficult to do for safety-related outcomes. The development of a critical view on the validity of inferences made is an important part of training in data analytics.

The Cost-Effective HealthCare project analysed emergency room data to try to improve treatment for patients with pneumonia symptoms. They aimed to build a system that could predict people who had a low probability of death, so they could be simply sent home with antibiotics. This would allow care to be focused on the most serious cases, who were likely to suffer complications. The neural network developed by the team had a very high accuracy but, strangely, it always decided to send asthma sufferers home. This conclusion was unexpected, since asthmatics are actually at high risk of complications from pneumonia. It turned out that asthmatics who arrive at the hospital with pneumonia symptoms are always admitted to Intensive Care. Because of this, the training data used to develop the neural network did not include any cases of asthmatics dying, and the model concluded that asthmatics were low risk, when the opposite was actually true. The model was very accurate, but if deployed in production it would certainly have killed people.

10.7 Conclusions

Big data analytics has significant potential to improve the detection of early warning signs of system failure, when compared with the use of standard statistical tools or of unassisted human monitoring of system data. However, while algorithms can improve human judgment, they will never replace it completely. Humans will remain important at every step in designing the process but also in its operation, in (automated or manual) data collection and in expert validation of correlations found. As correlations are not causation, investigation of assumptions made by big data analytics will remain a key activity.

Big data analytics can find patterns in data and derive models from data. This new opportunity does not reduce the importance of traditional analysis techniques, including modelling work to establish a cognitive representation of reality, analyse possible causal links, and extract expert decision-making heuristics. We should take heed from the lessons of the excessive optimism seen during the expert system era of the 1980s and 1990s: implementing big data techniques is not just a technological change that will magically produce results, but must be accompanied

---
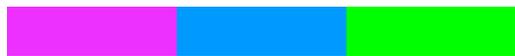
by critical analysis and expert assessment of the safety relevance of algorithmic predictions.

Effective use of predictive analytics faces a number of obstacles, and requires very specialized skills of the analysts who build the data collection and analysis tools. It also requires a critical viewpoint on the part of users, who will need to assess the relevance of predictions produced by the systems and counter the "black-box" effects of algorithms and integrate contextual factors that may not have been taken into account. The man-machine interface is of critical importance to allow step-by-step control of the data analysis process, back and forth, in a master-slave relationship. Automatic decision-making based on algorithms should be very carefully controlled due to many biases in quality of data (including under-reporting) and its treatment, the difficulty of handling data on human and organisational factors of system performance, which are often tacit and informal.

10.8   References

Alloing, C., Moinet N. (2016), *Les signaux faibles, du mythe à la mystification*, Hermès, La Revue, 2016/3, n°76, pp. 86-92.

Amanullah, M. A., Habeeb, A. S. M., Nasaruddin, F. H., Gani, A., Ahmed, E., Nainar, A. S. M., Akim, N. M., Imran, M. (2015). *Deep learning and big data technologies for IoT security*, Computer Communications, Volume 151, 1 February 2020, Pages 495-517, DOI : 10.1016/j.comcom.2020.01.016

Antonsen, S., (2009). *Safety Culture Assessment: A Mission Impossible?,* Journal of Contingencies and Crisis Management 17:4, DOI: 10.1111/j.1468-5973.2009.00585.x.

Artificial Intelligence High Level Expert Group – AIHLEG (2019), Ethics guidelines for trustworthy AI. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Bainbridge, L. (1983). *Ironies of automation*. Automatica, 19, issue 9, 775-779.

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

Dechy N., Blatter C. (2018), *Better addressing human and organisational factors in learning from experience: what benefits and challenges for big data and automatic language processing?* (in French) in Proceedings of the Institut pour la Maîtrise des Risques λμ21 conference, Reims, October 2018.

Duhigg, C. (2012). *How companies learn your secrets.* New York Times.

Espinosa, A. J., Kaisler, S., Armour, F. And Money, W.h. (2019). Big Data Redux: New Issues and Challenges Moving Forward, Proceedings of the 52nd Hawaii International Conference on System Sciences, HICSS 52, Maui, Hawaii, 2019. https://scholarspace.manoa.hawaii.edu/bitstream/10125/59546/0106.pdf

Esteva A., et al. (2017) *Dermatologist-level classification of skin cancer with deep neural networks,* Nature 2017; 542: 115-118.

Hagras H., (2018) *Toward Human-Understandable, Explainable AI* in Computer, vol. 51, no. 09, pp. 28-36, 2018. DOI: 10.1109/MC.2018.3620965

Ham, D.-H., Park, J. (2020). *Use of a big data analysis technique for extracting HRA data from event investigation reports based on the Safety-II concept*, Reliability Engineering & System Safety, Volume 194, February 2020, 106232, DOI: 10.1016/j.ress.2018.07.033

Huang, L., Wu, C., Wang, B. (2019). *Challenges, opportunities and paradigm of applying big data to production safety management: From a theoretical perspective*, Journal of Cleaner Production, DOI: 10.1016/j.jclepro.2019.05.245

Jouniaux P., Hadida D., Dechy N., Marle L., Billy F., Pierlot S., Parrennes F., Rouvière G., Husson D. (2014), *Detection, relevance and amplification of weak signals within the learning from experience*, (in French), in proceedings of the Institut pour la Maîtrise des Risques λμ19 conference, Dijon, October 2014

Khan, M. A., Uddin, M. F. and Gupta, N. (2014), *Seven V's of Big Data - Understanding Big Data to extract Value*, Proceedings of 2014 Zone 1 Conference of the American Society for Engineering Education (ASEE Zone 1), http://www.asee.org/documents/zones/zone1/2014/Professional/PDFs/113.pdf, ©2014 IEEE

Koornneef, F., and Hale, A. (2004), Organizational memory for learning from operational surprises: requirements and pitfalls, in Andriessen J. H., Fahlbruch B.,

How to manage experience sharing – From organisational surprises to organisational knowledge, Elsevier science, Amsterdam

Micheli M., Blakemore M., Ponti M. and Craglia M. (Eds.) (2018*), The Governance of Data in a Digitally Transformed European Society*, Second Workshop of the DigiTranScope Project, JRC114711, European Union, 2018. https://ec.europa.eu/jrc/communities/sites/jrccties/files/jrc_digitranscope_repo rt_-_oct_2018_data_governance_workshop_1.pdf

O'Neil, C., (2016), *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, ISBN: 978-0553418811.

Petrov, C. (2020), 25+ Impressive Big Data Statistics for 2020, Techjury Blog. https://techjury.net/blog/big-data-statistics/#gref

Qiao, W., Liu, Q., Li, X., Luo, X. and Wan, Y. (2018), *Using data mining techniques to analyse the influencing factor of unsafe behaviors in Chinese underground coal mines*, Resources Policy, Volume 59, December 2018, Pages 210-216, DOI: 10.1016/j.resourpol.2018.07.003

Rousseau, J.-M., Montméat A., Ben Ayadi M., Hébraud C. (2018), *Operating experience feedback and digitalisation – No pain, no big data!* (in French) in Proceedings of the Institut pour la Maîtrise des Risques λμ21 conference, Reims, October 2018

Rowley, J. (2007), *The wisdom hierarchy: representations of the DIKW hierarchy*, Journal of Information Science, 33(2) 2007, pp. 163–180, DOI: 10.1177/0165551506070706

Scheerlinck, J, Van Eeghem, F. and Loutas, N. (2018), D05.02 Big Data Interoperability Analysis, Report written as part of Action 2016.07, SEMIC (Semantic Interoperability Community) of the Interoperability solutions for public administrations, businesses and citizens (ISA²) Programme, Deliverable for task 5 - Capacity Building on Information Governance, Specific contract 508. https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/document/big-data-interoperability-analysis

Schwartz, Paul. M., (2010) *Data Protection Law and the Ethical Use of Analytics*. The Centre for Information Policy Leadership, Hunton & Williams LLP

Silver, N., (2012). *The Signal and the Noise: Why Most Predictions Fail – but Some Don't*. Penguin Books. ISBN: 978-1594204111.

World Economic Forum – WEF (2019), *How much data is generated each day?* https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f/

11 The Whistle-Blowers: Active-Actors in Foresight for Safety

Yves Dien, Collectif Heuristique pour l'Analyse Organisationnelle de Sécurité (CHAOS), France,
Paulo Maia, Energias de Portugal (EDP), Portugal,
Sever Paul, Agenția de Investigare Feroviară Română (AGIFER), Romania,
Sverre Røed-Larsen, SRL Health Safety Environment Consulting, Norway,
John Stoop, Kindunos Safety Consultancy Ltd, the Netherlands,
Eric Marsden Fondation pour une Culture de Sécurité Industrielle (FonCSI), France.

Executive summary

Even if industrial accidents are felt as a surprise, their investigation studies show that they do not occur by chance. They result from the degradation of safety. Issue to detect symptoms of degradation in order to act before the event. In addition to conventional foresight practices (i.e. "tools" for prevention) studies of accidents also show that, in many cases, some persons launched alerts about safety level decreasing. Take advantage of information provided by these persons could help for avoiding occurrence of accidents. Unfortunately, these whistle-blowers are not listened, not to say, worst, they are harassed or put aside by their management or even by their colleagues. Often, management either is unable or denies alerts, knowing that they make sense only after the event. Nevertheless, alerts and whistle-blowers have characteristics that allow them to be identified and differentiated from "moods" and "bad spirits". Remaining question, which is arising, is how to protect whistle-blowers from disciplinary sanctions and harassment.

11.1 Introduction

> "The freedom to speak the truth is one of the pillars of democracy"
> Polybius (c. 200 – c. 118 BC), The Histories, XII

> "O monstrous world! Take note, take note, o world
> to be direct and honest is not safe!"
> Shakespeare (1564 – 1616), Othello, III, iii

Current, relevant and interesting debates about industrial safety call into question the relevance of some concepts whose definitions and approaches have seemed, so far, to be widely shared. One such concept is that of safety. Does safety mean avoiding things that could go wrong or ensuring that things go right? Are causes of events to be found in failures, errors and malfunctions – the operational dark side – or should we consider that both expected and unwanted outcomes occur in the same way (Hollnagel, undated; Hollnagel, 2014)? In some countries, the concept of safety has been developed as a kind of "umbrella concept", covering both unwanted events within the safety field and intended events (security). Safety is also, in this new tradition, integrated in the modern SHE Safety, Health, Environment approaches and is strongly linked to risk and events which may occur both in the working environment and in the external environment and to both health risks of employees and to third parts.

In many of these discussions, the focus of safety approaches is still mainly on the avoidance of adverse events. In spite of undeniable progress in recent decades, many experts share the view that safety has reached an asymptote (Frantzen, 2004). Facing this problem, practitioners are trying to find new ways in order to improve safety management.

Does the problem arise in the same terms in the field of societal safety? First of all, we note that industrial safety is a part of the societal security domain that is global. In addition to technical failures, it includes protection of society and response to incidents, emergencies and disasters caused by intentional or unintentional Organisations are generally not a monolithic whole, a homogeneous entity. Sometimes, within the midst of the organisation, some dissident voices alert the powers-that-be about potential safety problems. Could these persons, whom we call "whistle-blowers", help to improve levels of safety? Could they help to meet the challenge of foresight for safety?

11.2 The Issue

Current Industrial safety approaches and practices mainly rely on two pillars: risk analysis and learning from experience.

Risk analysis can be broadly described as the process of risk identification and measurement. In that case, risk mitigation is a tool to avoid unwanted events or to minimize the impacts of their occurrence. Quantitative risk analysis seeks answers to questions such as the following:

- What are the events, with negative safety impacts, that could occur?
- What is their likelihood?
- What would be consequences of their occurrence?[59]

Risk analysis allows us to define the "*notionally normal starting points*" of the industrial process, meaning (i) "*initial culturally accepted beliefs about the world and its hazards*" and (ii) "*associated precautionary norms set out in laws, codes of practice, mores and folkways*"[60] (Turner and Pidgeon, 1997, p. 72). Because theoretical knowledge evolves with time, analysing risks is a continuous process.

In spite of substantial efforts in terms of methodology and successes in terms of results due to risk analysis, some events happen during production. These events are analysed in order to figure out causes of their occurrence and to determine and implement improvement(s). Industries, especially high-risk industries, have set up operating feedback systems for learning from experience. It is the second pillar of industrial safety approaches. Unfortunately, it seems that industries have reached a limit in terms of results. They hardly progress, they are ''*dancing a tango on asymptote*'' (Frantzen, 2004), meaning that, from year to year, numbers of safety records are more or less the same (either slightly higher or slightly lower). Does it mean that "learning from experience" is in a state of persistent deadlock?

Occurrence of an event can be described from two different points of view. On the one hand, the operating feedback system is **reactive** (the conventional approach), that is, an event is seen as a surprise, as an "exceptional set of unfortunate circumstances" (Finn, 2002). Nowadays, safety management is more foresight-oriented, considering a situation as "an accident waiting to happen", i.e., when we

are living during the "incubation period"[61] of an event. Indeed, "*[a]ny event is generated by direct or immediate causes (such as a technical failure or "human error"")*. *Nevertheless, its occurrence and/or its development is considered to be induced, facilitated or accelerated by underlying organizational conditions (complex factors) and some warning signals exist prior to the event*" (Dien, 2006, p. 148). So, the goal becomes to assess degradation of the safety level in detecting the warning signals, near-misses, and **weak signals**… In that sense, our operating feedback systems need to become **proactive**.

The concept of weak signals exists in several areas such as history, geology, medicine, acoustics… It was more recently coined by Vaughan (1996) in the domain of industrial safety after the space shuttle *Challenger* disaster: "*A weak signal is one conveyed by information that is informal and/or ambiguous, so that its significance […] is not clear*" (Vaughan, 1996, p. 355). Essentially, a weak signal is a symptom of a degradation of the production system.

Turner and Pidgeon (1997) describe these kinds of signals, "visible" during the incubation period, as a "set of events". They observed that these events go unnoticed. Indeed, unfortunately, even if detection and treatment of weak signals seems a promising way to go, it appears quite difficult to precisely define what a weak signal is. Its features are (Vaughan 1996):

- Qualitative (in contrast with quantitative);
- Subjective;
- Inconclusive;
- Giving partial information;
- Ambiguous, meaning several interpretations are potentially possible.

Furthermore, weak signals could be repetitive. In that case, repeatability itself is the criterion for identification. In this perspective, both qualitative and quantitative features are useful to validate its relevance in the context of the analysis (identification and selection) of weak signals. For example, a retrospective analysis of an accident can detect a weak signal based on its relevance (quality) and frequency (quantity) for a particular accident. Although it will not help prevent that accident, it can be useful to avoid similar accidents.

---

[59] *Qualitative risk analysis, as for it, uses words or colours to identify and evaluate risks or presents a written description of the risk*
[60] *Emphasis added.*

[61] *"Accumulation of an unnoticed set of events which are at odds with the accepted beliefs about hazards and the norms for their avoidance" (Turner and Pidgeon, 1997, p. 72).*

Detection of a weak signal relies on an engineer's feelings, intuition, perceptions rather than rational and scientific demonstration. In that sense, a weak signal is not in line with *"the norms of quantitative, scientific positivism"*[62] (Vaughan, 1996, p. 355). Indeed, it may even be in conflict with such norms and consequently, challenge the validity of such norms.

Furthermore, often, in terms of safety, a signal makes sense only after an event has occurred. In other words, the meaning of signs related to safety is not obvious, and organisations put in place systems for collecting and gathering signs that they do not really know what to do with except compiling statistics on accumulated data. Furthermore, companies have to cope with two concerns:

- Taking into account and treating a "wrong" signal (i.e., a signal that did not impact safety), which would lead to waste resources and time. However, such signals *could be* symptoms of other type of weaknesses or problems in companies, as weaknesses connected to the company culture, to shortcomings in leadership or management, to misconduct concerning social responsibility etc.
- Not detecting a relevant signal, which would be symptomatic of poor safety management and could lead to a major event.

So, here is a key question: Is it worth investing in the collection and treatment of weak signals, especially if we do not even recognise the weak signal? And here is another question: How should we define the relevant and accurate features of a weak signal?

The analysis of major events often shows that, in many cases, they were preceded by alerts, warnings launched by persons close to (or knowing) how a system technically functions.

Organisations are generally not a monolithic whole, a homogeneous entity. Sometimes, within the midst of the organisation, some dissident voices alert the powers-that-be about potential safety problems. Could these persons, whom we call "whistle-blowers", help to improve levels of safety? Could they help to meet the challenge of foresight for safety?

11.3   Definition of "Whistle-Blowers"

So far[63], there is no common legal definition of a whistle-blower, and a lot of different perceptions.

Nevertheless, before proceeding further, let's define the term "whistle-blower" (or whistleblowing). The implied definition mainly refers to the societal domain.

For Wikipedia, a "*whistleblower (also written as whistle-blower or whistle blower) is a person who exposes any kind of information or activity that is deemed illegal, unethical, or not correct within an organization that is either private or public*"[64].

For the British Government "*You're a whistleblower if you're a worker and you report certain types of wrongdoing. This will usually be something you've seen at work - though not always.*

The wrongdoing you disclose must be in the public interest. This means it must affect others, for example the general public"[65].

According to Near and Miceli (1985), whistleblowing is "*the disclosure by organization members (former or current) of illegal, immoral, or illegitimate practices under the control of their employers, to persons or organizations that may be able to affect action*".

Chateauraynaud and Torny[66] (1999) make a distinction between "prophets" whose message is future dedicated and "whistle-blowers" (denouncers) who condemn past and ongoing events. Nevertheless, in both cases, the aim is to avoid occurrence of unwanted events and/or negative outcomes.

ADIE (2008) added a notion explaining that a "*whistle-blower is anyone who discloses or helps to disclose fraud, irregularities and similar problems*". So, a whistle-blower is not only the one who acts, but also the one who supports.

---

[62] *Let's remember that when engineers of the space shuttle O ring manufacturer raised an alert concerning the performance of seals in cold temperatures, NASA decision-makers challenged them to prove it by quantifying their concerns!! (Vaughan, 1996).*
[63] *Chapter written in 2018.*

[64] *https://en.wikipedia.org/wiki/Whistleblower, retrieved March 31, 2018.*
[65] *https://www.gov.uk/whistleblowing, retrieved March 31, 2018.*
[66] *They were the first French scholars who tackled this issue. The French concept is "lanceur d'alerte" which means in a word-for-word translation "alert launcher".*

The Council of Europe (2014) considers that a whistle-blower is "*any person who reports or discloses information on a threat or harm to the public interest in the context of their work-based relationship, whether it be in the public or private sector*".

For the European Commission (2018), "*whistle blowers are people speaking up when they encounter, in the context of their work, wrongdoing that can harm the public interest, for instance by damaging the environment, public health and consumer safety and EU public finances.*"

For Edward Snowden[67] (2019), a whistle-blower is "*a person who through hard experience has concluded that their life inside an institution has become incompatible with the principles developed in - and the loyalty owed to – the greater society outside it , to which that institution should be accountable*". From his point of view such a person "knows that they can't remain inside the institution, and knows that the institution can't or won't be dismantled".

So, in conclusion, the definition of a whistle-blower seems to have developed in the last decades:

- From only related to internal company conditions/employees to issues related to institutions and organizations of public interest
- From narrow subjects (e.g. types of wrongdoing) to a wider group of threats or harms
- From a single actor to group action (supporters).

However, it is important to emphasize an important distinction: whistleblowing should not be confused with the statutory obligation of information established in a number of countries: many companies within different industrial branches, health and care institutions, transport operators, etc. and some occupational groups (e.g. doctors, nurses) have a special reporting obligation, including conditions that have safety relevance. In an attempt to distinguish between different reactions to negative working environment conditions among workers, a researcher distinguished between insulted employees, whistle-blowers, complaints and messages from employees favouring openness ("bell ringers").

11.4 Examples of Whistle-Blowers

Whistleblowing is not a recent concept. If we immerse ourselves in mythology, we already may find, in tales of ancient Greece, persons who warned their compatriots. Perhaps the most famous was Cassandra, Princess of Troy, daughter of King Priam and Queen Hecuba, who spoke true prophecies. Unfortunately, a curse struck by Apollo had the consequence that her true prophetic statements would never be believed. Laocoon, a Trojan priest, warned the citizens of the deceptive nature of the horse, but was killed with both of his sons by sea serpents, sent by Poseidon.

11.4.1 Whistle-Blowers in Industry

11.4.1.1 A Committed Nuclear Engineer

Let's return to our times, where we wish to draw your attention to a decision made in January 1996 by the US Nuclear Regulatory Commission (NRC) [68], to put the three units at the Millstone nuclear power plant (NPP) in Connecticut on the Watch List. This action allows the NRC to order the shutdown of a unit and to authorize its restart only under certain conditions.

This decision was motivated by serious unsafe practices in the operation of the plant (during the refuelling process). It was not the consequence of an incident nor **did it result from an investigation or an audit carried out by the Safety Authority.** It was the result of determined, voluntary and pugnacious action by a NPP senior engineer, named George Galatis. As early as 1992, he became concerned about the management of spent fuel that did not comply with regulatory safety requirements. He warned his hierarchy, but they did not take his alert into account. In the next two years, nothing changed, except that Galatis was isolated and bullied within the plant. In 1994, he took the initiative to directly alert the NRC, knowing that the NRC had been aware of the plant practices for the previous 10 years and had not taken any corrective action. Faced with the persistent apathy of the NRC, Galatis decided, in August 1995, and in connection with an NGO, to petition the NRC to suspend the Millstone I licence for 60 days and deny the company's request for an amendment of the regulatory requirements concerning

---

[67]*E. Snowden is a whistle-blower (see § 11.4.4.3).*

[68] *American Nuclear Safety Authority.*

fuel unloading. (Miller, 1995; Pooley, 1996). The pressure on Galatis redoubled, but the case became public, and the NRC was forced to react.

The "stubborn crusade" of this engineer earned him a long article and the cover of the American magazine TIME.

11.4.1.2   A Product Engineer Involved in Safety

On 3 March 1974, Turkish Airlines Flight 981 crashed over the Ermenonville Forest, north of Paris, few minutes after its taking off from Orly airport. The 346 people on board of the DC-10 airplane died.

The direct cause of the accident was an explosive decompression, due to a broken cargo door at the rear of the plane. It led to a collapse of the passenger compartment floor that cut all wires necessary to control the aircraft. The plane became uncontrollable and crashed to the ground.

A similar event had happened two years before. On 12 June 1972, the rear cargo door of American Airlines Flight 96 DC-10 blew off while flying over Windsor, Canada. Because they were fewer passengers (67 persons), decompression led to (only!) a partial collapse of the compartment floor with (only!) a partial restriction of the controls. In spite of the situation, the pilot was able to land safely.

Fifteen days after this event, Dan Applegate, Director of product engineering for Convair, a McDonnell Douglas subcontractor involved in the DC-10 design, wrote a document known as the "Applegate Memorandum". Applegate gave it to his immediate supervisor. In the memo, he mentioned some concerns. The long memo stated, among other things:

"*The potential for long term Convair liability has been causing me increasing concern for several reasons:*

- *The fundamental safety of the cargo door latching system has been progressively degraded since the program began in 1968.*
- *The airplane demonstrated an inherent susceptibility to catastrophic failure when exposed to explosive decompression of the cargo compartment in 1970 ground tests.*

[…]

"*Since Murphy's Law being what it is, cargo doors will come open sometime during the twenty-plus years of use ahead for the DC-10*"

[…]

"*I would expect this to usually result in the loss of the aircraft*"

[…]

"***it seems to me inevitable that, in the twenty years ahead of us, DC-10 cargo doors will come open and I would expect this to usually result in the loss of the airplane***" [69](Eddy et al., 1976, pp. 183-185)

Applegate's supervisor considered that it was needed to "*look the "other side of the coin""* (Eddy et al., 1976, p. 186).

Convair vice-president in charge of the DC-10 project convened a meeting to decide the company's policy regarding this issue. Convair management thought that changes requested from the memo would be costly and it was not sure which company would pay the bill (Convair or McDonnell Douglas). During this meeting, it was acknowledged that Applegate was closer than his supervisor to the engineering of the DC-10. Nevertheless, the reasoning of the supervisor was preferred and the "*interesting legal and moral problem*" was resolved "*by deciding that Convair must not risk an approach to Douglas*". […] *most of the statements made by Applegate were considered to be well-known to Douglas and there were nothing new that was not known to Douglas* (Eddy et al., 1976, p. 187). So, Douglas was never officially informed about Applegate's concerns.

11.4.1.3   A Field Journalist

On the night of 2 - 3 December 1984, a toxic cloud of methyl isocyanate (MIC) spread over the city of Bhopal, Madhya Pradesh, 600 kilometres south of Delhi. The cloud made its way especially into and around the shanty towns located near the Union Carbide India Limited (UCIL) pesticide plant. The disaster eventually created about 600,000 victims, including more than 12,000 deaths.

The cause of the disaster is still under debate. Nevertheless, we could assume that slack management leading, among other things, to deferred maintenance which created a situation where routine pipe maintenance caused a backflow of water

---

[69] *Emphasis added by authors Eddy et al.*

into a MIC tank, triggering the accident [70]. Before the accident, the plant was idling with reduced staff (Shrivastava, 1992; Lapierre & Moro, 2001).

Several serious events preceded the catastrophe. On 23 December 1981, a phosgene (toxic gas) leak occurred during a maintenance shutdown and caused the death of Mohammed Ashraf, foreman of the plant. Union Carbide concluded that the causes of the accident were two human errors. However, the trade unions claimed that the accident resulted from a deterioration of the plant's safety levels since the rules of procedure prohibited the storage of phosgene when the treatment unit was out of service. On 10 February 1982, a new gas leak occurred on a phosgene pump: 25 people were intoxicated[71]. Factory workers launched a strike.

Rajkumar Keswani, owner of and reporter for the local newspaper, the "Rapat Weekly", was an acquaintance of Mr. Ashraf. He wanted to know if his death was an accident or the consequence of internal failures at the pesticide plant. With the collaboration of plant workers, he was able to visit it illegally. After consulting scientific books, he came to the conclusion that "*tragedy was only a matter of time*" (Lapierre and Moro, 2001, p. 264). He also obtained results of an audit carried out in May 1982 by three engineers from the technical centre of the parent company in the United States. Its conclusions concerning safety of the plant were alarming. The audit report revealed hundreds of deviations from both operational and safety rules. He also underlined the high staff turnover, the lack of training and insufficient operating procedures.

With this information at the end of his investigation, Keswani tried to alert the public by writing a series of articles with prophetic titles:

- "*Please, spare our city*", on 17 September 1982. In this article, he warned: "*If one day misfortune happens, do not say you did not know.*"
- "*Bhopal: "we are all sitting on the crater of a volcano*", on 30 September 1982.
- "*If you refuse to understand, you will be reduced to ashes*", on 7 October 1982.

Keswani became a modern-day Cassandra. His articles gave rise to indifference and at worst to denial. Thus, the Madhya Pradesh State Minister of Labour said: "*There is no reason to worry about the presence of Carbide because the phosgene it makes is not a toxic gas*" (Lapierre and Moro, 2001, p. 266-269).

Bored by the attitude of his fellow citizens, the journalist left Bhopal shortly after, but before the tragedy of December 1984.

11.4.1.4 A Conscientious Operations and Safety Director

On 5 October 1999, two trains on the same track collided head-on at the Ladbroke Grove Junction a few kilometres west of Paddington Station, London. The accident cost 31 lives and injured more than 400 people.

A Public Inquiry was launched after the accident. The Investigation Commission chaired by Lord Cullen conducted a detailed and thorough analysis of the event. The immediate and direct cause of the accident was a signal (SN 109) passed when it was red. It brought to light that beyond the direct cause, the accident was rooted in the shortcomings of organisation and poor management of safety in this railway sector (Cullen, 2000).

The investigation showed in particular that the SN 109 signal had been passed eight times when it was red in the six years preceding the accident[72]. During this same period, 46 cases of signal passed at red were recorded in the railway zone of the accident.

The Commission of Inquiry noted the existence of a whistle-blower in the person of Mrs. Forster. She was the Operations and Safety Director of the rail company operating at Paddington. In February 1998, a train of her company passed the SN 109 signal when it was red. She was informed that a train from another company had also passed the same red signal in early August.

This information worried her. So, she wrote at the end of August 1998 to the chairman of a working group in charge of proposals for improvements in signal safety. She shared her concerns about the SN 109 signal and she asked what action could be taken "*to mitigate against this high-risk signal?*" In view of the dilatory

---

[70] *Union Carbide Corporation, owner of the plant at the time of the accident, claimed it was due to sabotage.*
[71] *Six other serious incidents, which led to a dozen victims (dead and wounded), occurred before the disaster. Some of these events were in connection with the MIC.*

[72] *It means that with this single signal, there is an annual risk of collision of 7.2%, that is to say, the risk of a collision every 14 years. It seems that, sometimes, even "scientific" data are not enough for an organisation to make the (right) decisions!*

response of the chairman[73] and his move to another position, she wrote to his successor to reiterate her concerns about "*a serious problem with drivers misreading signals*" in the Ladbroke Grove zone. The new chairman promised her "*a full risk assessment*" through a future study that a consulting firm would have to carry out. No contract was ever signed on the subject and the "new" chairman of the working group left office. Mrs. Foster wrote again to the third chairman four months before the accident. Her letter remained unanswered, the addressee confessing after the accident that "*he was not aware of the remit which had been given*" to the working group (Cullen, 2000, p. 117-118).

11.4.1.5   A Seismologist Warning about Tsunami

On 11 March 2011, a powerful earthquake struck Japan, triggering a tsunami and a nuclear accident. It was an earthquake with a magnitude of 9.0 on the Richter scale. The tsunami, with waves more than 10 meters, impacted a wide area of the Japanese north-eastern coast. It caused huge damage to buildings and infrastructure. The earthquake and tsunami caused great loss of life and widespread devastation in Japan. As of May 11, 2011, the death toll of the earthquake and tsunami is 14,981 dead and about 9,850 disappeared according to the Japanese police. Three months after the disaster, there were 23,500 dead and missing, with no hope of finding missing survivors.

The tsunami specially impacted 3 NPPs: From north cost to south, it was Onagawa NPP (3 reactors), Fukushima Daini NPP (4 reactors) and Fukushima Daiichi NPP (6 reactors). The anti-tsunami seawall of Fukushima Daiichi NPP (called Fukushima in the rest of the section) was 10 meters high, with about 6 meters above the sea level. The 15 meters high waves of the tsunami submerged the seawall. Waves flooded and totally destroyed the emergency diesel generators and every other power generation systems of the plant. The loss of electricity led to an insufficient cooling of the reactors and nuclear meltdowns in Units 1, 2, and 3 (from 12 March to 15 March). Loss of cooling also caused the pool for storing spent fuel from Reactor 4 to overheat (15 March). It is difficult to assess consequences of the nuclear disaster. Indeed, ionizing radiations and life of radioactive elements are a very slow decaying process that may take decades and centuries.

In 2009, the NISA[74] held meetings with panel of experts to discuss the safety needs of the Japanese NPPs. During the meetings, issue of tsunamis was never on the agenda. In 2007, an earthquake with a magnitude of 6.6 impacted the west cost of Japan. It caused radioactive leaks at Kashiwazaki-Kariwa NPP, owned and operated by TEPCO[75], as Fukushima, and water from a pool of nuclear wastes entered the Sea of Japan. When case of Fukushima NPP was addressed, the panel focused on earthquake. Dr Yukinobu Okamura, a respected seismologist, was invited to a meeting in order to present his findings. It was concerned because NISA did not see tsunamis as likely enough to be considered in the Fukushima area. Data used for preventing effect of earthquake were taken from the largest earthquake recorded in 1938 with a magnitude of 7.9. It caused a small tsunami and TEPCO had built a seawall able to stop this kind of tsunami. Okamura explained to the panel that this earthquake was not the biggest. An earthquake that occurred in year 869 was more important and Okamura did not understand why it was not mentioned. The TEPCO representative said that it did not cause much damage. Okamura disagreed and said that damage had been severe. Discussion were focus on earthquakes, not on tsunamis. Furthermore, for TEPCO the earthquake occurred in 869 was simply "historical" without certified data. Eventually, the safety report for Fukushima was approved. It did not consider the 869 earthquakes in model used for updating Fukushima safety guidelines (Clarke and Eddy, 2017).

We note that Okamura was not the only person raising concerns. For instance, a geologist, Masanobu Shishikura told the government before the Fukushima disaster, that north-eastern Japan was overdue for a huge wave (McKie, 2011).

11.4.2   Whistle blowing in the military
In the maritime and aviation sector, based on a long military tradition, operational feedback on safety matters to top management was established as a distributed and delegated responsibility of the captain and his officers. Such feedback covers both the individual level of Good Seamanship and Good Airmanship and the institutional level of legal disciplinary actions by a Maritime Court in case of misconduct. Social sciences have elaborated this good operatorship concept into High Reliable Organisations, while the Maritime Court concept evolved into

---

[73] *"I have commissioned a special study to determine what causes can be identified which contribute… I expect a report in the near future and this will ensure that effective solutions are identified for early implementation…"* However, no such report was ever produced.

[74] *Nuclear and Industrial Safety Agency, the Japanese Safety Authorities.*
[75] *Tokyo Electric Power COmpany.*

independent safety investigation agencies. Such operational feedback can be both prospective and retrospective.

11.4.2.1 The Joe Kennedy air crash: a prospective case

On August 12th 1944, a thunderous explosion destroyed a B-24 Liberator over the coast of England on its way to its targets in France, the underground Fortress of Mimoyecques launching sites of the German V3 missiles. The B-24 was part of Operation Aphrodite, equipped as a drone; a massive flying bomb with rudimentary control equipment for guiding the drone to its target. The pilots were supposed to bail out after bringing the aircraft up to 2000 feet where it was supported further along to its target by another aircraft. Several minutes short of the planned bailout, an electrical fault in the wiring harness of the warning device of the Liberator caused the 21.170 pounds of Torpex to detonate. The massive explosion dispersed the aircraft and the two airmen on board; Joseph Kennedy – the older brother of John F. Kennedy- and Wilford Willy. The shockwave almost brought down the trailing Mosquito flying 300 feet above and about 300 yards to the rear of the robot. The photographer on this aircraft was injured and the aircraft was damaged slightly by the explosion. The aircraft belonged to a unit under the command of the son of President Roosevelt who claimed to be aboard this trailing aircraft.

An electronics officer, Earl Olsen, who believed the wiring harness had a design defect, had warned Kennedy of this possibility the day before the mission. The electronics officer implored the 29-year-old pilot and veteran of 25 combat missions not to go on this one. He discovered a fault in the remote-control arming device aboard his aircraft. He warned Kennedy's Chief Officer to no avail. Olsen pleaded to call off the flight, but was ignored. Since hearing about John's exploits in the South Pacific, Joseph badly wanted to top his brother heroics. While John's patrol boat sunk by a Japanese destroyer, he led the 10 survivors to an island where they were eventually rescued. John received the Navy Cross for his actions and was celebrated as a genuine war hero. It started his political career and run for Presidency. Operation Aphrodite was a complete disaster. Of more than a dozen missions, only one plane caused damage to the target, and only because it

crashed somewhat close. In September 1944, Canadian troops raided the V-3 base and found it abandoned since July (Wikipedia 1).

11.4.2.2 The Aerolinee Itavia air crash: a retrospective case

On June 1980, an inexplicable explosion destroyed an Italian DC-9 aircraft of Aerolinee Itavia, flight 870. It crashed into the Tyrrhenian Sea near Ustica. All 77 passengers and 4 crew members perished. The crash remained a source of conspiracy for many years with reports that contradicted each other. Only on December 20th 2017, a former US military, Brian Sandler, felt confident after 37 years to speak up in public about the crash (Telegraaf 2017). Sandler recalled a return to the US aircraft carrier Saratoga of two Phantom jets without their air-to-air missiles, because they had shot down two Libyan MIG jets that had approached them aggressively. The DC-9 had been hit accidentally during the interception, but the American attack had been denied until his revelation, 37 years later.

11.4.3 Note on the Role of Whistle-Blowers in Industry

The role and importance of whistle-blowers in the domain of safety is not yet fully acknowledged[76]. For instance, Rajkumar Keswani (see § 11.4.1.3) is not cited in the accident analysis seen as a reference by scholars (Shrivastava, 1992). His action is "only" described in a general audience book (Lapierre and Moro, 2001). You could not find the name of Dan Applegate (see § 11.4.1.2) in the official accident report (Secrétariat d'État aux Transports, 1976): to know the existence of his warning, you must read a book written by journalists (Eddy et al., 1976). The same story has happened to the alert launched by Carlyle Michelson (see § 11.5): it is expressed in a technical report drafted for the NRC (Rogovin and Frampton, 1980) and not in the "official" report of the President's Commission on the accident (Kemeny et al., 1979).

We have also to note that it is difficult to find documentation in scientific literature about cases for which a warning was successfully listened and treated.

Taking whistleblowing into account does not belong to a statistical or probabilistic paradigm. Event occurrence and whistle-blowers belong to the domain of "outliers of the curve" treatment. It takes effort to dig as deep as possible during an event

---

[76] *One reason could be because event reports are anonymous (people are not named), disembodied. In numerous reports, not to say the largest majority, it seems that there is no human being with flesh and blood present at the time of the event! (see Llory, 1996).*

analysis to highlight the existence of whistle-blower(s). We assume that the game is worth the candle because events would be analysed in a more systematic way and it would allow us to define more precisely features or alerts of whistle-blowers.

National, monetary, corporate or military vested interests may obstruct a timely transparency in precursors and causes of accidents, preventing further analysis, research and development. An exception exists in the aviation industry, where Good Airmanship and High Reliable Organisations were created to institutionalize timely operational feedback without fear of retaliation or exclusion of the messenger (McCall 2017).

11.4.4   Whistle-Blowers in Civil Society

11.4.4.1   The "First" Whistle-Blower

To be whistle-blower should it be in our DNA? Can we educate ourselves in this direction? Should it be a civil obligation? Is it something good or bad to be a whistle-blower? What or who should determine us to become a whistle-blower? Is it a matter of courage? What risks can we assume?

After a decade since the end of the Second World War and practically the beginning of the "Cold War", more precisely on November 1, 1955, an armed conflict started in South East of Asia, in Indochina Peninsula. This is known as the "Vietnam War" or the "Second Indochina War". People from Vietnam, called it the " War of Resistance Against America". It is still known as the "American War" and in fact, it was a fight against the two parts of Vietnam, North and South. The army of North Vietnam (The Democratic Republic of Vietnam) was supported by the Soviet Union, China and other communist allies, and the army from South Vietnam (The Vietnam Republic) was supported by the United States, South Correa, Australia and other anti-Communist allies (Wikipedia 3, undated).

The period in which this war occurred, until April 30, 1975[77], was an extremely difficult one for the four successive Presidents. The whole period was characterized by big protests organised by anti-war associations and most of the Americans considered that the war was "unjustified", indefensible (Ely, 1990).

At the end of the 60s, the US Secretary of Defence, Robert McNamara, set up a Commission for drafting a realistic analysis about Vietnam situation ("a study"). Goal was to have an "encyclopaedic history of the Vietnam War". According to his point of view, this report, officially entitled "History of U.S. Decision-making in Vietnam, 1945-1968". He believed, he later said, that a written record of the key decisions that led to the U.S. involvement in Vietnam would be of great value to scholars. (Linder, 2011).

The Commission, was composed of specialists from the Pentagon, the State Department, academics and some "think tanks", such as RAND Corporation[78] (Wikipedia 4, undated). They had access to many documents and records, from the White House, the Secretary of Defence personal notes and from the CIA.

The study was finished by the beginning of 1969. The result of this work, 7000 pages contained in 47 volumes, showed that the USA was involved in that war by the Truman administration, which decided to offer military support to France in its colonial war against the Viet Minh. Also, the analysis revealed that the next three administrations (Eisenhower, Kennedy, Johnson), have intensified the war and made decisions which were hidden to the American people or which had, at that time, negative observations from the "US Intelligence Community" such as "the bombing of North Vietnam in 1965" (Encyclopaedia Britannica, undated).

Although, the study could have been a historical study, considering the disclosures made, the it was classified "Top Secret" by the Pentagon and only fifteen copies were published. The authorities "were worried" about the possible negative consequences if the public came to know about the whole output. (Linder, 2011).

Daniel Ellsberg was one member of the Commission. He was a strategic analyst at the RAND Corporation. In 1954, he enlisted in the U.S. Marine Corps and served in the army for three years. From his point of view, the study showed that the United States faced a difficult choice between "the bad and the worst" ". He was not optimistic regarding a victory of his country in Vietnam. (Linder, 2011).

In August 1969, Daniel Ellsberg met Randy Kehler, an opponent of the war, on a meeting of the "War Resisters International" (an international anti-war organisation) organized by the Haverford College. Kehler's speech, in which he showed his availability to go to prison (finally, he was charged with a federal

---

[77] *Fall of Saigon, the capital of South Vietnam. This event marks the end of the Vietnam War.*

[78] *RAND Corporation is an American non-profit global policy think tank created in 1948 by the Douglas Aircraft Company.*

crime") to fight against the war, impressed him and was an important event for his future decision (Wikipedia 4 & 5, undated).

In this context, considering that this war had done enough victims for all parts involved, and probably it would never stop, after Kehler's conference, he decided to do something, assumed any risk, including his act being considered as a crime and to be punished for that.

At the end of September 1969, he decided to copy that report and to give it to the press. In 13 of June 1971, the New York Times published parts of this study. (Encyclopaedia Britannica, undated)

The Washington Post and another 18 publications, did the same. They called them the "Pentagon Papers".

On June 28, 1971, Ellsberg publicly surrendered to the United States Attorney's Office in Boston. He admitting to have given the documents to the press.

He faced charges under the Espionage Act of 1917 and other charges including theft and conspiracy, carrying a total maximum sentence of 115 years. The trial began in Los Angeles on January 1973. Ellsberg tried to claim that the documents were illegally classified to keep them not from an enemy but from the American public. However, that argument was ruled "irrelevant". Ellsberg was silenced before he could begin.

In spite of being effectively denied a defence, Ellsberg began to see events turn in his favour when the break-in of Fielding's office was revealed to the Judge Byrne.

On May 9, further evidence of illegal wiretapping against Ellsberg by the FBI was revealed in court. Furthermore, the prosecution had failed to share this evidence with the defence

Due to the gross governmental misconduct and illegal evidence gathering, Judge the Judge dismissed all charges against Ellsberg on May 11, 1973 after the government claimed it had lost records of wiretapping against Ellsberg.

Ellsberg's action led to a decrease in confidence from the American people in the government, reducing also the influence capacity of the American authorities in that region. Finally, the war finished in April 1975, after the fall of Saigon (Wikipedia 4, undated).

Daniel Ellsberg is recipient of the inaugural Ridenhour[79] Courage Prize as well as the Gandhi Peace Award, the Right Livelihood Award, the Dresden Peace Prize, and the Olof Palme[80] Prize.

Daniel Ellsberg became an anti-war activist (against US-led war in Iraq, against US military action against Iran, …). He is also a support for American whistle-blowers (he says that the existence of WikiLeaks[81] helps to build a better government, he is taking part in demonstrations against Manning's incarceration[82], …) (Wikipedia 4, undated).

He is also concerned be nuclear weapons. For him: "*As long as the world maintains large nuclear arsenals, it is not a matter of if, but when, a nuclear war will occur*". (Canfield, 2017)

11.4.4.2   A Physician with a Global Vision

Irène Frachon is a French pulmonologist, posted at the Brest hospital since 1996. She developed a support activity for people suffering from arterial pulmonary hypertension. During early 1990s, as young doctor she trained in a unit of a hospital in (the suburb of) Paris. This unit was specialised in pulmonary arterial hypertension and was recently receiving and treating a number of young women with a lethal very high pulmonary hypertension. This disease was related to the consumption of the drug Isomeride® (which is an "appetite suppressant") for 20% and later for 30% of them. Isomeride®, manufactured by Servier laboratories, was commercially marketed in the 1980s and was withdrawn from the market in the late 1990s due to serious side effects, including heart valve disease and pulmonary arterial hypertension.

---

[79] *In 1969, Vietnam veteran Ron Ridenhour wrote a letter to Congress and the Pentagon describing the horrific events at My Lai – the infamous massacre of the Vietnam War – bringing the scandal to the attention of the American public and the world.*
[80] *Olof Palme was a former Prime Minister of Sweden, assassinated in 1986 while he was in office. The crime remains unsolved.*

[81] *International non-profit organisation, founded by Julian Assange, that publishes classified media provided by anonymous sources.*
[82] *Chelsea h Manning was an intelligence analyst assigned in 2009 to an Army unit in Iraq who disclosed to WikiLeaks about 750,000 classified, or unclassified but sensitive, military and diplomatic documents.*

In June 2006, Dr Frachon reads in a medical journal an article criticising the retention of a drug marketed since 1976 under the name of Mediator®[83]. Manufactured also by the Servier Laboratories, it should be prescribed for cases of non-insulin-dependent diabetes. Nevertheless, it is widely distributed in France[84], because it is prescribed and used as an "appetite suppressant" with recommendation as a simple adjunct to a diabetic diet. Mediator® is a drug of the same family as Isomeride®[85]. They are amphetamine derivatives.

In February 2007, she received a patient with an unusual pulmonary arterial hypertension. Looking at her prescription, she realised that the patient has been on Mediator® for several years. She then contacted her former Parisian colleagues who tell her that they do not know what to do because they have only a few observations. Together, they decided to report the facts at Afssaps (the French Health and Safety Authority for Health Products): it is a pharmacovigilance declaration. They also decided to make a scientific communication about the cases they will gather. But it would be complicated because there are not many cases, and it will be a little tricky and to attribute causes. Furthermore, they start to get a little scared because Servier laboratories are known for their legalism, bullying, and lobbying. So, she looked for accurate and detailed information about Mediator® in order not to get into trouble with laboratories (in case of wrong information or error). Laboratories explained that there was no relationship between Mediator® and Isomeride®, for both the chemical structure and the metabolic pathway. Actually, the molecules are not the same, but, according to Dr Frachon, to say that there is no chemical resemblance between the two is "to be in bad faith". Norfenfluramine is the active ingredient ("appetite suppressant") of one and the other of these products.

Throughout the years 2007-2008 Dr Frachon led an informal team that she set up. It was made of some cardiologists with whom she worked in the frame of arterial pulmonary hypertension studies. They alerted her to cases of valvopathy for patients who took Mediator®. Dr. Frachon doubted about pure coincidence. This opinion was shared by her colleagues in the hospital's IT department. In parallel that she goes to see in order to get the list of patients suffering from valvopathies

and thus exploit epidemiological statistics. Relationships appear between valvopathies and Mediator® intake.

In October 2008 and February 2009, she made two scientific papers about this issue at conferences. These communications did not receive a considerable echo.

Then a report of 15 cases is sent to Afssaps. The report is accompanied by an email saying, "*We are very worried, very concerned about these valvopathies*". The team was received for the first time by Afssaps on June 3, 2009. During the meeting, the team presented everything they have gathered with photos of the valvopathies. It also compares with the 1998 American dossier about Isomeride®, that shows the same pathologies. It also presents a dossier of a Spanish case in 2003: a patient who, under Mediator® had a very retracted valve. The Spaniards sounded an alarm by saying that they saw valvopathies under Mediator® as under Isomeride®[86]. So, the results of the study are shown to Afssaps committees. The committee members said: "This is very good, but we will have to do additional studies beforehand to confirm this causal link, we do not withdraw a Marketing Authorisation like this". Therefore, the team sets up a case-control epidemiological study. It consists in taking into account all the valvopathies leading to insufficiencies and requiring hospitalization at the Brest hospital from 2003. Then the unexplained valvopathies were isolated, assuming that it is among them that there is the best chance of finding this causal factor. Then the patients were questioned extremely rigorously to know if they had been exposed to the Mediator®, their surgeon was also interviewed, and their medical file was compared with patients who had valvopathies for commonly known causes. The team realises that it was 70% exposure to Mediator® (unexplained valvopathies) compared to 6% (valvulopathies for known causes). These results are shown, during an in-camera meeting in September 4, 2009 to Afssaps management, which rushed ... to criticise thoroughly the study because of – supposedly - bias in the study. The study is presented to Afssaps dedicated committee on September 29, and the study worries the experts of Afssaps. It is noted that results of another study are added to the first results. This second study was collected thanks to the

---

[83] *The Mediator® was initially prescribed as part of a treatment for diabetes before being recommended for weight loss (although it was not an "official" appetite suppressant).*

[84] *From 1976 to 2010, the Mediator® was a top selling drug: bought by at least 2 million consumers and with 7 million boxes sold per year. Main (most serious) pathology linked to the Mediator® are valvopathies for some people (and only those people).*

[85] *While drugs from the amphetamine family have been withdrawn from the market two years after the Isomeride® ban, the Mediator® is not concerned since it is not officially an appetite suppressant.*

[86] *Servier laboratories immediately withdrew its drug from Spain (probably so that the noise does not come to France !!). But they did not remove it from France.*

"arterial pulmonary hypertension network" that Dr Frachon activated during the month (Frachon, 2010; Bensadon et al., 2011; Hermange, 2011).

On September 29, 2009, based on these two studies, Afssaps voted that it is not possible to leave patients exposed to this risk. The effective withdrawal is decided to be effective on November 30, 2009.

Dr Frachon realises that there will be no information for public about the reasons of the drug withdrawal. As victims who will never be fully informed, it will make difficult for them to file a lawsuit in front of the Servier laboratories in order to obtain recognition of their responsibility and then compensation. She decided to write a book with the goal to have a document that was not too big, accessible and extremely well documented. The book was published in June 2010 with the title "Mediator 150 mg" and subtitle "How many deaths?". The censorship of the subtitle is ordered right after its released by a court at the request of the Servier laboratories. In January 2011, a judgement by a Court of Appeal annulled the censorship of the title of the book.

Regarding medical consequences of the scandal of Mediator®, at first, Afssaps advanced the figure of 500 deaths minimum. A latest forensic expertise on the subject estimates that between 1,500 and 2,100 people died of the adverse effect of this drug.

In 2011, Irène Frachon received a "Citizens Whistle-Blower" trophy award from an NGO whose role is to rehabilitate representative democracy, to promote ethics in politics, to fight against corruption and tax evasion.

Finally, the Servier Laboratories and the "National Agency for Drug Security"[87] are being sent to the correctional court for the Mediator case, on September 2017. They are charged respectively for "aggravated deception, fraud, injury and manslaughter and trading in influence" and "Injury and manslaughter".

11.4.4.3   A Citizen Sensitive to the Protection of Privacy

Edward Snowden is an American computer scientist who revealed the details of several US mass surveillance programs.

In 2006, Edward Snowden was hired by the Central Intelligence Agency (CIA) as a computer engineer to maintain the computer system's network security, having a top-secret security clearance. In 2009, he resigned from the CIA and is, then, hired as a contractor by Dell. He is assigned to National Security Agency (NSA) facilities (in Japan and in Hawaii). While he was working in Hawaii, he copies on a USB stick ultra-confidential information. In January 2013, he contacts Laura Poitras[88] and Glenn Greenwald[89] anonymously.

Edward Snowden fled to Hong Kong in May 2013. In June he flew to Moscow where he was granted temporary asylum in July[90]. In January 2017, his asylum was renewed for three more years

In June 2013, the British newspaper "*The Guardian*" started publishing some the revelations. Then, many newspapers from all over the world published leaks originally provided by Snowden.

The volume of documents transmitted is a controversial issue. According to different sources it is between 15,000 and 20,000; another estimate goes up to 1,7 million.

Snowden's leaks revealed, among other things, that:

- there were many NSA programs for mass surveillance of telephone calls and online exchanges;
- many foreign leaders are wiretapped;
- many data on "ordinary" American citizens were collected;
- although focused on surveillance for an anti-terrorist purpose, the NSA was also involved in economic and industrial espionage.

The US government indicted Snowden on three charges: theft of government property, unauthorised communication of national defence information and wilful

---

[87] New name for the Afssaps (which was reorganised after the Mediator® scandal).
[88] Laura Poitras is a director, documentary producer, journalist and American photographer. Her 2014 movie about Snowden history (Citizenfour was awarded the Oscar for Best Documentary in 2015).

[89]Glenn Greenwald is a political journalist, lawyer, blogger and American writer.
[90] In the meantime, he had applied for asylum in twenty-one countries.

communication of classified intelligence to an unauthorised person: the last two charges fall under the 1917 Espionage Act (Burrough et al., 2014; Lefébure, 2014).

11.4.4.4 The Panama Papers: Action of An Anonymous Whistle-Blower

Panama Papers report, in August 2015, the leak of more than 11.5 million confidential documents from the Panamanian law firm Mossack Fonseca, detailing information on more than 214,000 offshore companies and the names of the shareholders of these companies. Among them are politicians, billionaires, top athletes or celebrities and even mobsters and smugglers.

The documents were provided by an anonymous and unpaid whistle-blower (known only under the pseudonym of "John Doe"). "John Doe" chose to send information to Bastian Obermayer, a reporter from "Süddeutsche Zeitung", a German newspaper. The Süddeutsche Zeitung decided to analyse the data in cooperation with the International Consortium of Investigative Journalists (ICIJ)[91]. When "John Doe" contacted the Süddeutsche Zeitung, he put a couple of conditions for providing the information because he considered that, doing that, he put his life in danger:

- to chat only over encrypted files;
- no meeting, ever.

"John Doe" said that reasons why he leaked information was to make "*these crimes public*".

Volume of data leaked was approximately 2,6 Terabits – more than the combined total of the Wikileaks Cablegate[92], Offshore Leaks[93], Lux Leaks[94], and Swiss Leaks[95]. The data primarily comprises e-mails, pdf files, photo files, and excerpts of an internal Mossack Fonseca database. It covers a period spanning from the 1970s to the spring of 2016.

On 3 April, 2016, 109 newspapers, TV and online channels, published simultaneously, in 76 countries, first conclusions to inform about schemes used by Mossack Fonseca clients for hiding for money in some tax havens, for tax evasion and for money laundering. (Obermayer and Obermaier, 2016; 2017).

Because all these people and actions seems to have connections with the law office Mossack Fonseca from Panama, this investigation was called "Panama Papers" in reference to the "Pentagon Papers", and is probably, as Edward Snowden said, "*the biggest leak in the history of data journalism*" (on Twitter, April 3, 2016).

11.4.4.5 Social Media and Whistle-Blowing

The emergence of social networks (Facebook, Twitter, YouTube, Instagram, …) has changed the game in allowing to dissemination of information in real time on a large scale, thus bringing an audience to many citizens who claim themselves as whistle-blowers. In that case, social networks can be important relays for an alert receives a significant audience. For instance, in 2017, when an interim truck driver posted a video on YouTube showing that an iron and steel company spilling acid in a slag near the plant[96], he did not suspect for a moment that the images would become viral, be seen by millions of people, and thus trigger an environmental scandal[97].

So far, the alert launching and the warning role of the general public was rather played by a single person expert in a specific domain or by consumer or environmental organizations, as well as unions and journalists. But the current atmosphere of mistrust between people and "official institutions" has not spared

---

[91] *ICIJ is an international network launched in1997 by the Center for Public Integrity (non-profit investigative journalism organization). ICIJ was spun off in February 2017 into a fully independent organisation which includes more than 200 investigative journalists and 100 media organizations in over 70 countries (Wikipedia)*

[92] *In 2010, leak of 251287 diplomatic telegrams which were exchanged between nearly 300 embassies since 1966 (1.7 gigabytes of data).*

[93] *Name of an ICIJ report disclosing details of 130,000 offshore accounts in April 2013. It was seen as the biggest hit against international tax fraud of all times.*

[94] *Financial scandal revealing the content of several hundred very advantageous tax agreements concluded with the Luxembourg tax authorities by audit firms on behalf of many international clients such as multinational corporations Apple, Amazon, Heinz, Pepsi, Ikea and Deutsche Bank. The scandal is revealed in November 2014 following investigations by the ICIJ. The judicial aspect of the scandal*

[concerns people prosecuted by the Luxembourg courts for leaking documents and which resulted in 2016 in the conviction of the two whistle-blowers.]

[95] *Disclosure in 2015 of a giant tax evasion scheme operated with the knowledge and encouragement of the British multinational bank HSBC via its Swiss subsidiary. Disclosure has been triggered by leaked information from a computer analyst Hervé Falciani on accounts held by over 100,000 clients and 20,000 offshore companies with HSBC in Geneva.*

[96] *Liquid spilt was nots not iron mud, as indicated by the delivery notes, but "used acid". The cargo should instead be driven to a suitable recycling centre, an hour and a half far from the plant. A longer and much more expensive procedure.*

[97] *The truck driver was fired for "breach of commercial discretion". Since then, he has not been able to find another job.*

intermediate bodies. This phenomenon, associated with the development of social networks that allows all citizens to speak without intermediary, marks a break and explains the proliferation of ordinary committed citizens who behave as whistle-blowers, eager to engage themselves.

Social networks have taken up a lot of space today. It is an important relay so that an alert could receive a large audience. Being a whistle-blower leads to a kind of authenticity label for many causes, meaning that case/information provided is relevant and important.

In the past, the tools used by citizens to express themselves implied the use of traditional media, for example by being interviewed by journalists. Today, social networks can transmit information in real time, as evidenced by the many movies posted live by users on Twitter. The "traditional" media then intervene in a second phase, taking up some information to increase its visibility.

Citizens becoming whistle-blower take risks, which can put their lives in danger. Thus, Daphné Caruana Galizia, a Maltese journalist, columnist for several media and editor of a popular blog, denounced the corruption that ruled in Malta among politicians of all political parties, up to members of the Government. She was murdered in October 2017, in the explosion of her car, under which a bomb was put.

On the other hand, is denouncing an abuse or a danger on the web enough to make you a whistle-blower? Social networks have brought a lot of confusion up to misinformation. So, the Cambridge Analytica scandal showed that Facebook has decisively influenced the results of the US presidential election in 2016 and on the result of the referendum on Brexit in Great Britain in 2016 (Lewis and Hilder, 2018; Szadkowski, 2019).

11.4.4.6   Crisis Situation and Whistle-blowing

We have just seen that being a whistle-blower is not an easy position whether in a business or in civil society. One might think that this is partly due to the fact that

he/she warns of a hypothetical event for which the possibility of occurrence is difficult for others to comprehend. Do crisis situations restore the role of whistle-blower in giving it back the place it deserves? In other words, in a crisis situation, the dreaded event (the threat) has occurred. Therefore, the question is no longer whether the alert is relevant but to appreciate if the perception of the danger linked to the consequences is shared.

In December 2019, an outbreak due to the coronavirus covid-19 started in the Wuhan region of China. It quickly turned into a global pandemic leading to more than 2 million people infected and the death of more than 160,000 people worldwide[98]. This crisis gave the opportunity to see on the one hand a few whistle-blowers appear and in the other hand the behaviour of decision makers (Bodet and Chaverou, 2020; Johnson, 2020; Gafni and Garofoli, 2020):

- On December 30, 2019, Li Wenliang shared with former medical students, on a social network, the report sent by Doctor Ai Fen[99]. Two days later, in the middle of the night, he was arrested with seven other doctors for "spreading rumours" and "seriously disrupting social order". Questioned for several hours, he was forced to sign a letter of reprimand for spreading rumours on the internet. He must promise not to commit "acts contrary to the law". Only then is he allowed to return to work. On January 10, Li Wenliang cares for a patient with glaucoma, without knowing that she is infected with the coronavirus. He tested positive on February 1. Hospitalized two days later, he was transferred to an intensive care unit and placed on respiratory assistance. On February 6, Chinese national television CCTV and the daily newspaper Global Times announced his death, before removing this information from social networks following the denial of the Wuhan central hospital. A few hours later, the establishment confirmed his death[100];
- On March 25, 2020, Sergei Satsouk, editor of the online daily Ejednevnik in Belarus, was arrested and charged with "corruption", a crime punishable by ten years in prison. Three days earlier, Ejednevnik, well

---

[98] *Data from April 2020, paragraph writing period. At the end of the crisis, the figures will be much more significant.*

[99] *She is seen as the first whistle-blower regarding covid-19.*

[100] *Authorities are accused, on social media, of delaying the announcement of his death. Some criticize the Chinese government for covering the scale of the outbreak and demand more freedom of expression. On March 19, an official Chinese investigation disavowed the Wuhan police for having reprimanded Li*

Wenliang and his seven colleagues. In late January, the Supreme Court had already rehabilitated Doctor Li and other whistle-blowers in an article published in the press. Li Wenliang is now a national hero for part of the Chinese population for alerting colleagues when the virus first appeared, while the authorities sought to stifle his revelations. He was 34 years old. He and his wife were expecting their second child

knowns for its inquiries about the country's health care system, published an editorial questioning official statistic on the Covid-19 outbreak. The article also criticizes President Lukashenko's order to "deal with" the media covering the epidemic;

- Ruth Michaelson, a journalist with the British daily The Guardian, has worked in Egypt since 2014. On Sunday, March 15,2020, she reported in the newspaper on research by infectious disease specialists from the University of Toronto, as well as public health and information reports that indicate that Egypt is much more affected by the coronavirus than the government says. The day after publication, the journalist is summoned for three and a half hours by the Egyptian State Information Service. On March 17, Ruth Michaelson lost her accreditation. She was expelled from the country three days later. In Egypt, the government has stepped up censorship, officially to fight "fake news»[101];

- Ana Lalić, journalist for the Nova.rs news site in Serbia, publishes an article about the hospital in Novi Sad, in northern Serbia. Its title: "Voïvodine clinical centre at the breaking point, no protection for nurses". Ana Lalić describes there "a chronic shortage of basic equipment" and "chaotic" working conditions. On condition of anonymity, a doctor said that "the nurses rebelled and refused to enter the patients' rooms because there was no protective equipment". The article states that employees of the emergency centre and the intensive care unit, including those in the operating rooms, are only entitled to one protective mask per day. The hospital denied this information and filed a complaint against Ana Lalić for defamation, shouting her "indignation at the inaccurate, unverified and malicious reports" of Nova.rs. The day after her article, Ana Lalić is arrested by six police officers who search her apartment from top to bottom, seize her computer and mobile phone and make her undergo two hearings;

- In Great Britain, draconian measures were decided/applied in order to prevent some healthcare professionals discussing their work during covid-19 outbreak. Healthcare professionals are being silenced and threatened with disciplinary action for speaking out about their work during the coronavirus outbreak. Healthcare professionals are being silenced and threatened with disciplinary action for speaking out about their work during the coronavirus outbreak. Workers who have spoken to the journalists say they fear being disciplined. Several professionals said they worried about losing their jobs. Examples include an email signed by the chief executive of one NHS (National Health Service) trust forbidding all staff from talking to the media. In some cases, staff suspect emails and social media accounts are being monitored. Many NHS staff are increasingly concerned that their ability to share stories about their work is being restricted by a clampdown on speaking out publicly. It follows reports of doctors and nurses being gagged by hospitals and other NHS bodies from speaking out about widespread shortages of personal protective equipment. It has included threatening emails, the possibility of disciplinary action, and some people even being sent home from work. f suspect emails and social media accounts are being monitored;

- Brett Crozier, the commander of the American nuclear aircraft carrier Theodore Roosevelt, can be seen as the whistle-blower's symbol in the US Navy. It all started on March 30. After a stopover in Guam, Brett Crozier alerted his hierarchy about the meteoric progression of the disease on board the ship: a hundred men were affected out of nearly 5,000 crew members. "This will require a political solution but it is the right thing to do", Crozier wrote. "We are not at war. Sailors do not need to die. If we do not act now, we are failing to properly take care of our most trusted asset — our Sailors." Crozier also said that only a small contingent of infected sailors have been off-boarded. Most of the crew remain aboard the ship, where following official guidelines for 14-day quarantines and social distancing is impossible. Due to a warship's inherent limitations of space, we are not doing this," Crozier wrote. "The spread of the disease is ongoing and accelerating." Eventually, he requested "compliant quarantine rooms" on shore in Guam for his entire crew "as soon as possible". Few days later, acting secretary of the US

Navy, Thomas Modly, announced that captain Brett Crozier was relieved of his command of the USS Theodore Roosevelt, stationed in the Pacific, for showing "extremely poor judgment" by widely disseminating a memo about the coronavirus infection spreading quickly on the vessel with 4,800 crew members. Thomas Modly accused Crozier of "misrepresenting the facts" and took him to task for disobeying the chain of command[102]. Even the President ot the USA, Donald Trump, called the letter a "mistake" that had worried families and showed "weakness".

All these examples show that even in crisis situations whistle-blowers are rarely listened to or even worse, harassed undergoing censorship attempts to silence them.

We can notice that a crisis situation "favours" the emergence of whistle-blowers - here a worldwide crisis with whistle-blowers in many countries, whatever the political system. We also note the speed at which they disseminate information using massively social networks and the media (presence of journalists playing the role of whistle-blower.

The authorities' reaction is also very rapid and very strong in order to silence the whistle-blowers who are seen as troublemakers. In addition, they have little room for their defence with a lot of cover-ups, administrative decisions or "botched" trials. Whistle-blowers shake the certainties of authorities in power who fear that their authority and decisions will be criticized or called into question Is it really surprising? Not so much when we keep in mind that whistle-blowers point to system failures (or weaknesses) that only a few people agree to see, and that crisis situations are the moments when these weaknesses become "obvious" (visible).

11.5 Features of Whistle-Blowers and of Whistleblowing in Industry

This chapter has only addressed a few iconic cases. We could have talked more in detail about Carlyle Michelson, a nuclear engineer who worked part-time for the NRC and who took, in 1977, the initiative to study behaviour of the process in case of a small break in a specific location of the reactor primary circuit (top of the

pressurizer). Results were far beyond design (designers) assumptions, yet few people read about them. A reviewer in NRC prepared a memo based on Michelson's concerns and based on a previous incident that occurred at Davis Besse NPP (Ohio). Michelson's study and the memo did not circulate widely because the issue was not identified as a generic safety problem for operating plants. Eventually the memo was filed away (Rogovin, 1980). About one year later, a major accident occurred at the Three Mile Island NPP (Pennsylvania). The scenario was similar to that imagined by Michelson.

We could also have told about the story of Roger Boisjoly, one of the most well-known whistle-blowers in the "history" of industrial safety. He was a mechanical engineer at Morton Thiokol, the manufacturer of the solid rocket boosters for the Space Shuttle program. In July 1985, he wrote a first memo about their weaknesses, arguing that if, unfixed, it could lead to a catastrophe. He wrote several other memos on that matter, but no action was taken. On the eve of the launch of the 25th Space Shuttle flight, on 28 January 1986, he tried with some colleagues to convince the NASA management to postpone the flight because of the cold temperature. They felt that this would jeopardize the safety of the mission, and potentially lose the shuttle. No one listened to them. The Space Shuttle exploded 73 seconds after lift-off, killing the seven astronauts on board (Vaughan, 1996).

Even if the search for whistle-blowers is not yet a major concern of event analysis, we could still provide an outline of whistle-blowers and of whistleblowing features:

- "*Whistle-blower lacks a legitimate base of power for making the change*" (Near Miceli, 1985, p. 2);
- Whistleblowing deals with degradation of safety and could prevent occurrence of some events;
- Duration of an alert is variable: It can last days, months or years;
- A whistle-blower is either inside or outside the organisation (company / plant), but he / she is always close to the technical aspects of operations;
- The position of a whistle-blower in the organisation could be from the bottom (e.g., a field operator) to the top (e.g., a manager) and expertise.

---

[102] *Eventually, Thomas Modly has resigned, fallout from the ongoing controversy surrounding the Navy's handling of a coronavirus outbreak on the naval ship USS Theodore Roosevelt.*

The whistle-blower has the power of influence, but is not a decision-maker regarding the alert launched;

- For informing about the alert, the whistle-blower uses internal channels (within organisation), or (often then) the Safety Authorities, or the media, or NGOs;
- Alerts are technically oriented and safety oriented and they can be repeated, sometimes in different ways;
- Most of the time, alerts are issued by people close to the technical field or having information from field personnel;
- Warnings can be issued before events or can disclose information afterwards that was suppressed before the event.

We have to stress that alerts are not a scientific "expert opinion", since a whistle-blower is personally involved and ethically committed. Typically, an alert is not a simple denunciation since the alert is developing. This is not a prediction because an alert relates to the symptoms of deterioration of safety.

This first set of features might help to make a difference between alerts and background noise, i.e., to figure out relevant safety alerts among the numerous alerts that are launched.

11.6 Position of the Company Towards Whistle-Blowers in Industry

It seems that very often, companies are not ready to listen to alerts. There are several underlying cultural reasons explaining this "behaviour". They are summarized by the US CSB[103] (U.S. CSB, 2007, p.160):

- *"The incentives used in* [the] *workplace may encourage hiding mistakes."*
- *"[…] work under pressures that lead us to miss or ignore early indicators of potential problems."*
- *"Bad news is not encouraged."*

As we saw, very often, organisations do not listen to whistle-blowers[104]. Two apparent reasons that lead to this result are on the one hand the inability to

identify the relevance of alert, and on the other hand, the will not to detect or to identify the alert.

Several tactics to cope with whistle blowing exist:

When an organisation is unable to identify or accept the alert, it will have an attitude of denial in claiming that whistle-blowers are dissatisfied or displeased. The organisation will deny the risk (e.g., Keswani, Okamura) or engage in delaying tactics (e.g., Forster).

When an organisation does not want to acknowledge the significance of an alert, it becomes obstructive in isolating or bullying the whistle-blower (e.g. Galatis).

In every case, the implicit message is that the organisation denies the expertise and competence of the whistle-blowers.

We also note that, in some cases, whistle-blowers are isolated by their colleagues who consider them as "traitors" (e.g. Galatis, Boisjoly). Often, they have to leave the company they were working for. For example, George Galatis worked for the pharmaceutical industry and in the maritime sector after being pushed out of his job. Similarly, Roger Boisjoly resigned from the NASA and became a speaker on workplace ethics.

In recent years, a number of strands of research in the social sciences have emphasized the importance of diversity of viewpoints to effective decision-making, and have identified a number of features of organizational culture that encourage the expression of concerns and their effective management. The more normative recommendations that have emerged from this research are being integrated into management practice in various ways:

- The importance of requisite variety of expertise to learning within organizations, the risks generated by attempts to eliminate divergent opinions and the implications in terms of power struggles within companies have been described by researchers in organizational studies (Antonsen 2009).

- Researchers analysing what they called *high reliability organizations* identified the importance of features within the organizational culture

---

[103] CSB: **C**hemical **S**afety and Hazard Investigation **B**oard.

[104] *Unfortunately, as we already said, we do not have enough data concerning alerts listened and treated.*

including *chronic unease*, the opposite of complacency with respect to risks (Weick and Sutcliffe 2001), preoccupation with failure and sensitivity to nuances that can lead to failure. Managers are encouraged to actively seek out information about potential vulnerabilities in the system and to encourage front-line workers and experts to raise any concerns they may have related to safety.

- The nuclear power sector refers to an important attribute of the organizational culture, the *questioning attitude* (INSAG 4), which encourages people to continuously challenge existing conditions and activities in order to identify discrepancies that might result in error or inappropriate action.

- A study in the finance sector[105] pointed to the dangers associated with a *risk glass ceiling*, a situation where information on risks does not reach top-level managers who have the power to allocate resources necessary for risk prevention. The "glass ceiling" can be caused by a tendency in the organizational culture not to share bad news, by a position of safety units within the organizational chart which is too distant from decision-makers, and by top-level managers who do not make the effort to seek out contrary opinions and signs of organizational vulnerability. A report[106] by the UK Financial Reporting Council (concerning the audit profession) suggests that board members have a responsibility to leave the board room to speak with front-line staff and establish their own impression concerning risks facing the organization, as well as to ensure that internal reporting channels are working correctly.

- The notion of *psychological safety* (Edmondson, 1999) is important in encouraging people within an organization to speak up and raise their concerns. Psychological safety is a shared belief, within a group, that one will not be punished or humiliated for speaking up with ideas, questions, concerns or mistakes. Training courses such as the CRM programmes implemented in many high-hazard industries have been developed both to improve team managers' ability to encourage debate and the voicing of concerns within the work collective, and to foster workers' ability to speak up about their concerns.

- Firms are encouraged to establish whistleblowing policies and set up communication channels to allow reporting of concerns in a trusted environment[107].

For further reading on this topic, we refer to the chapter on resilience to explore theories and concepts of identifying and managing safety and risk.

11.7 Features of Whistle-Blowers and of Whistleblowing in the Civil Society

As for the whistles-blower in industry who are close to the technical dimension of operations that they deal with, whistle-blowers in civil society are close to the sources of information they disclose. They usually have direct access to the data.

Actions of whistle-blower (disclosure of "hidden" information) is often a matter of ethics. Indeed, the whistle-blower wants to share "material" with public opinion for making if know the non-ethical, not to say immoral, behaviour of the institution he/she is working for/dealing with/in contact with. In revealing the case they are concerned about, whistle-blowers know that they take personal legal risks (see § 11.8).

The channel used for dissemination of information is mainly specific media involved in investigative journalism. Another "tool" available to them begins to be widely used by whistle-blowers: it is WikiLeaks[108].

We also note that if monitoring or control authorities exist in the domain concerned (e.g. health system), the whistle-blower will first contact them. If the alert does not lead to "effects", then the whistle-blower will address the mass

---

[105] Report *"Roads to ruin. A study of major risk events: their origins, impact and implications"*, 2011, Cass Business School on behalf of Airmic.
[106] FRC report on *"Corporate Culture and the Role of Boards"*, available from https://www.frc.org.uk/directors/the-culture-project.

[107] An example of such guidance is the report *"Effective speak-up arrangements for whistle-blowers: a multi-case study on the role of responsiveness, trust and culture"* by the Association of Chartered Certified Accountants (ACCA), available online.
[108] WikiLeaks is a non-profit organisation, initiated in 2006, that publishes secret information provided by anonymous sources.

media. It may happen that contact with the mass media is provided through non-governmental organisations.

11.8 Position of "Society" Towards Whistle-Blowers in Civil Society

The position of "Civil Society" has a twofold aspect.

On the one hand, public opinion is grateful to the whistle-blowers for keeping it informed of the malfeasance of certain institutions, as these wrongdoings may have a more or less direct impact on health, on freedoms, etc.

On the other hand, the institutions in question have an attitude of denial, and bring lawsuits against the whistle-blowers. For instance:

- Servier Laboratories sued Dr. Frachon for publishing a book on the Isomeride® case (Frachon, 2010);
- Daniel Ellsberg and Edward Snowden were prosecuted by the American Administration;
- Antoine Deltour, Raphaël Halet and Edouard Perrin[109] have been sued by the Luxembourg justice for the Luxleaks affair[110];
- …

These prosecutions (legal proceedings) can create a halo, an atmosphere of threat, not to say of fear, and lead to a ***chilling effect*** which inhibits or discourages the legitimate exercise of information to the public about malicious acts or malfeasance.

11.9 Treatment of An Alert

An alert is not a prediction of when the event will occur. It is even not sure that an event connected to this alert will happen. Nevertheless, it is an information about a (potential) degradation of safety level. As such, it deserves to be treated at an institutional level. It means that it has to be listened to and investigated by a person with the authority/capability to make decision (whether in industry or in the civil domain).

The treatment of an alert has to begin as soon as possible after it is received through channels made available to whistle-blowers or used by them.

The duration of the treatment (i.e. release of the results, whatever they are) have to be issued as quickly as possible. Indeed, it is not possible to determine how far from the event the situation is. So, the quicker the potential corrective measures are implemented, the better. This neither means nor implies that the investigation must be neglected and / or carried out in a shoddy manner.

Furthermore, the whistle-blower must be kept informed (or better, associated) all along progress of the investigation.

11.10 Protection of Whistle-blowers

Protection of whistle-blowers is a recent concern. In the 1990s, the issue was still seen almost invariably in a hostile light. The term was most frequently used to describe public officials who had paid a heavy penalty for leaking information, usually to the media. Whistle-blowers were presented, if not as villains, as loners (Public concern at work, undated).

On the other hand, recent cases ("Dieselgate", "Luxleaks", "Panama Papers") have shown that whistle-blowers can play an important role in uncovering illegal activities that are detrimental to the general interest and well-being of citizens and society. So, their protection became an issue since a decade or so in several countries. A few examples are given hereafter.

The goal of protection is that the whistle-blower is not blamed or worst (e.g. pestered, laid off), because he/she released information about malfunctioning or fraud.

Some companies have put in place mechanisms that allow employees to issue alerts which could be seen as "shields" for whistle-blowers. For example, in:

---

[109] *Antoine Deltour and Raphaël Halet are the former employees of the consulting firm PricewaterhouseCoopers who have uncovered the large-scale tax optimization practiced by multinationals via Luxembourg. They had communicated the documents to journalist Edouard Perrin, a member of the organisation behind the revelations (an international consortium of investigative journalists).*

[110] *For this case, see footnote 92.*

- Offering "channels" to the alert launcher (e.g. confidential to the line manager, the human resources Department or the legal Department of the company, anonymous, directly to the management of the company through a dedicated Department);
- Setting up a Committee for Ethics and Social and Environmental Accountability to which any employee can report and submit cases in a confidential manner, if an employee considers there are breaches of the rules enacted by the company code of conduct;
- Setting up hotlines to warn of possible behaviour contrary to the values of the company.

It should be noted that very often in some companies, alerts should be limited to the violation of financial and accounting rules.

We note that, as an output of the 2011 G20 Summit in Cannes (France), the G20 leaders provided support to the compendium of best practices and guiding principles for whistle-blower protection legislation, prepared by the OECD, as a reference for enacting and reviewing, as necessary, whistle-blower protection rules by the end of 2012[111] (G20, 2011).

Hereafter, few examples of laws/regulations in domain of whistle-blower protection for some countries are given.

11.10.1 At the European (Union) level
The European Commission proposed (in 2018) a new law to strengthen protection of whistle-blowers across the European Union (EU). It will ensure a high level of protection for whistle-blowers who report violations of EU law by setting new Union-wide standards. This new law will see the establishment of safe channels for reporting both within an organisation and the public authorities. It will also protect whistle-blowers against dismissal, demotion and other forms of retaliation and will compel national authorities to inform citizens and train public authorities to accompany whistle-blowers.

The project proposes to set up internal structures to denounce illegal acts. These structures will have to be installed in companies with more than 50 employees, or with a turnover of more than 10 million euros, as well as in all public administrations, from the highest level (the State) to the municipality of more than 10,000 inhabitants.

The draft of the European executive will be submitted to the two European legislators, the Council (the member States) and the Parliament.

Until then, there was only a recommendation which states *"that member States have in place a normative, institutional and judicial framework to protect individuals who, in the context of their work-based relationship, report or disclose information on threats or harm to the public interest". To reach this goal, "the recommendation sets out a series of principles to guide member States when reviewing their national laws or when introducing legislation and regulations or making amendments as may be necessary and appropriate in the context of their legal systems"* (Council of Europe, 2014).

The EU law has been ratified April 19, 2019.

Every Member State of the European Union will have to transpose the Directive into their national legislation.

11.10.2 United Kingdom
The Public Interest Disclosure Act 1998 is an Act of the British Parliament which is a "shield" for whistle-blowers against detrimental treatment by their employer. It is an amendment to the Employment Rights Act 1996. It applies to cases where:

- a criminal offence has been committed, is being committed or is likely to be committed;
- a person has failed, is failing or is likely to fail to comply with any legal obligation to which he is subject;
- a miscarriage of justice has occurred, is occurring or is likely to occur,
- the health or safety of any individual has been, is being or is likely to be endangered;
- the environment has been, is being or is likely to be damaged, or
- information tending to show any matter falling within any one of the preceding paragraphs has been, is being or is likely to be deliberately concealed.

---

[111] *The compendium was endorsed by the G20 Anticorruption Working Group.*

Nevertheless, despite undeniable progress, some workers do not (yet) qualify for the whistle-blower protection (e.g. Jobseeker, volunteers, interns, non-Executive Directors, foster carers members of the armed forces and security services, self-employed workers, …).

It has also to be noted that a disclosure of information is not a qualifying disclosure if the person making the disclosure commits an offence by making it.

Against the background in the 1990s of both serious accidents (sinking of Herald of Free Enterprise, the Clapham rail crash) and an expressed negative attitude to whistle-blowers (Villains/loners) organizations to support and advise whistle-blowers and business firms/industry companies were established. Example of such an organization is e.g. "Public Concern at Work" in 1993 (www.whistleblowing.org.uk). In 2018, a group of parliamentarians formed an All-Party Parliamentary Group/Whistleblowing (APPG/Whistleblowing), which has proposed a series of reforms to update and modernize the British legislation (including expanding the whistleblowing concept to include all citizens and establish an Independent Office for Whistleblower).

11.10.3 France

The law of 9 December 2016, known as the Sapin 2 law, relating to transparency and the fight against corruption, created a protection for the whistle-blower by requiring companies with more than 50 employees to set up procedures for collecting data and reports issued by employees or by outside and casual collaborators.

The granting of the legal status of whistle-blower is subject to the following conditions:

- Have personal knowledge of the facts;
- To act in good faith;
- Do not profit or draw compensation from the alert issued;
- Do not try to harm.

If his / her status of whistle-blower is acknowledged, the person concerned will benefit from special protection. The law provides that no person may be excluded from a recruitment procedure or access to an internship or a period of professional training, no employee may be sanctioned, dismissed or subject to a discriminatory

measure, direct or indirect after an alert. Nevertheless, some trade unions and some NGOs consider that Sapin 2 has some limits:

- The whistle-blower must denounce the facts by internal channels of the company;
- Legal persons (a company, an association, a trade union, etc.) cannot have the quality of whistle-blower;
- The whistle-blower cannot be defended by a staff representative or a union in its alert procedure.

As a consequence, in France, 42% of executives, out of the 36% who witnessed "illegal or unethical practices", did not report these abuses. If they did not warn about these shortcomings, it is because they do not have confidence in the declaration procedure. Furthermore, 51% of executives consider that it is risky to denounce unethical practices in their business. (AFP 2019).

The European Directive should address shortcomings of the current French regulation.

11.10.4 The Netherlands

In the Netherlands, investigating accidents as a timely feedback of safety mishaps in reality has a long history of about 100 years. Started in 1909 in the maritime, other modes of transportation followed in inland shipping, (1931), aviation (1937) and railways (1956). In 1999 the call for an independent and permanent investigation agency lead to the establishment of the intermodal Transportation Safety Board. Due to a series of major disasters -in particular disco fires, firework explosions-, the scope of independent investigations was extended to other industrial and social sectors by establishing the Dutch Safety Board in 2005. There was felt no need to further institutionalize whistle blowing after establishing independent safety investigation organisations. However, mishaps and deficiencies triggered several cases of whistle blowing, in particular in the military and public governance sectors which were not covered by the safety board mandate. At the initiative of the chairman of the Dutch Safety Board, a separate arrangement for individual whistle blowing was established in 2016 with the Whistle Blowing Clearing House Act, providing a clearing house for re porting of mishaps and abuse. This system for individual whistle-blowers however, has not yet functioned properly due to a mismatch between staff and functions and close working relations with the Ministry of Internal Affairs (Wikipedia 2). Consequently,

in the Netherlands, facilitating whistle blowing has been successful only on an institutional level by the establishment of independent safety investigation agencies, albeit restricted to a retrospective approach.

11.10.5 Norway

Norway has a long history of whistleblowing – from individuals within a number of areas; such as industry, health and social sectors, the education sector, the armed forces, civic activities and government management. But it was only in 2007 that rules on whistleblowing were established in the Working Environment Act. The regulations included the right to whistle-blow, the way in which whistleblowing should occur, and prohibition of retaliation (which includes the right to remedy). In the present legislation, there is a duty for employees to notify in some circumstances: all workers are obliged to notify if a colleague is discriminated against or harmed or in conditions that could endanger life and health. Elected representatives at the workplace should assist workers who alerts. A representative may also notify on behalf of the notifier so that the notifier may be anonymous to the employer. A safety representative (a legally mandatory function in companies) has self-notification obligations in certain cases, such as concerning injury and illness and the risk of life and health. The regulations were partly changed in 2017. Now, businesses with more than five employees are also required to develop internal notification procedures. In addition, there are separate rules in the Equality and Discrimination Act: there is a prohibition of retaliation against anyone who has complained about discrimination, sexual harassment or any other form of harassment, such as ethnicity, disability or sexual orientation. There are also separate rules for employees in shipping.

The present legislation is undergoing change. A public exposition (NOU 2018:6) has been out for public consultation, and the Government promoted in April 2019 proposals to the Stortinget (The Parliament) concerning changes in the law, including that the scope of the field would be expanded to include certain groups that are not workers, clarifications of key concepts, sharpening the employer's duties after receiving notice and introduction of objective liability for financial loss after retaliation.

An association, the Zola Association (based on Emile Zola's defence of Dreyfus and famous article – J'accuse-) has since 1998 annually given the Zola prize to "people who openly and unafraid have uncovered or counterworked conditions that threaten human dignity, democracy and rule of law in Norway ", including many whistle-blowers.

11.10.6 Romania

Through law of 14 December 2004, No. 571, have been regulated some measures for the protection of persons who have complained or noticed violations of the law within the public authorities, public institutions and other units, committed by persons with management or execution functions in the authorities, public institutions and other budgetary units. (Parlamentul României, 2004).

The warning had to be made in respect of any act that involves a violation of the law, professional ethics or principles of good administration, efficiency, effectiveness, economics and transparency.

We have to note that the law covers only the worker from the public sector and not those from private sector. The general principles refer more to corruption facts or offenses against the financial interests of the European Communities or public institutes.

The protection of whistle-blower refers to:

- The presumption of good faith, until proven otherwise;
- In case of a disciplinary inquiry for a whistle-blower, the inquiry commissions should invite the press and a representative of the union;
- Identity protection if the denounced is a direct or an indirect manager or he have control or evaluation tasks for the whistle-blower.

11.10.7 Portugal

Currently, Portugal only provides partial protection to whistle-blowers from specific sectors of economic activity (e.g., financial services) or certain categories of employees (e.g., civil servants).

The law no. 19/2008 of 21st April, that set out measures to combat corruption is the only one that explicitly specifies provisions concerning the protection of whistle-blowers. However, this law only refers to offences in general terms, failing in specifying which ones are included. On the other hand, as this law addresses measures to combat corruption, it can be inferred that the offences should be related to corruption and alike. So, if the term corruption is used here in its broad sense, besides the legal and criminal sense, it may also include abuse of power,

damaging management, financial participation in a business, money laundering, embezzlement, unfair advantage or influence peddling. In this context, it can be assumed that the option by the term 'offences' in the law is a deliberate decision of the legislator to encompass a wide range of crimes related with the corruption world. In addition, the legislator option by a generic term ('offences') may also include any type of irregularities that the employee may be aware of by having access to the information through the exercise of his employment, profession or duties. By way of example, this includes criminal offences such as tax-related ones, prevarication or even sexual abuse committed by an employee's colleague or a hierarchical superior, as well as administrative offences, such as labour or environmental infractions.

In all situations mentioned, the denunciation can cause damages to the whistle-blower, including disciplinary action up to dismissal, as well as when the infraction is an offence of passive corruption, in which the denunciation would deserve the same type of protection. Although it is the most appropriate understanding to reach a wide and effective protection of the whistle-blowers, this is not the solution provided for by the law no. 19/2008 of 21st April. (J.A.A. de Matos Ramos, 2018).

11.10.8 United States of America

As far we know at this moment, the history of whistle-blower protection in USA started on July 30, 1778, when the Continental Congress, enacted by "*unanimous consent*", America's first whistle-blower protection law (Snowden, 2019). The law declared that "*is a duty of all person in the service of the United States, as well as all other inhabitants thereof, to give the earliest information to Congress or any other proper authority of any misconduct, frauds, or misdemeanors committed by any officers or persons in the service of these states, which may come to their knowledge*".

The United States has an ambiguous policy towards whistle-blowers. Famous whistle-blowers have been prosecuted and jailed like Chelsea Manning[112], and yet there is a US law since 1989, the Dr. Chris Kirkpatrick Whistle-blower Protection Act[113] (GPO, 1989). This law protects federal government employees from

retaliatory action for voluntarily disclosing information about dishonest or illegal activities occurring in a government organization. This text reinforces the protection of US public service agents if they denounce activity of their administration, in violation with laws or regulations. The law was completed in 2007 and allows federal agents to submit evidence of violation of the law, heavy waste of public money, abuse of authority, danger to health or public safety. There are restrictions to this law. The law does not apply neither to employees of the Federal Bureau of Investigation (FBI), neither to employees of the Federal Police, nor to employees of the National Security Agency (NSA). It was again amended in 2017 with the aim of providing greater whistle-blower protections for Federal employees, increased awareness of Federal whistle-blower protections, and increased accountability and required discipline for Federal supervisors who retaliate against whistle-blowers.

On the other hand, if a person works in a company that is fraudulent to the American tax authority, he (she) is strongly encouraged to launch the alert, to denounce his(her) employer. The US tax office promises informants up to 30% of the amounts recovered through their information.

In private companies, the protection is often less effective than for federal agents. When whistle-blowers lodge complaints against their employer reprisals, the legal proceedings are often slow and inefficient.

11.10.9 Remarks on Protection of Whistle -Blowers

Legislations protecting whistle -blowers are recent. Their goals is mainly to ensure transparency and to fight against corruption and embezzlement of (public) funds. Many conditions must be met to be considered a whistle-blower and, therefore, "protected".

Moreover, with regard to the European Directive, it is not clear to what extent it conflicts with the Directive (EU) 2016/943 (June 2016) which deals with the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure. Indeed, information

---

[112] *Chelsea Manning (born Bradley Manning) is a former United States Army soldier. She was convicted by court-martial in 2013 of violations of the Espionage Act and other offenses, after disclosing to WikiLeaks nearly 750,000 classified, or unclassified but sensitive, military and diplomatic documents. She was imprisoned between 2010 and 2017.*

[113] *Chris Kirkpatrick was a Department of Veterans Affairs psychologist and whistle-blower who committed suicide by gun on July 2009, he was fired from the Tomah Veterans Affairs Medical Centre in Tomah (Wisconsin). Previously, he complained about overmedicating patients at the facility.*

revealed by a whistle-blower can be claimed as "sensitive" information by the incriminated institution.

Despite legal protection of whistle-blowers, it is possible to inflict damage to the reputation of individuals and organisations that serve the goal of feedback and foresight.

Recent research suggests that the proportion of whistle-blowers exposed to reprisals and retaliation is increasing — despite more or less explicit law protection. And although whistleblowing can happen via multiple channels, which has a positive effect on the scope of whistleblowing, the vast majority of whistle-blowers reports internally and to the immediate leader. Only about 2% of the whistle-blower notices to the supervisory authority as the first step.

Furthermore, the protection of whistle-blowers can be thwarted as long as their positive contribution in security (of citizens and of industrial processes). For example, some academics (e.g. Amalberti, 2013) question the role of whistle-blowers, and in particular the one played in detecting weak signals, arguing that alerts are probably more useful to the social positions of those who raise them than to the risk analysis itself.

Heroes or villains?

The role of whistle-blowers as adversary opponents to a consensus perception of what went right is frequently submitted to framing, blaming and shaming of the individual whistle-blower. Killing the messenger remains a persistent tradition after the Greek example of Laocoon.

In our times, this blaming, shaming and framing may even create an inverse picturing of the actual role and function of safety investigation agencies (Wilson and Straker 2018). In the movie 'Sully" on the ditching of UA 1549 in the Hudson river, the director Clint Eastwood transformed the role of the NTSB into the role of villain. Despite of the proof of the contrary, given by both the NTSB report on case UA 1549 and Sullenberger in his book "Highest Duty: My search for what really matters". In the movie, the NTSB was portrayed as discrediting the role of captain 'Sully' Sullenberger: Sullenberger, meanwhile was portrayed as the All-American Hero in saving the lives of the crew and passengers of flight UA 1549. The NTSB was chosen to serve the role of villain. Both the professional pilot community and the general public raised doubts about the integrity and credibility of this US governmental organisation. As a consequence of such portraying as villain, the reputation of the NTSB as an independent and blame free safety assessor was jeopardized. (Wilson and Straker 2018).

In the realm of safety, namely occupational safety, two administrative measures would help in protecting the whistle-blowers and possibly to alleviate some of the prejudice burden associated to them. The first one would be based on a communication channel/reporting system that would enable any employee to report (anonymously or not) a safety observation, intended to correct, improve or eliminate a potentially dangerous or harmful situation related to the health and safety at work. In this context, safety observation means a communication tool available to all employees and subcontractors to report hazards or to propose safety recommendations in the workplace. The system would then analyse the safety observations submitted and propose the adequate actions, whenever appropriate. Publicizing internally all the safety observations received and the measures adopted for each one, would demonstrate that any opinion is important and motivate others to participate. The second measure, especially targeted for large-sized enterprises, would include the appointment of a safety ombudsman, in line with the ethics ombudsman figure already available in some of them. The safety ombudsman would be committed to fully respecting, amongst others, the health and safety principles at work and convey through the appropriate channels, complaints reported by employees and subcontractors. Alternatively, this role could also be played by the ethics ombudsman, with the advantage that besides the issues related to health and safety at work, all the other issues would also be reported to just one person.

11.11 Conclusion

In industry, it turns out that listening to whistle-blowers is a way to detect major degradation of safety level and, so, potentially to prevent major events. Also, middle and minor safety related unwanted events or intended events (social security) may be prevented or consequences reduced.

In that sense, listening to whistle-blowers must be an integral part of safety and / or citizens protection processes. Nevertheless, to listen to whistle-blowers does not mean to agree with them. However, listening to them should lead to open debates about industrial safety and its current and actual practices. Debates about safety could naturally, not to say mechanically, lead to an increase in safety because the organisation's mindset would change.

Taking account of whistleblowing requires the adoption of a new paradigm: to see beyond quantitative approaches and to leave room for "alternative voices" and field expertise, which is one feature of highly reliable organisations (Weick and Sutcliffe, 2001)[114].

The solution goes through a bottom-up approach (i.e., decision-makers listening to the technical experts and front-line workers) to complement the top-down approach (i.e., decision-makers asking questions), recommended, for instance, by Conklin (2012).

Whistle-blowers cannot be an official position, a box in the organization chart. To be a whistle-blower is a specific moment in a professional career.

The entire safety burden cannot be carried by whistle-blowers. Listening to whistle-blowers seems a necessary but not sufficient condition for maintaining and increasing safety. Whistle-blowing must just be one (more) tool in the toolbox for prevention. Every sign or event, near-miss… must continue to be treated in order to increase safety. For instance, in the six months preceding the DC-10 airplane accident[115], 1,000 incidents related to the cargo door were reported (it means about 10 incidents by DC-10 aircraft in service in the USA). It seems to "sign" a poor safety culture and safety flawed approaches in the aviation domain at that time So, it is not a big surprise that warning of Dan Applegate was lost in an "ocean of indifference", not to say an "ocean of denial" to safety. The curse of Cassandra lives on.

In the social domain, listening to whistle-blowers helps to fight against potential fraud and to be informed about malicious actions or misintentions of citizens or governments. They disclose information in the general interest, allow the prevention or the revelation of the flaws and dysfunctions of our States, our economies, our political and financial systems. In particular, their action has led to considerable progress in the fight against corruption. Their action has made it possible to disclose certain lies, be they lies of States or private companies, as well as certain breaches of privacy. So, their role is very important for maintaining and improving democracy.

In both cases, industry and civil society, protection of whistle-blowers is of utmost importance.

Analysing information provided by whistle-blowers should be part of Safety Management System, as a part of foresight for safety.

Foresight is about reducing uncertainty and predicting future performance. New approaches, theories, technologies and notions are still open and their desirability, feasibility and applicability is still undetermined. Their functioning may be revealed during operations as "emergent" properties with unanticipated consequences. Whistle-blowers can be seen as a specific form of operational feedback and early warnings of future mishaps in the functioning of systems and sectors.

Whistle-blowers - both individual and institutional - may serve three primary goals:

- Provide subject matter expertise
- Represent voices in a democratic participation process
- Support a multiple (ethical) value driven adaptation process which should be practiced in the context of the system in which they are applied.

Their functioning can be integrated in socio-technical systems at various levels of control. This integration could make the outsiders role of whistle-blowers obsolete and could reinstall their role as subject matter experts from within the system. Such a transition poses challenges on creating a shared repository of expertise, experiences and knowledge management, combining feedback and feed forward loops to design and operations of complex systems. As such, it may benefit foresight in safety by helping to identify early warnings of system degradation

11.12 References

ADIE (2008), Whistleblowing in Action in the EU Institutions, RBEUC, Tallinn, https://www.whistleblower-net.de/pdf/ADIE Whistleblowing EU.pdf, retrieved March 03, 2018.

AFP (2019), Lanceurs d'alerte : Un tiers des cadres confrontés à des « pratiques illégales » au travail, des dérives peu signalées, 20 Minutes, https://www.20minutes.fr/societe/2646379-20191107-lanceurs-alerte-tiers-cadres-confrontes-pratiques-illegales-travail-derives-peu-signalees,   retrieved November15, 2019.

---

[114] *For differences between "reliability" and "safety", see Llory and Dien, 2006.*

[115] *See § 11.1.4.2*

Amalberti, R. (2013), Porteurs d'alerte et signaux faibles : à la mode… et après ?, Tribunes de la sécurité industrielle – 2013, n°01, https://www.foncsi.org/fr/publications/tribunes-securite-industrielle/porteurs-alerte-signaux-faibles-a-la-mode-et-apres/Tribune-signaux-faibles.pdf, retrieved September 7, 2019.

Bensadon, A-C., Marie, E. and Morelle, A. (2011), Enquête sur le MEDIATOR®, Inspection générale des affaires sociales, http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/114000028.pdf, retrieved April 12, 2018.

Bodet, L. and Chaverou, E. (2020), Covid-19 : ces lanceurs d'alerte menacés pour avoir dit la vérité sur la pandémie, https://www.franceculture.fr/societe/covid-19-ces-lanceurs-dalerte-menaces-pour-avoir-dit-la-verite-sur-la-pandemie, retrieved April 17, 2020.

Burrough, B., Ellison, S. and Andrews, S. (2014), The Snowden Saga: A Shadowland of Serets and Light https://www.vanityfair.com /news/politics/2014/05 /edward-snowden-politics-interview, retrieved May 18, 2019.

Canfield, K. (2017), 'The Doomsday Machine', by Daniel Ellsberg, https://www.sfgate.com/books/article/The-Doomsday-Machine-by-Daniel-Ellsberg-12413956.php#item-85307-tbla-3, retrieved October 9, 208.

Chateauraynaud, F. and Torny, D. (1999), Les sombres précurseurs - Une sociologie) pragmatique de l'alerte et du risque, Éditions de l'École des Hautes Études en Sciences Sociales.

Clarke, R.A. and Eddy, R.P. (2017), Warnings – Finding Cassandras to Stop Catastrophes, Harper Collins Pubishers.

Conklin, T. (2012), Pre-Accident Investigations – An introduction to Organizational Safety, CRC Press – Taylor & Francis Group.

Council of Europe (2014), Protection of Whistleblowers - Recommendation CM/Rec(2014)7 and explanatory memorandum, https://rm.coe.int/ 16807096c7, retrieved March 03, 2018.

Cullen, W. D. [Lord] (2000,) The Ladbroke Grove Rail Inquiry, Part 1 Report, HSE Books, http://www.railwaysarchive.co.uk/docsummary.php?docID=38, retrieved April 10, 2018.

de Matos Ramos, J. A. A. (2018) A Proteção de Denunciantes de Corrupção e Criminalidade Conexa, Master Thesis, Lisbon University, Faculty of Law of Lisbon.

Dien, Y. (2006), Les facteurs organisationnels des accidents industriels, In: Magne, L. et Vasseur, D. (Coordonnateurs), Risques industriels – Complexité, incertitude et décision : une approche interdisciplinaire, pp. 133-174, Éditons TED & DOC, Lavoisier.

Eddy, P., Potter, E. and Page, B. (1976), Destination Disaster – From the Tri Motor to the DC10: The Risk of Flying, Quadrangle/The New York Times Book Co.

Edmondson, A. (1999), Psychological Safety and Learning Behavior in Work Teams, Administrative Science Quarterly, Vol. 44, No. 2 (Jun.,), pp. 350-383.

Encyclopaedia Britannica, (undated), U.S. Decision-Making in Vietnam, 1945–68, https://www.britannica.com/topic/Pentagon-Papers, retrieved September 25, 2018.

European Commission (2018), Whistleblower protection, https://ec.europa.eu/info/sites/info/files/placeholder 11.pdf, retrieved May 06, 2019

Finn, P. (2002), Crash Described as "Exceptionally Unlucky", The Washington Post, July, 3.

Frachon, I. (2010), Mediator 150 mg : Combien de morts ?, 2nde édition,Éditions-dialogues.fr. Dialogues,

Frantzen, C. (2004), Tango on an Asymptote, 13th SRA Europe Annual Conference, Paris, 15-17 November.

Gafni, M. and Garofoli J. (2020), Exclusive: Captain of aircraft carrier with growing coronavirus outbreak pleads for help from Navy, San Francisco Chronicle https://www.sfchronicle.com/bayarea/article/Exclusive-Captain-of-aircraft-carrier-with-15167883.php, retrieved April 4, 2020.

GPO (1989), Whistle blower Protection Act, Public Lax 101-12, https://www.govinfo.gov/content/pkg/STATUTE-103/pdf/STATUTE-103-Pg16.pdf, , retrieved May, 22, 2018.

G20 (2011), G20 Anti-Corruption Action Plan - Protection of Whistleblowers, https://www.oecd.org/g20/topics/anti-corruption/48972967.pdf, retrieved September 3, 2019.

Hermange, M-T. (2011), Rapport d'information fait au nom de la mission commune d'information sur : "Mediator : évaluation et contrôle des médicaments", Rapport du Sénat n° 675, http://www.senat.fr/rap/r10-675-1/r10-675-11.pdf, retrieved March 03, 2018.

Hollnagel, E. (undated), From Safety-I to Safety-II: A brief introduction to resilience engineering, http://safetysynthesis.com/onewebmedia/ Introduction%20to% 20S-I%20and%20S-II.pdf, retrieved March01, 2018.

Hollnagel, E. (2014), Safety-I and Safety-II: The Past and Future of Safety Management, CRC Press – Taylor & Francis Group.

Kemeny, J. G., Babbitt, B., Haggerty, P. E., Lewis, C. D., Marrett, C. B., Mc Bride, L., Mc Pherson Jr, H., Peterson, R., Pigford, T. H. and Trunk, A. (1979), The Need For Change – The legacy of TMI, Report of the President's Commission On The Accident At Three-Mile Island, Government Printing Office, Washington DC.

Johnson, S. (2020), NHS staff forbidden from speaking out publicly about coronavirus, The Guardian, April 9.

Lapierre, D. and Moro, J. (2001), Il était minuit cinq à Bhopal, Éditions Pocket Robert Laffont Pocket.

Lefébure, A. (2014), L'affaire Snowden - Comment les États-Unis espionnent le monde, Éditions La Découverte.

Lewis, P. and Hilder, P. (2018), Former Cambridge Analytica exec says she wants lies to stop, The Guardian, https://www.theguardian.com/uk-news/2018/mar/23/former-cambridge-analytica-executive-brittany-kaiser-wants-to-stop-lies, retrieved November 10, 2019

Linder, D. O. (2011), The Pentagon Papers (Daniel Ellsberg) Trial: An Account, http://law2.umkc.edu/faculty/projects/ftrials/ellsberg/ellsbergaccount.html, retrieved October 9, 2018.

Llory, M. (1996), Accidents industriels : le coût du silence - Opérateurs privés de parole et cadres introuvables, Éditions L'Harmattan.

Llory, M. and Dien, Y. (2006), Les systèmes sociotechniques à risques : Une nécessaire distinction entre fiabilité et sécurité, Partie 1 : Performances n°30, septembre – octobre, pp. 20-26, Partie 2 : Performances n°31, novembre – décembre, pp. 9-13, Partie 3 : Performances n°32, janvier – février, pp. 20-26.

McCall, J. (2017), Modern day heroes: a multiple case study of how successful flight crew and air traffic control coordination helped prevent disaster. International Journal of Current Research Vol 9, Issue 11, pp. 61268-61275, November 2017

McKie, Robin (2011), Japan ministers ignored safety warnings over nuclear reactors, The Guardian, http://www.theguardian.com/world/2011/mar/12/ japan-ministers -ignored-warnings-nuclear, retrieved May, 12, 2019.

Miller, J. (1995), Millstone's Neighbors in a Quandary, The New York Times, November 5.

Near, P. and Miceli P. (1985), Organizational dissidence: The case of whistle-blowing, Journal of Business Ethics, pp1-16.

Obermayer, B. and Obermaier, F. (2017), The Panama Papers: Breaking the Story of How the Rich and Powerful Hide Their Money, Oneworld Publications.

Obermayer, B. and Obermaier, F. (2016), Panama Papers. Die Geschichte einer weltweiten Enthüllung, Kiepenheuer & Witsch GmbH

Parlamentul României – Legea Nr.517/2004 privind protecţia personalului din autorităţile publice, instituţiile publice şi din alte unităţi care semnalează încălcări ale legii.

Pooley, E. (1996), Nuclear Warriors, Time Magazine, vol. 147 n° 10, March 4.

Public concern at work, (undated), The Whistleblowing Charity, http://www.whistleblowing.org.uk/business-support, retrieved September 9, 2019.

Ringdahl, L.H. (2004), Relationships between accident investigations, risk analysis, and safety management, Journal of Hazardous Materials, Vol 111 n°1-, pp):13-19

Rogovin, M. and Frampton, G. T. (1980), Three Mile Island - A Report to the Commissioners and to the Public, Vol I, NRC Special Inquiry Group, NUREG/CR-1250, Washington DC.

Schein, E. (2016), Whistle Blowing: A Message to Leaders and Managers. Comment on "Cultures of Silence and Cultures of Voice: The Role of Whistleblowing in Healthcare Organizations", International Journal of Health Policy and Management, 2016, 5(4), 265–266

Secrétariat d'État aux Transports (1976), Rapport final de la Commission d'Enquête sur l'accident de l'avion D.C. 10 TC-JAV des Turkish Airlines survenu à ERMENONVILLE, le 3 mars 1974, Journal Officiel de la République Française, Éditions des documents administratifs, N° 27 du 12 Mai.

Shrivastava, P. (1992), Bhopal Anatomy of a Crisis, 2nd edition, Paul Chapman Publishing.

Snowden, E. (2019), Permanent Record, Metropolitan Books.

Szadkowski, M. (2019), « The Great Hack » : plongée dans les eaux troubles du marketing politique de Cambridge Analytica, Le Monde, https://www.lemonde.fr/pixels/article/2019/07/24/the-great-hack-plongee-dans-les-eaux-troubles-du-marketing-politique-de-cambridge-analytica_54 92703_4408996.html, retrieved November 10, 2019.

Telegraaf (2017), Raadsel vliegramp opgelost? (Air crash mystery solved?) Newspaper clipping Dutch Newspaper De Telegraaf pp T18. 21 December 2017

THE WHISTLE BLOWERS PROTECTION ACT, (2014), http://legislative.gov.in /sites/default/files/A2014-17.pdf, retrieved May, 6, 2019.

Turner, B. and Pidgeon, N. (1997), Man-Made Disasters, 2nd edition, Butterworth Heinemann, Oxford [1rst edition: Turner, B. (1978), Wykeham Publications].

U.S. CSB (2007), Investigation Report, Refinery Explosion and Fire, BP – Texas City, Texas, March 23, 2005, Report N°2005-04-I-TX.

Vaughan, D. (1996), The Challenger Launch Decision - Risky Technology, Culture, and Deviance at NASA, The Chicago University Press.

Wells T. (2009), Wild Man - The Life and Times of Daniel Ellsberg, Palgrave Macmillan

Weick, K. and Sutcliffe, K. (2001), Managing the Unexpected – Assuring High Performance in an Age of Complexity, Jossey-Bass Publishers.

Wikipedia 1 (undated), Joseph P. Kennedy Jr. wikipedia.org/wiki/Joseph _P._Kennedy_Jr, retrieved June 7, 2018.

Wikipedia 2 (undated), www.ad.nl/binnenlands/huis-voor-klokkenluiders-kostte-negen- milioen-en-loste-nul0zaken-op~a0e6d73f/, retrieved June 7, 2018.

Wikipedia, 3 (undated), Vietnam War, https://en.wikipedia.org/wiki/Vietnam War, retrieved November 24, 2018.

Wikipedia 4 (undated), Daniel Ellsberg, (https://en.wikipedia.org/wiki /Daniel_Ellsberg, retrieved September 15, 2018.

Wikipedia 5 (undated), Randy Kehler, https://en.wikipedia.org/wiki/ Randy_Kehler, retrieved September 8, 2018.

Wilson K. and Straker D., 2018. Fiction versus reality: The Impact of Hollywood on Accident Investigation. ISASI Forum, Oct-Dec 2018. P 24-26.

Wilson, K. and Straker, D. (2018), Fiction versus reality: The Impact of Hollywood on Accident Investigation. ISASI Forum, Oct-Dec 2018. Pp 24-26.

12 Role of Technology in Foresight for Safety - Technological potentials and challenges to enhance foresight in safety

Zdenko Šimić, European Commission, Joint Research Centre (EC JRC), the Netherlands.

Executive Summary

This chapter identifies the major existing and emerging technologies relevant for foresight in safety, based on a systematic literature review. The chapter presents technologies, domains and applications in use to improve safety directly and by enabling the use of foresight. The review identifies potentials, limitations and difficulties associated with the application and the use of advanced technologies for enhancing safety and enabling foresight in safety.

New technologies are mainly based on the availability of advanced sensors and growing computing power, communication bandwidth and storage capacity. These basic technologies are improving software and hardware solutions and allowing the use of more advanced technologies like computer aided hybrid development, real time modelling, advanced simulations and artificial intelligence.

Technological advances are useful for all stages of the life cycle of safety related systems, i.e. design, verification, validation, production, testing, commissioning, operation, maintenance, emergency response and decommissioning. Improved designs assure system safety with an optimised use of resources.

The use of realistic simulators for the training of operators increases both performance and safety.

Monitoring and predictive diagnostics also improve availability and reduce the risk.

The use of technology for enhancing safety and foresight is scalable and widely applicable to many domains (e.g. transportation, power generation, construction, process industry, etc.). Different applications allow for the optimisation without compromise in safety. Advanced risk modelling with big data analytics and knowledge management allows for the integration of foresight in safety with conventional approaches. Simulations and virtualisations are improving the design, operation, safety and planning, while creating more realistic accident management. Improved realism, extended scope (with scenarios and time coverage) and easy use are together enabling foresight in safety.

The role of technology in enabling foresight in safety is to complement conventional approaches with the wider consideration (including less likely events and scenarios for long-term time horizons) as well as with means for a wider participation.

Some of the problems associated with the use of advanced technologies are related to the increased technical complexity and required connectivity. The problem with complexity is that software and digital instrumentation and control is more challenging to verify and validate. The necessary connectivity, wireless and over internet, increases both privacy and cyber security related risks.

The potential benefits from the use of new and emerging advanced technologies are continuously increasing. This is especially important for safety related applications where optimisation should not compromise safety. Foresight in safety is easier to apply with advanced technological tools like modelling and simulations over the whole life cycle of the system. It is important to assure that these benefits are greater than the threats coming from the use of many new technologies.

12.1 Introduction

Modern life is more and more changing because of the ongoing and increasing digitalisation of the world. The change to society is mainly digital (new software and more powerful hardware) but it is also complemented with the development of novel and inexpensive sensors, networks and communication systems. When combined, these technologies enable connectivity and base for many new applications, i.e. wide band global communication systems, affordable remote data storage, global positioning system (GPS), cloud computing power and internet of things (IoT). The resulting development is generally improving everyday life, the economy and society as whole but is also causing serious unintended consequences and issues.

The role of technology in safety and foresight is of special importance because technological advances can potentially influence safety related systems through all

stages of their life cycle. The design and operation of safety related systems can benefit significantly from technology, with a potential for continuous safety assurance through the application of advanced hardware and software solutions with optimised use of resources. A more integrated life cycle of safety related systems allows for example the use of advanced risk assessment modelling and the easier inclusion of complementary foresight thinking.

A specific example of a framework for an integrated nuclear digital environment, in [1] illustrates the potentials. The UN Sendai Framework for Disaster Risk Reduction (DRR) 2015-2030, on the other hand, emphasises the general use of science and technology, **Error! Reference source not found.**.

This chapter portrays a systematic review, which aims to identify major existing and emerging technologies with tangible potential safety benefits applicable to different life cycle phases (i.e. design, verification, validation, production, testing, commissioning, operation, maintenance, emergency response and decommissioning) of various systems.

The goal is also to identify domains of application and to provide typical examples of potential safety benefits emphasising the relevance for foresight in safety.

New technologies, while solving many problems, also introduce new challenges related to the use and safety (e.g. with unintended consequences). This duality raises many questions about the optimal development, regulation and implementation of new technologies.

The relevance of new technologies for the foresight in safety will be measured by how the conventional approach to safety can be extended and enhanced in relation to the inclusion of less likely and long-term scenarios proposed by a multidisciplinary team.

Foresight in safety can potentially greatly benefit from different new technologies (e.g. advanced simulations, visualisations and virtualisations). It is important to note that the development, use and valorisation of new technologies could also benefit from the foresight thinking and approaches.

The chapter is organised in the following way: the first section describes the approach and the scope of this systematic review; the following section presents findings and discussion; final section contains the concluding remarks.

Life cycle where technologies for foresight in safety are used
The role of technology is key for foresight in safety during all phases and related systems' activities – concept development, design, production, commissioning, operation, and decommissioning; validation, verification, testing, monitoring, education and training.

12.2 Approach

The role of technology in safety and foresight is inherently connected to both the use of technology in general and to the safety related systems specifically. This results in many different safety applications and related approaches to the development and use of regulations. Learning about the safety related use of technology and understanding the value of numerous applications in many different domains is a significant challenge.

A literature review was selected as an approach to define the systematic multi-domain role of technology in safety and foresight. Considering the author's background and the specific importance of safety the review is more focused on nuclear power. Many other fields are also included in an effort to make the review more comprehensive and to illustrate the generic value of many technologies that are useful for safety and foresight. Similar examples from different domains are used in order to illustrate common approaches and solutions. Reviewing different domains is also valuable in identifying different issues and potential limitations.

Google Scholar (scholar.google.com/intl/en/scholar/about.html) online search tool was used to access a list of related literature. The tool gives simple and flexible open access to most comprehensive cross-domain literature databases. References from many disciplines and sources are included, i.e. journal articles, theses and books from academic publishers and other online repositories like societies, universities and libraries. Google Scholar ranks articles based on the content, the publisher, the authors and the citation. The location of documents (in publishers and other repositories) and different versions are also available. Based on some estimates Google Scholar is the most comprehensive academic search engine with 389 million records [3].

Performing a literature search with Google Scholar is as easy as doing a regular web search with additional special functionalities. The search for chosen key words can be applied for selecting a time range. Search results contain a total number of references ordered based on the relevance. Every reference contains a number and the access to the searchable list of cited by articles. Additionally, search results provide a list of related articles. This is all web browser based and conveniently hyperlinked. With a Google account, it is possible to save interesting references.

The search for this chapter was made mostly for recent years, with only a few exceptions for some domains, in order to capture the broader development. The initial search was made with the key-words "technology" and "safety". Based on the titles and an abstract review the most relevant references were selected from the list as pointers for a further, more refined search and review. Over 100 papers were initially selected for further review. The selected references were then grouped in domains including "miscellaneous". A special group was related to "issues" from all domains.

The findings from this review are presented in the next section covering six dimensions as follows: domains where these general technologies are used; general groups of available technologies; list of specific applications; identified parts of life-cycle with related activities where technologies are used; technologies especially related and useful for the foresight in safety; identified issues related to preventing use of technologies or issues with potential to introduce new safety problems.

The selection and review are representative for the role of technology in safety and foresight. However, this is far from the most representative or comprehensive review considering rapid developments, as well as the number of domains and applications.

Domains where technologies for foresight in safety are used
Technology is used for safety and foresight everywhere: transport, power generation, medicine, construction, mining, military, industry, food, meteorology, security, communication, internet, research & development, smart cities, disasters risk reduction and society in general.

12.3 Findings and discussion

The present findings about the role of technology in safety and foresight are based on the more detailed review of 60 references. The presentation is divided into three subsections. The first subsection presents the fundamental technologies with example applications in different domains. The second subsection presents the major issues related to safety and foresight from the use of technologies. The final subsection covers the role of technology in foresight for safety.

The total yearly number of English articles resulting from the search of the term "technology safety"[116] has increased to the point that it doubled during the last ten years to over 300 thousands in 2019. This illustrates the increasing importance of technology for safety. The yearly number of publications including the term "safety foresight" is steady over time at about five thousands (and it is similar when the term "technology" is added). Perhaps this is an indication of a constant, if not strong, interest for the foresight in safety.

12.3.1 Findings about the role of technology in safety and foresight

The role of technology in safety and foresight is reviewed through examples from literature in nuclear power and several other domains. The findings are presented as short explanations of the technology used, the specific application, the part of the life-cycle in which it is used and how it contributes to the foresight in safety. Applications are usually combining several technologies and the grouping of technologies selected in this review emphasises the role of computing power and software.

Technologies useful for foresight in safety
Computing power, software, cloud computing, sensors, laser scanning, radars, machine learning, artificial intelligence, smartphones, social networks, internet, internet of things, global positioning system, geographic information system, virtual reality, augmented reality, 3D printing, big data analytics, knowledge management and blockchain.

---

[116] *These numbers were derived using Dimensions online tool (www.dimensions.ai/). This resource is similarly big to Google Scholar with a functionality for generating yearly statistics of publications.*

12.3.1.1   Computing power and advanced software

Computing power is an enabling factor for many other technologies and applications, e.g. safer and affordable design needs the fastest possible computing in order to explore numerous alternatives and test them against multiple hazards. A nuclear power plant (NPP) design, for instance, requires both the highest level of safety and economic competiveness. High performing computing with advanced modelling and simulation is necessary to include multi-physics "core simulation" (e.g. radiation transport, thermal-hydraulics, corrosion chemistry, etc.), which requires algorithms for robust numerical solutions and uncertainty quantification [4]. The access to enormous computing is significantly improved with so-called *cloud computing*. Significant and affordable computing power enables the application of many other technologies including advanced software technologies.

Advanced software consists of algorithms and the necessary hardware. The hardware is usually a combination of high computing power, sensors and some other peripherals. Several advanced software technologies are presented here, both specific (i.e. simulators, building information modelling and virtual and augmented realities) and general (i.e. visualisation and knowledge management).

Plant simulators are proven tools allowing for a better training of operators of complex systems like airplanes, NPPs and process plants, [5]. Methods for designing a human system interface evolve together with the development of technologies and include more than just an improved user interface, [6]. Simulators improve in two different directions in order to make them fully realistic and affordable. So-called full-scale simulators can not only present the full operational characteristics of the plant but also the accident conditions and the development of related scenarios, [7]. Simplified simulators can run on a single personal computer and still represent most of the plant operation, including emergency conditions. This improves both the education and training of plant engineers and operators, [8]. Plant simulators enable foresight in safety with a wider involvement and engagement of staff and stakeholders for the evaluation of different plant conditions and for testing various what-if conditions.

Virtual and augmented realities (VR and AR) are the most advanced software developments with the potential to improve education, training and operation. Both VR and AR could be applied independently or combined with other technologies like simulations. For instance [9] and [10] suggest a virtual environment and simulation as means to improve the safety during both the work

and the decommissioning of an NPP. In [11] the use of augmented reality is evaluated for safety signs in the working environment. The use of AR for generating safety awareness and enhancing emergency response for construction, earthquakes and driving is reviewed in [12]. VR and AR have a potential similar to plant simulators (and especially in combination with simulators) for enabling foresight in safety with a wider participation and the consideration of less probable plant conditions.

Visualisation and *multimedia* are demonstrated to be beneficial, for example in the construction industry (e.g improved safety management and training, hazards identification, monitoring and warnings, [13]), and hospitals (e.g. preventing surgery mistakes, [14]). The *Building information modelling* (BIM) framework is used in construction design, implementation and operation for different domains (e.g. for nuclear [1] and general waste [15]) and many applications like construction risk management [16] and fire protection [17]). BIM is also used for the planning and building of the first high-level radioactive waste final disposal facility by Posiva in Finland, [18]. The risk management potentials for BIM are further enhanced with ontology and web semantic technologies, [19]. Visualisation and multimedia enable foresight in safety by allowing a wider participation and safety deliberation. BIM is therefore important for foresight in safety because it enables a long time consideration.

Knowledge management (KM) is increasingly important for complex systems during the whole life cycle. KM has the potential to improve the safety economy with a better design and efficient operation and decommissioning, with a better use of accumulated knowledge and experience, [20]. Successful KM relies on many building blocks like information systems, databases, collaborative networking, expert systems, ontologies, web semantics and organizational culture. KM enables foresight in safety by allowing a wider involvement and consideration of greater operating experience (i.e. from other plants and industries, including less significant events).

Computing and software related technologies do not always depend on high computing power or sophisticated solutions. *Novel approaches* and *advanced algorithms* solutions could result in safety improvement like in the central control of trains to avoid rear-end collisions in [21]. However, the computing power and the software needed to design and test these solutions are still necessary.

Applications of technologies useful for foresight in safety
Technologies applications for foresight in safety: optimised design without safety compromise; enhanced validation and verification; virtual/augmented experience for better design, operation and emergency planning; improved and effective education, training, operation and maintenance.

12.3.1.2 Sensors, internet, communication, "big data" and artificial intelligence

Sensors are (essential) components required for an efficient and safe operation (they are critical e.g. for avoiding dangerous situations and for reducing unwanted consequences). The requirements for sensors (like precision, speed, robustness, connectivity, energy consumption and cost) depend on the domain and application. One example in security checking for explosives, where both speed and sensitivity are required, is the use of thermo-desorption mass spectrometry, [22]. Further examples are the use of hyperspectral imaging technique for automated non-destructive analysis and assessment applied to a wide range of food products (for disease detection and quality control), reviewed in [23]. The values measured by sensors also depend on the software capabilities to interpret signals and diagnose conditions, and to predict developments. In [24] the use of a distributed equation and artificial immunity system is proposed for the online monitoring and prediction in condensate and feed water system of the NPP. Sensors enable foresight in safety by expanding the possibilities of collecting small signals.

Internet, as a network of computers, sensors and people, has a growing potential for technologies and applications in many domains. Information about online search queries is useful for various applications, e.g. early detection of food related epidemics, [25]; perception and prediction of viral and other outbreaks, [26] and [27]. Together with sensor equipped devices like smartphones, this presents an additional potential for the use of technology to improve safety, e.g. monitoring health behaviour, [28]; managing construction, [29]; and for collision warning while driving, [30]. The traceability of (and thus the possibility to prove) the origin of safety parts could be solved with new software technology, like blockchain, by assuring the validity of records, [31]. Other technologies like cloud computing rely on the internet for accessibility and affordability. Internet and *communication* enhance foresight in safety by enabling wide and instant participation.

Geographical information system (GIS) is useful for area integrated risk management like regional risk assessment in [32]. The optical, radar and other satellite data obtained with GIS are useful as a support for emergency response services for natural, technology and social related hazards, [33]. The disaster planning, warnings and response incorporate the use of social networking like tweets, [34]. The combined use of an increasing number of satellites could improve both the resolution and the responsiveness (i.e. to hours). *Global positioning system* (GPS) has many applications from industry to personal use and it is critical for real time use. However commercially available resolution still limits some new applications, like autonomous driving, [35].

Data from *video* and *mobile* sensors are useful for improving safety in many applications, like intersection monitoring for safety analysis, [36]. Wearable personal devices with biosensors (e.g. for hart beat, movement, sleep behaviour) are able to track physiology data that make health diagnostics and decisions about therapies easier, [37]. The accumulation of data from an increased number of sources presents an opportunity for a better understanding of complex systems and for providing new insights to safety science, [38].

The analysis and interpretation of huge volumes of data ("*big data*") requires and enables the use of new techniques like *machine learning* (ML) and *artificial intelligence* (AI). Impressive recent AI results, surpassing humans in GO game and medical diagnostics, show huge promises. However, the limits and timescale for the development of the further potential of AI are not easy to predict. About 50% of the experts believe that high-level machine intelligence will be developed in the next 30 years and that superintelligence might be developed in the subsequent 30 years, [39]. New AI applications like automatizing human work are increasingly available, e.g. restaurants food safety check and simple news writing, [40]. ML and AI enable foresight in safety by analysing all available data and by identifying important patterns that are hardly noticeable for humans. A wider participation in safety assessment is possible with user friendly tools based on ML and AI.

Modern vehicles are increasingly equipped with safety technologies assisting drivers (e.g. automatic emergency braking, blind spot monitoring, road line support system, objects recognition). Fully *autonomous vehicles* could be commercially available in several years. Automated vehicles are result from the

implementation of leading edge technology solutions, including advanced sensors, computing power and edge AI, [41].

The number of technologies having the potential to enable foresight in safety is significant. Two more technologies are mentioned here to illustrate this growing and varied potential. *Unmanned aerial vehicles* (UAV, drones) are useful for numerous applications in remote monitoring (e.g. 3D radioactive contamination mapping, [42]). Eye movement recording and analysis allow experts like pathologists to learn and improve their diagnostics, [43]. Advanced 3D printing is used in many domains for the preparation of difficult tasks, for producing custom complex parts, and for education and training, e.g. in medicine [44].

12.3.2 Issues with the use of technology for safety and foresight

Some potentials and promises of new technologies for improving safety and enabling foresight need testing before wide adoption. This is necessary even for simple applications like material condition monitoring ([45]) and especially for complex solutions like digital control rooms (DCR), [46]. An example of DCRs show that the potential might be different for various domains depending on many elements like the implementation and the operators' age. For instance [46] documents potential side effects reducing the operators' reliability in DCRs for NPPs.

The verification and validation (V&V) for digital technologies is an open problem. While by nature digital technologies allow for realistic virtual testing, the existence of an immense number of possible states makes a full testing practically impossible. This is the case for example with the autonomous car [47] and with the nuclear digital instrumentation and control, [48]. Experience proves that hardware and software induced failures are inevitable in complex digital systems and this should always be factored into the design redundancy and the system's recovery function, [49]. Completely new problems arise from the limited capability to provide for adequate reasoning and arguments for the results created by AI. A number of recommendations for the research and development prioritisation in the development of NPPs relates to the adaptation of digital technologies addressing V&V and other issues, [50].

The Internet and related social networks are both useful and cause many bad unintended consequences (e.g. effective dissemination of false information). The reliability of information is important for a better functioning society and especially during an emergency situation when it can have detrimental effects, as it was tragically illustrated during and after the Fukushima Daiichi nuclear accident, [51]. The possibility that online data are imperfect, incomplete and changing should be always considered, [52]. Human and AI based solutions are in development to help with this problem. However, the problem of information reliability is increasing and additionally complicated with other issues like free speech, who is provider, etc.

While smartphones allow for easy communication and access to information, they are also a distraction for important activities like driving, and could be the cause of accidents, [53]. This is regulated in some countries and easy to identify through the availability of recording of activity by the smartphone before any accident.

The cost limits the introduction of some technologies with proven benefits. Usually a wide use would make them affordable. The cost and potentials depend on many factors specific to each application, country or situation. E.g. the difference in perception of the so-called "value of prevented fatality" justifies the installation of commute bus crash avoidance systems in the U.S. but not in Colombia, [54].

Cybersecurity is one of the major issues for many internet and wireless based technologies because a perfect protection is impossible without losing functionality, e.g. for autonomous vehicles [55]. Hacking is an increasing problem on the Internet and it might reduce the trust in some new technologies like various applications of IoT and AI, e.g. for autonomous vehicles and medical assistance devices, [56].

Issues with the use of technologies for safety
Use of technology for safety and foresight in safety has number of issues: cost, complexity; verification & validation; faster change cycles; cyber security; disinformation; distraction; proved benefits; privacy; AI explained.

Solutions for the issues mentioned are not trivial and will require continuous development. For some of these issues the solution is technology itself, either already built in (e.g. communication for UAV collisions, [57]) or complemented

with other solutions (e.g. documenting scientific software for nuclear safety applications, [58]). Another part of the solution is learning by doing (e.g. for health information technology, [59]) after accepting a new technology with simple criteria in order to prove that it is at least as good as the technology already in use. Some issues with a new technology will require the development of new methods which will help to prevent unwanted consequences, e.g. for detecting promoted social media campaigns, [60].

12.3.3   Discussion about the role of technology in safety and foresight

The examples of the current and potential benefits from the use of technologies presented in this review demonstrate their usefulness for the whole life cycle of various safety related systems. The potential for the role of technology in improving safety seems is vast. Technology is beneficial for more efficient and safer operation. Advanced technologies enable foresight in safety in all three dimensions: plausibility, scope and inclusiveness. Enhanced analytical capabilities enable the consideration of less probable events and scenarios. Affordable advanced simulations allow for the consideration of long-term developments. The Internet, big data analytics and visualisation make it feasible for more actors to participate and to include the views and assumptions of non-experts' in the considerations.

Use of technologies for foresight in safety, examples:

Foresight in safety use of technology examples: improving the participation, the scope and the realism of safety analysis and related simulations; continuous improvements based on data, simulations and a wider participation; prompt and appropriate accident management based on the improved assessment and a wide participation; designing faster and better emergency response, based on the realistic assessment and a wide participation to minimise consequences and prevent societal disruptions; use of simulations and a wider participation to improve accident prevention and emergency preparedness.

Technology enables and improves foresight in safety with an extended scope of assessment, a long-term consideration and a wider participation. Some examples of how technology improves foresight in safety are the improved analysis and simulations to identify and anticipate safety issues; implementing adaptive maintenance to prevent failures; enabling continuous safety improvements with operating data assessment; assuring accident prevention with timely preparation and appropriate response; helping prompt and appropriate accident management with real time assessment; supporting faster and better emergency response with appropriate organisation and communication, [61]; improving learning from accident investigation; preventing societal disruptions with proper communication.

12.4   Conclusions

A systematic literature review provided many examples of technology used for improving safety and enabling foresight in safety. The numerous examples presented in this chapter demonstrate how various technologies, both individually and combined, could improve the safety and enable foresight in safety.

Many benefits are already in realised while some benefits are still in development.

The role of technology in enabling foresight in safety is to complement the conventional approaches with the consideration of a more extensive scope (including less likely events and considering scenarios for long-term horizons) as well as with means for a wider participation.

There are also issues caused by complexity and too fast introduction of new technology without sufficient regulation. Some of these issues (related to design and testing) can be solved by using advanced technologies, while others (related to safety assessment) require further intrinsic development. These issues will require the solution of specialists, but they also depend on the regulation, on the users' participation and on the change of perception.

The adaptation of new technologies might improve by accepting relative criteria in respect to existing technology, i.e. by keeping the same safety requirements for new technologies as it is applied for current technology, and by adopting a "learning while doing" approach with a demonstrated minimum initial safety level.

12.5 References

[1] Patterson EA, Taylor RJ, Bankhead M. *A framework for an integrated nuclear digital environment*. Progress in Nuclear Energy. 2016 Mar 31;87:97-103.

[2] Aitsi-Selmi A, Murray V, Wannous C, Dickinson C, Johnston D, Kawasaki A, Stevance AS, Yeung T. *Reflections on a science and technology agenda for 21st century disaster risk reduction*. International Journal of Disaster Risk Science. 2016 Mar 1;7(1):1-29.

[3] Gusenbauer M. *Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases*. Scientometrics, 2019;V118;I1;177-214.

[4] Turinsky PJ, Kothe DB. *Modeling and simulation challenges pursued by the Consortium for Advanced Simulation of Light Water Reactors (CASL)*. Journal of Computational Physics. 2016 May 15;313:367-76.

[5] Colombo S, Golzio L. *The Plant Simulator as viable means to prevent and manage risk through competencies management: Experiment results*. Safety Science. 2016 Apr 30;84:46-56.

[6] Hugo JV, Gertman DI. *A Method to Select Human – System Interfaces for Nuclear Power Plants*. Nuclear Engineering and Technology. 2016 Feb 29;48(1):87-97.

[7] Li Y, Lin M, Yang Y. *Coupling methods for parallel running RELAPSim codes in nuclear power plant simulation*. Nuclear Engineering and Design. 2016 Feb 29;297:1-4.

[8] NEI. *Educational revolution: An integrated suite of training simulators running on standard PCs is transforming initial training for all types of nuclear workers*. NEI; 2015 Nov 19. (www.neimagazine.com/features/featureeducational-revolution-4731118/)

[9] Jeong KS, Choi BS, Moon JK, Hyun DJ, Lee JH, Kim IJ, Kang SY, Choi JW, Ahn SM, Lee JJ, Lee BS. *The safety assessment system based on virtual networked environment for evaluation on the hazards from human errors during decommissioning of nuclear facilities*. Reliability Engineering & System Safety. 2016 Dec 31;156:34-9.

[10] Liu YK, Li MK, Peng MJ, Xie CL, Yuan CQ, Wang SY, Chao N. *Walking path-planning method for multiple radiation areas*. Annals of Nuclear Energy. 2016 Aug 31;94:808-13.

[11] de Amaral LR, Duarte E, Rebelo F. *Evaluation of a Virtual Environment Prototype for Studies on the Effectiveness of Technology-Based Safety Signs*. International Conference on Applied Human Factors and Ergonomics 2017 Jul 17 (pp. 100-111). Springer, Cham.

[12] Agrawal A, Acharya G, Balasubramanian K, Agrawal N, Chaturvedi R. A *Review on the use of Augmented Reality to Generate Safety Awareness and Enhance Emergency Response*. International Journal of Current Engineering and Technology, 2016 Jun; 6(3):813-820.

[13] Guo H, Yu Y, Skitmore M. *Visualization technology-based construction safety management: A review*. Automation in Construction. 2017 Jan 31;73:135-44.

[14] Dixon JL, Mukhopadhyay D, Hunt J, Jupiter D, Smythe WR, Papaconstantinou HT. *Enhancing surgical safety using digital multimedia technology*. The American Journal of Surgery. 2016 Jun 30;211(6):1095-8.

[15] Akinade OO, Oyedele LO, Munir K, Bilal M, Ajayi SO, Owolabi HA, Alaka HA, Bello SA. *Evaluation criteria for construction waste management tools: towards a holistic BIM framework*. International Journal of Sustainable Building Technology and Urban Development. 2016 Jan 2;7(1):3-21.

[16] Zou Y, Kiviniemi A, Jones SW. *A review of risk management through BIM and BIM-related technologies*. Safety Science. 2016 Jan 23.

[17] Cheng MY, Chiu KC, Hsieh YM, Yang IT, Chou JS, Wu YW. *BIM integrated smart monitoring technique for building fire prevention and disaster relief*. Automation in Construction. 2017 Dec 31;84:14-30.

[18] NEI, *Engaging with BIM*, 2016, Nov 17. (www.neimagazine.com/...-with-bim-5672206/)

[19] Ding LY, Zhong BT, Wu S, Luo HB. *Construction risk knowledge management in BIM using ontology and semantic web technology*. Safety science. 2016 Aug 31;87:202-13.

[20] Wang M, Zheng M, Tian L, Qiu Z, Li X. *A full life cycle nuclear knowledge management framework based on digital system*. Annals of Nuclear Energy. 2017 Oct 31;108:386-93.

[21] Wang J, Wang J, Roberts C, Chen L, Zhang Y. *A novel train control approach to avoid rear-end collision based on geese migration principle*. Safety science. 2017 Jan 31;91:373-80.

[22] Zhao Q, Liu J, Wang B, Zhang X, Huang G, Xu W. *Rapid screening of explosives in ambient environment by aerodynamic assisted thermo desorption mass spectrometry*. Journal of Mass Spectrometry. 2017 Jan 1;52(1):1-6.

[23] Liu Y, Pu H, Sun DW. *Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications*. Trends in Food Science & Technology. 2017 Nov 1;69:25-35.

[24] Wang H, Peng MJ, Wu P, Cheng SY. *Improved methods of online monitoring and prediction in condensate and feed water system of nuclear power plant*. Annals of Nuclear Energy. 2016 Apr 30;90:44-53.

[25] Bahk GJ, Kim YS, Park MS. *Use of internet search queries to enhance surveillance of foodborne illness.* Emerging infectious diseases. 2015 Nov;21(11):1906.

[26] Petersen J, Simons H, Patel D, Freedman J. *Early detection of perceived risk among users of a UK travel health website compared with internet search activity and media coverage during the 2015–2016 Zika virus outbreak: an observational study*. BMJ open. 2017 Aug 1;7(8):e015831.

[27] Bates M. *Tracking Disease: Digital Epidemiology Offers New Promise in Predicting Outbreaks*. IEEE pulse. 2017 Jan;8(1):18-22.

[28] Ernsting C, Dombrowski SU, Oedekoven M, LO J. *Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey*. Journal of medical Internet research. 2017 Apr;19(4).

[29] Azhar S, Jackson A, Sattineni A. *Construction apps: a critical review and analysis*. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction 2015 Jan 1 (Vol. 32, p. 1). Vilnius Gediminas Technical University, Dep. of Construction Economics & Property.

[30] Botzer A, Musicant O, Perry A. *Driver behavior with a smartphone collision warning application – A field study*. Safety science. 2017 Jan 31;91:361-72.

[31] Zheng Z, Xie S, Dai HN, Wang H. *Blockchain Challenges and Opportunities: A Survey*. Work Pap. 2016.

[32] Zhao M, Liu X. *Regional risk assessment for urban major hazards based on GIS geoprocessing to improve public safety*. Safety science. 2016 Aug 31;87:18-24.

[33] Denis G, de Boissezon H, Hosford S, Pasco X, Montfort B, Ranera F. *The evolution of earth observation satellites in europe and its impact on the performance of emergency response services*. Acta Astronautica. 2016 Nov 30;127:619-33.

[34] Landwehr PM, Wei W, Kowalchuck M, Carley KM. *Using tweets to support disaster planning, warning and response*. Safety science. 2016 Dec 31;90:33-47.

[35] Murrian MJ, Gonzalez CW, Humphreys TE, Pesyna Jr KM, Shepard DP, Kerns AJ. *High-precision GPS Vehicle Tracking to Improve Safety*. TR-1115, D-STOP, University of Texas. 2016 Sep.

[36] Shirazi MS, Morris BT. *Looking at intersections: a survey of intersection monitoring, behavior and safety analysis of recent studies*. IEEE Transactions on Intelligent Transportation Systems. 2017 Jan;18(1):4-24.

[37] Li X, Dunn J, Salins D, Zhou G, Zhou W, Rose SM, Perelman D, Colbert E, Runge R, Rego S, Sonecha R. *Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information*. PLoS biology. 2017 Jan 12;15(1):e2001402.

[38] Ouyang Q, Wu C, Huang L. Methodologies, principles and prospects of applying big data in safety science research. Safety Science. 2018 Jan 1;101:60-71.

[39] Müller VC, Bostrom N. *Future progress in artificial intelligence: A survey of expert opinion*. Fundamental issues of artificial intelligence 2016 (pp. 553-570). Springer.

[40] Thurman N, Dörr K, Kunert J. *When Reporters Get Hands-on with Robo-Writing: Professionals consider automated journalism's capabilities and consequences*. Digital Journalism. 2017 Feb 26:1-20.

[41] McGehee DV, Brewer M, Schwarz C, Smith BW, Jensen M, Tudela A, Row S, Krechmer D, Flanigan E. Review of Automated Vehicle Technology: Policy and Implementation Implications. UoI, RB28-015, IDoT, 2016 Mar.

[42] Martin PG, Kwong S, Smith NT, Yamashiki Y, Payton OD, Russell-Pavier FS, Fardoulis JS, Richards DA, Scott TB. *3D unmanned aerial vehicle radiation mapping for assessing contaminant distribution and mobility*. International Journal of Applied Earth Observation and Geoinformation. 2016 Oct;52:12-9.

[43] Brunyé TT, Mercan E, Weaver DL, Elmore JG. *Accuracy is in the eyes of the pathologist: The visual interpretive process and diagnostic accuracy with digital whole slide images*. Journal of biomedical informatics. 2017 Feb 28;66:171-9.

[44] Marro A, Bandukwala T, Mak W. *Three-dimensional printing and medical imaging: a review of the methods and applications*. Current problems in diagnostic radiology. 2016 Feb 29;45(1):2-9.

[45] Boguski J, Przybytniak G. *Benefits and drawbacks of selected condition monitoring methods applied to accelerated radiation aged cable*. Polymer Testing. 2016 Aug 31;53:197-203.

[46] Liu P, Li Z. *Comparison between conventional and digital nuclear power plant main control rooms: A task complexity perspective, Part I: Overall results and analysis*. International Journal of Industrial Ergonomics. 2016 Feb 29;51:2-9.

[47] Kalra N, Paddock SM. *Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?*. Transportation Research Part A: Policy and Practice. 2016 Dec 31;94:182-93.

[48] Li Y, Lin M, Yang Z, Hou Y, Yang Y. *Methods of applying nuclear simulation technology to the dynamic site testing of digital I&C system—I: Scheme of OLVT*. Annals of Nuclear Energy. 2017 Jun 30;104:157-65.

[49] Fan CF, Yih S, Tseng WH, Chen WC. *Empirical analysis of software-induced failure events in the nuclear industry*. Safety science. 2013 Aug 31;57:118-28.

[50] McCarthy K. *Research, Development and Demonstration (RD&D) Needs for Light Water Reactor (LWR) Technologies* A Report to the Reactor Technology Subcommittee of the Nuclear Energy Advisory Committee (NEAC) Office of Nuclear Energy US Department of Energy. Idaho National Laboratory, Idaho Falls, ID (United States); 2016 Apr 1.

[51] Utz S, Schultz F, Glocka S. *Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster*. Public Relations Review. 2013 Mar 31;39(1):40-6.

[52] Diaz F, Gamon M, Hofman JM, Kıcıman E, Rothschild D. *Online and social media data as an imperfect continuous panel survey*. PloS one. 2016 Jan 5;11(1):e0145406.

[53] Boudette NE, *Biggest Spike in Traffic Deaths in 50 Years? Blame Apps*. New York Times. 2016 Nov 15. (www.nytimes.com/2016/11/16/business/tech-distractions-blamed-for-rise-in-traffic-fatalities.html)

[54] Mangones SC, Fischbeck P, Jaramillo P. *Safety-related risk and benefit-cost analysis of crash avoidance systems applied to transit buses: comparing New York City vs. Bogota, Colombia*. Safety science. 2017 Jan 31;91:122-31.

[55] Straub J, McMillan J, Yaniero B, Schumacher M, Almosalami A, Boatey K, Hartman J. *CyberSecurity considerations for an interconnected self-driving car system of systems*. System of Systems Engineering Conference (SoSE), 2017 12th 2017 Jun 18 (pp. 1-6). IEEE.

[56] Hengstler M, Enkel E, Duelli S. *Applied artificial intelligence and trust — The case of autonomous vehicles and medical assistance devices*. Technological Forecasting and Social Change. 2016 Apr 30;105:105-20.

[57] Dill ET, Young SD, Hayhurst KJ. SAFEGUARD: *An assured safety net technology for UAS*. In Digital Avionics Systems Conference (DASC), 2016 IEEE/AIAA 35th 2016 Sep 25 (pp. 1-10). IEEE.

[58] Smith WS, Koothoor N. *A document-driven method for certifying scientific computing software for use in nuclear safety analysis*. Nuclear Engineering and Technology. 2016 Apr 30;48(2):404-18.

[59] Fong A, Howe JL, Adams KT, Ratwani RM. *Using Active Learning to Identify Health Information Technology Related Patient Safety Events*. Applied clinical informatics. 2017;8(1):35-46.

[60] Ferrara E, Varol O, Menczer F, Flammini A. *Detection of Promoted Social Media Campaigns*. In ICWSM 2016 Mar 31 (pp. 563-566).

[61] Xie T, Li CD, Wei YY, Jiang JJ, Xie R. *Cross-domain integrating and reasoning spaces for offsite nuclear emergency response*. Safety science. 2016 Jun 30;85:99-116.

13 The Role of Safety Authorities in Providing Foresight

Tuuli Tulonen, Tukes - Safety Technology Authority, Finland,
Zdenko Simic, European Commission, DG-Joint Research Centre, Energy, Transport & Climate, (JRC), the Netherlands,
Eric Marsden, Fondation pour une Culture de Sécurité Industrielle (FonCSI), France,
Frank Verschueren, Federal Public Service Employment, Labour and Social Dialogue, Belgium,
Sever Paul, Agenția de Investigare Feroviară Română (AGIFER), Romania,
Ana Lisa Vetere Arellano, European Commission, DG-Joint Research Centre, Space, Security & Migration, (JRC), Italy.

13.1 Executive summary

Safety authorities and regulators have a unique role in governance, which provides them with specific opportunities to generate and disseminate foresight. From the viewpoint of safety, this means foresight concerning risks from accidents and other deviations from safety. This chapter describes some daily activities of safety authorities and discusses the current and potential future possibilities of authorities to provide foresight during these activities. We concentrate here on those authorities that work in the field of safety, especially industrial safety, including safety investigation agencies. Parts of the discussion also apply to public organizations who aid authorities in some sectors.

Authorities working in the field of safety in different countries have different mandates, structures and cultures. This means that their perspective may be somewhat different. In case of an accident, some authorities have the responsibility to find out whether there were any violations of laws or regulations and whether there is reason to fine or prosecute, while others have no mandate to investigate issues of responsibility, but rather to find the root causes of the accident and ways to prevent similar occurrences in the future. Despite these differences in mode of operation, all safety authorities have the common endeavour of working towards a safer future.

Foresight is gained through the authorities' numerous contacts with other actors in the field of safety. Multi-level cooperation produces insight in various safety areas, which may then be communicated to the sectors, areas or fields that are under the authority's supervision. The discussions between actors that follow this communication produce foresight that e.g. companies can use to maintain and enhance the safety of its personnel, process and products.

13.2 Introduction

This chapter discusses the role of regulatory bodies working in the field of safety, and the authorities' potential in developing and using foresight – for instance by identifying emerging threats to safety and early warning signs of an accident – and sharing this foresight with interested parties such as industry, other authorities and the general public.

An authority or a regulatory body or can be a public entity or a government agency at a sub-national, national or supranational level that has a mandate established by a specific legal act or acts. Such a body typically aims to protect stakeholders in a given sector by supervising that they follow given norms and procedures. Examples are:

- Sub-national level: Transport Scotland aims to deliver a safe and sustainable transport system for the people of Scotland guided by the Waverley Railway Act;
- National level: Finnish Safety and Chemicals Agency (Tukes) supervises the safety and reliability of products, services and industrial activities in Finland, enforcing e.g. the Act on the Safe Handling and Storage of Dangerous Chemicals;
- Supra-national level: European Food Safety Agency (EFSA) provides scientific advice to protect consumers, animals and the environment from food-related risks under Regulation 178/2002.

With the aim of accident prevention and safety promotion, one of the tasks of safety authorities is to detect and communicate early warning signs. Traditionally authorities have carried out analyses of past accidents in order to learn from them. Lessons learned are shared with companies and the public to give insight of risks that are present in their activities and to verify that these risks are managed in an acceptable manner. Using accident-related hindsight, authorities and other regulatory bodies can gain insight and enhance various areas, such as legislation, guidelines and inspection practices.

Safety authorities have wide-ranging cooperation in different ways with numerous other actors that work in the field of safety: standardization organizations, academia, research institutions, and unions, for instance. Some of these stakeholders are listed in table 1.

Table 1: Cooperation network of regulators

Stakeholders	Regulators interact with many stakeholders
Governmental organizations	Local, regional authorities; Military; Supranational authorities (EU institutions); Standardization organizations; Academia and research institutions
Social and political organizations	Non-governmental organizations; Society (citizens); Media
Economic organizations	Companies; Unions; Parties active in advancement of technology

In some countries, such as the United Kingdom, Singapore and The Netherlands, a shift towards strategic foresight has been observed, attempting to cut across the traditional segmentation of problems and their allocation to specific organizations or departments (Habegger, 2010). These countries have realized that focusing on a single issue at a time makes dealing with emerging threats very difficult. Due to the interconnectedness between social, political, economic, environmental and technological sectors, a multi-disciplinary approach in looking at risks through foresight creates increasing margins for improving preparedness and resilience towards identifying weak signals and risk scenarios. Thus, the work done by authorities can become more anticipatory rather than reactive in approach. It can be seen that *a posteriori* methods are insufficient in capturing weak signals to alert companies on time to better prepare themselves for possible adverse effects.

Given their position and role in the governance of risks, authorities are well positioned to identify general industry-wide or societal trends that are likely to lead to safety degradations in the future. They are notified of incidents and can undertake trend analysis. They can observe the evolutions of external constraints (economic conditions, trends in the societal acceptance of specific hazards…) and anticipate their impact. Some examples of this in recent years:

- The rapid increase in the popularity and technical capabilities of civilian drones poses increasing challenges to air traffic control. Regulators in the EU, USA, and other countries are working to change regulations concerning the use of civilian drones to reduce the risk of collisions with general aviation traffic (Cracknell, 2017), and to improve both detection capabilities and enforcement of zoning regulations, including developing drone destroying capabilities[117].
- The increased availability of high-power laser pointers poses significant hazards for general aviation pilots (airplanes, helicopters), with many cases of pilots being blinded by pointers during airport approach. Regulators may be able to identify a trend in these new forms of threats and work with airlines to find risk mitigation strategies.
- The chemical industry has seen a trend to reduce the quantities of hazardous materials stored on site, following inherent safety principles. This leads to an increase in the transport of hazardous materials, with consequences for safety of road/rail transport that can be anticipated by authorities. Such effects of new trends can and should be identified early on to avoid unintended consequences.
- Electric scooters have become more and more popular, also with adults in their daily movements. This may mitigate air pollution in big cities, but the scooter speeds of 25-30 km/h cause new problem areas that should be taken into account in e.g. appointed areas to use scooters, traffic rules and accident insurance.

13.3   Types of foresight-enabling activities

13.3.1   Foresight possibilities during daily work.

In order to be effective in carrying out foresight within a safety authority, it is essential to have foresight-enabling activities in place during daily work. This would build and strengthen a *foresight looking culture* that would help stakeholders become anticipatory in managing safety. Foresight should promote thinking about

---

[117]*For instance, the* regulatory framework timeline *of EASA on this issue extends into 2023.*

and visualising future alternative scenarios, whilst engaging stakeholders to actively work together to debate the future and contribute to shaping it[118].

When carrying out daily work there are various activities that could enable foresight culture in safety authorities:

- Setting up and sustaining a systematic knowledge base. As there are many forms of knowledge storage (emails, forum discussions, social media exchanges, local and shared drives, cloud, working groups, experts, etc.), having one reference point can be beneficial. Depending on the topics that need to be monitored, stakeholders should co-design processes to systematically capture, index, store, curate, analyse, visualise, apply and disseminate knowledge (old and new; tacit and explicit). Building a knowledge management culture is a necessary component to enable a foresight culture.

- Building a systematic data analytics capability. By having analytical processes in place to systematically understand the past, i.e. what happened (descriptive analytics); to gain the insight on why it happened (diagnostic analytics) and what will happen (predictive analytics) [see also Chapter 10], safety authorities would routinely be able to detect trends, patterns and emerging change. This insight generation process would foster informed decision making and increase foresight capacity of an organisation.

- Encouraging organisational learning. When a safety authority motivates a systematic implementation of knowledge management processes within its organisation, knowledge is embedded into its organisational processes. This way the organisation builds its continuous learning capacity (in terms of practices and behaviours) whilst achieving its corporate objectives. In other words, organisational learning is a sustainable way to improve knowledge utilisation. Building a lessons learning culture is a necessary component to enable a foresight culture.

- Reducing learning barriers. Safety authorities should bring stakeholders together to regularly identify learning barriers and find ways to overcome them. This will help foster organisational learning. Addressing this will enable a foresight culture.

- Putting more emphasis on organisational factors in the development of early warning signs. [see also Chapter 8.] Safety authorities often read lessons learned reports. It can be observed that such documents tend to focus on more technical causes and lessons learned. These result in the development of more technical-oriented indicators that help identify early warning signs instead of organisational-oriented ones. However, when reading such reports, it can be observed that there are underlying organisational factors that have not been explicitly communicated. Focus on the latter will help increase capacity to capture more tacit knowledge.

- Promoting inclusive multi-stakeholder and multi-disciplinary teams. When a strategic foresight related issue needs to be addressed, safety authorities should, as a habit, include in this process many stakeholders from various disciplines. Doing this regularly would ensure a more effective design of a foresight strategy and action plan, thanks to dialectic debate and inclusion of different perspectives.

13.3.2 Inspections and site visits

Authorities can adopt different strategies when interacting with regulated organisations such as operating companies, ranging from a "policeman" attitude which is focussed on identifying gaps between practices on a site and the regulatory requirements, to an "advisor" role which involves discussion with operating companies on how to interpret the regulatory requirements and strategies to attain compliance and improve safety even further. In the academic literature, this differentiation in attitude depending on what inspectors perceive of the motivation of the operating company is called *responsive regulation* (Ayres & Braithwaite 1982). When operating in an advisory role, which leads to richer interactions between inspectors and companies, inspectors are more likely to generate foresight than by operating in a "policeman" role.

Inspectors can help the operating companies identify where procedures, tools and systems could be improved. The challenge for regulators is to aid the company to improve its foresight capacity. As an example, the inspector can look at the elements of the company's training program as training employees in hazard identification and reporting will lead to better insight of the safety status, discussions of risks and needs to improve, and through these, foresight.

---

[118] http://www.foresight-platform.eu/community/forlearn/what-is-foresight/

Safety authorities can identify the fact that the technology on a site is lagging far behind the state of the art and suggest or mandate changes. Regulations that are expressed as obligations to implement best available technologies when possible help make this approach systematic. However, new obligations are bounded by a legal framework and the objective of avoiding adding unnecessary burden.

Safety authorities can audit/inspect the systems in place in operating companies for handling events and detecting warning signals and the organizational and cultural features that are known to be necessary for a learning culture and a mindful organization. For example, they may be able to detect underreporting of significant events by talking with front-line workers about the incidents they experience and the accidents that they remember, and to compare these with the formal record contained in the company's experience feedback database.

Authorities help generate foresight when they distribute to e.g. small and medium-sized enterprises good safety practices, information of identified risks and other safety aspects they have identified during inspections to pioneering companies.

In some cases, where for different systems (aviation, railway, maritime), same regulation stipulates the condition for medical and psychological examination for employees with responsibility in safety traffic (e.g. Romania), the safety and investigation authorities should cooperate and exchange information if an accident occurred having as cause or as contributing factor, an issue relating with the provision of that regulation.

13.3.3 Feedback to legislation

The activities safety authorities and regulators perform have a strong link to legislation. Regulators and other safety authorities may have a mandate to write, monitor, support or update the legislation and associated regulations. Outdated regulations are sometimes a contributing factor in large accidents, where new technologies on the market are not covered by existing regulations and are used without sufficient thought concerning the safety design and related impacts. An example is provided by the fire at Grenfell Tower (London) which caused multiple fatalities in 2017. Companies involved in the refurbishment of the building used combustible building materials which were not specifically covered by existing building codes. (Grenfell Tower Inquiry, n.d.)

"Policy development often lags well behind technological advances" (Lee, 2019). By working with legislators in a proactive manner, safety authorities can help minimize the temporal lag between the appearance of new risks and the development of appropriate regulations.

Likewise, regulatory bodies need foresight to minimize the lag between societal changes and legislation required to protect society and the environment. For instance, increasing life expectancy is likely to lead to increases in the retirement age in most countries, and to the presence of older workers in the workplace. This change may require modifications to labour laws and related ergonomics standards, for example to account for physical differences, and to government support for lifelong learning programmes.

As soon as a new scientific invention reaches the regulatory bodies, they should be working towards getting that into legislation. There should be certain specific groups that focus on anticipating new innovations, designing processes that inject new knowledge into legislation in a timely manner (horizon scanning).

13.3.4 Market surveillance

Market surveillance done by regulators is targeted at e.g. products that are on the market and available to consumers, product documentation, markings and labels on products, and procedures for demonstrating compliance. Surveillance can be either risk-based or based on random selection. Regulators aim to find and check products with the greatest safety risks as they are only able to check a small proportion of all the products that are on the market.

The authorities have insight through EU cooperation concerning safety-related events, and they can use it to help companies improve foresight. In early 2019, the Finnish Safety and Chemicals Agency carried out a survey on the safety of so-called escape rooms — a popular game for kids and adults —based on a dramatic accident that had occurred in Poland. As a result, Finnish escape room operators significantly improved customers' ability to leave the room in case of a real emergency (Tukes 2019).

The authorities may also use the insight they have to inform consumers to make safer choices — and to gain the needed foresight to do this: In a project called "At your own risk" (Tukes 2018/1) numerous Finnish authorities and other organizations worked together to inform consumers of the responsibility and risks they take when purchasing products from outside the EU regulatory framework. In this project in 2018, e.g. the Finnish Safety and Chemicals Agency tested

products it had ordered online from countries outside the EU. Only 1 out of 32 products fulfilled European requirements (Tukes 2018/2).

13.3.5   Accident database management

National and local authorities often maintain accident and incident databases. Through these databases the authorities can monitor the level of safety on a larger scale and identify trends or e.g. new issues that raise concern. As different safety and investigation authorities have different mandates, their accident (incident, near miss) databases, knowledge and knowhow complement each other. Through accident-related information exchange the authorities can obtain a better view of the situation as a whole. Additionally, with the rise of open data, big data and related methods, existing information can be combined to produce new information that the authorities and others can use to help companies gain foresight to identify emerging problem areas and prepare for new risks.

International databases contain information about risks identified in other countries, and/or different activities (systems) making it possible for the authorities and others to learn from major incidents that have already been realized elsewhere.

When they run an incident database, the regulator is well placed to identify rare events. For example, the icing threat on certain types of pitot tubes that was a causal factor in the AF447 Rio-Paris crash had been detected by EASA prior to the crash. EASA had not yet decided to mandate a change to the equipment, but some airlines had decided to replace the pitot tubes by another model which was thought to be less susceptible to high-altitude icing. In the case of the affected aircraft, Air France was in the process of replacing the pitot tubes, but the change had not yet been implemented on that specific aircraft (BEA 2012).

Some European databases that are utilized widely include:

- The European Commission has established a Clearinghouse for collecting and analysing operating experience from nuclear power plants in order to provide feedback for EU regulators to improve nuclear safety (https://clearinghouse-oef.jrc.ec.europa.eu/).
- The European Commission maintains the Major Accident Reporting System (eMARS), houses lessons learned reports of chemical accidents and near misses from EU, EEA, OECD, and UNECE countries. eMARS event

reporting by EU Seveso Competent Authorities is mandatory. (https://emars.jrc.ec.europa.eu)

- The French Ministry of Environment/General Directorate for Risk Prevention developed the Analysis, Research and Information on Accidents (ARIA) database, which contains (https://www.aria.developpement-durable.gouv.fr/the-barpi/the-aria-database/?lang=en
- The German database Zentrale Melde- und Auswertestelle für Störfälle und Störungen in verfahrenstechnischen Anlagen (ZEMA) contains annual reports of all events which must be reported to the authorities pursuant to the 12th Federal Emissions Control Ordinance. (https://www.infosis.uba.de/index.php/en/index.html)
- Even smaller countries with less resources are able to develop a simple database. This is illustrated by the Belgian Database "Lessons of Accidents" of the Belgian Competent Authority for Seveso Industries (https://emploi.belgique.be/fr/themes/bien-etre-au-travail/seveso-prevention-des-accidents-majeurs/publications-sur-la-5)

13.3.6   Horizon scanning and adversarial approaches

Given their position and role in the governance of risks, safety authorities are well positioned to identify general industry-wide or societal trends that are likely to lead to safety degradations in the future. They may be able to detect signs of practical drift or normalization of deviance (Vaughan 1996) through their interactions with companies, thanks to their "outsider" view and their mandate to provide a critical analysis. Indeed, authorities typically work with a range of company roles within high-hazard industry sectors, as well as with representatives of civil society (local government officials, members of local communities). Authorities are notified of safety-related incidents, participate in audits and investigations, and can undertake trend analysis. They can observe the evolutions of external constraints (economic conditions, trends in the societal acceptance of specific hazards…) and anticipate their impact.

The following points illustrate horizon scanning and adversarial activities that can be undertaken by authorities and other organizations:

- Brainstorming sessions amongst inspectors after an accident or near miss to come up with a list of questions to ask that may enable to identify

establishments with similar root causes and problems. These discussions can be enriched by operational experience feedback data, and can help to identify potential pathways to an accident (scenario-based approach) that have not yet been identified. These could be the basis for the development of lagging indicators.

- Brainstorming sessions amongst inspectors to come up with a list of leading indicators.
- Implement "red team" type exercises with volunteer firms. These adversarial exercises, developed in the military planning sector, consist of establishing two teams, a "red" team which searches for "holes" in the organization's defences and event sequences that can lead to an accident, and a "blue" team which is responsible for defence (Bloomfield & Whaley 1963). These exercises actively challenge an operating firm's beliefs and the assumptions underlying its risk analyses, and can help reduce complacency. A well-known example of this practice is the stress tests used by financial regulators to ensure that banks and insurance companies have sufficient capital reserves to deal with extreme events[119]. A similar approach is taken to review the safety of nuclear power plants in Europe with stress tests defined and organized at the EU level after the Fukushima Daiichi accident. This included self-reporting on defined questions from national regulatory bodies and independent public expert review and conclusions.

With this information, authorities are in a position to:

- Provide additional guidance to operating companies, warning them of emerging risks and potential preventive actions. An example of this activity is provided by the UK Health and Safety Executive's annual Foresight reports (UK HSE 2018).
- Update their inspection checklists to integrate new threat types.
- Suggest changes to regulations that can be made by the legislator.

13.3.7   Insights from research, other organizations and the industry
Efforts to improve risk governance exist in all domains (e.g. transportation, process and the food industry) where modern government tries to reduce cost and assure safety and benefits. Risk governance approaches and experience from different domains contains universally useful values for improving regulatory efficiencies and foresight (IRGC, 2017). This cross-domain exchanges could be useful for different elements of regulatory framework, i.e. to improve policy, legislation, enforcement, inspections, experience feedback practice and foresight. Broadly, there are three sources of risk regulation experience from different domains potentially useful to improve risk governance and achieve social and environmental goals, i.e.: 1) research; 2) international organizations; and 3) similar hazardous industries (IAEA 2020).

Research about risk governance (policy development, implementation and regulatory process) exists in all domains and its results could be universally valuable. The value of research could go from developing approaches to evaluate regulatory efficiency to identification of best practices to improve specific regulatory features and include evaluation of the regulatory framework designs.

Regulatory efficiency could be measured by assessing costs and benefits of regulation which could be useful for deciding about introducing specific regulatory requirements or selecting alternative approaches, e.g. Robinson at. al. 2008. It is important but also challenging to measure efficiency and effectiveness of regulation with interconnected impact of co-existing regulatory features, e.g.: communication, consultation, consistency, flexibility, independence, accountability and transparency (Berg, 2001). Identifying and explaining best practice can help improve regulatory efficiency.

Comparison of regulatory framework designs can contribute to improved regulatory decision making. Different regulatory designs aim to enforce compliance and to improve efficiency from collaboration with regulated organisations. Systematic empirical research into the applicability and effectiveness of different regulatory types for different problems and under different conditions is lacking (NASEM, 2018).

International organizations like OECD and the European Commission (EC) are facilitating risk governance experience exchanges from different domains. The EC has organized workshops with regulatory experts from different domains, e.g.

---

[119]These stress tests are run by the European Banking Authority in collaboration with the European Systemic Risk Board, the European Central Bank and the European Commission.

shipping, aviation and nuclear industry (EC 2008). This kind of activity presents opportunity for exchange of applicable best cross cutting regulatory practices.

OECD facilitates exchanges between many different regulatory domains in order to improve policy and governance. The OECD Council on Regulatory Policy and Governance published recommendations and tools for effective and efficient regulatory policy, governance and management (OECD 2012). Another OECD report provides guidance on improving regulatory enforcement and inspections with examples of good practices and principles, i.e.: evidence-based enforcement, selectivity, risk focus and proportionality, responsiveness, long-term vision, transparency, information integration, fairness, compliance promotion, and professionalism (OECD 2014). The OECD has also developed indicators of regulatory policy and governance covering three principles (stakeholder engagement, regulatory impact assessment and ex-post evaluation) and providing a baseline measurement to track status and progress[120].

More specific experience and more directly applicable insights are coming from domains (industries) which share some similarities. The governance of safety in all hazardous industry could produce experience and insights applicable in other domains. This could include all different aspects from regulatory organization and activities to specific technical and human. Findings from investigations of major accidents in other hazardous industries should be included within the scope of experience feedback. The role of the regulator was assessed e.g. for offshore safety following the Macondo disaster (Weaver, 2014). Novel activities from regulators in other hazardous industries could be considered for adoption and as a source for improvements. For example, process industry regulators from different EU countries organized mutual joint inspections in order to exchange insights and best practice (Wood, 2014).

Risk governance experience insights from different industries, especially high hazard ones, are potentially universally applicable and should be regularly reviewed. Existing activities and available information from different industries could help risk governance in many segments (from policy development, through implementation, and regulatory process) including foresight. There are however many challenges to fully utilizing all these potential opportunities related to the applicability of findings, uncertainties of results, and the need for additional resources. International organizations like the OECD and EC are providing arrangements to identify, scrutinize and disseminate such cross domains risk governance experience.

13.4 Conditions for success

13.4.1 Authorities working with companies

Compared with the classical operating mode of many safety regulators, which is primarily focused on verifying compliance with prescriptive requirements and investigating incidents and accidents, the adoption of foresight-informed approaches often involves changes to the way in which the authorities operate and interact with regulated organisations, legislators and the public. This foresight-informed regulatory approach also requires the development of new competencies for the authorities' staff.

To increase foresight, authorities will need to adopt a cooperative relationship with operating companies, advising and working in collaboration towards safety, rather than a relationship focussed on enforcement alone. This requires the development of trust, openness and positive collaboration, all features which cannot be imposed but rather which develop with sustained effort over time. Both actors need to foster a constructive and open safety environment where early warning signs can be identified and dealt with in a transparent and efficient manner. Here the safety culture of the regulatory organization also plays a role (NEA 2016).

Effective foresight development also requires specific skills and competencies for the inspectors and other regulatory personnel, such as the ability to anticipate risks, knowledge of methods such as scenario development and horizon scanning, and communication skills. Maintaining and developing these competencies requires specific attention at the organizational level. These competencies are easier to develop in authorities which maintain specialist expertise in the areas they are overseeing, rather than delegating part of their supervisory authority to industry personnel (delegation of this type by the US aviation regulator has been

---

[120]*Indicators are available online for 2015 and 2018 at www.oecd.org/gov/regulatory-policy/indicators-regulatory-policy-and-governance.htm.*

heavily criticized after the Boeing 737 Max disasters in 2018 and 2019). Also the interface between the company and the regulatory body should be developed.

Regulators must follow up on their recommendations to ensure that they are implemented within an appropriate timeframe. Otherwise, companies may omit to implement recommendations, as for example at BP Texas City, where OSHA had identified a number of safety management deficiencies during various inspections but did not enforce their recommendations (CSB 2007, page 20).

It is important to note that the role of regulators in generating and disseminating foresight is not necessarily positive: if a regulatory body is excessively conservative, and does not work towards modernizing legislation, the regulatory framework and inspection practices, if the regulatory body promotes an outdated view that equates safety with compliance (minimum demands fulfilled), it may constitute an obstacle to foresight activities within companies, by preventing the implementation of novel technologies and organizational strategies.

13.4.2 Companies working with authorities
Between the authority and the company it supervises there should be regular discussion and follow-ups of lessons learned, with a focus on near misses. The questions addressed in these discussions should also be discussed inside the company.

- What could have been done to prevent this near miss: at individual, organizational, corporate level?
- What can we monitor (which indicators) on a regular basis to ensure that we could anticipate and prevent such an incident from occurring?
- Who are the key actors to ensure that such a process is designed, with an integrated follow-up mechanism?

Once indicators have been identified, encouraging all actors (company, inspectors, and other stakeholders) to continuous learning:

- ensure awareness on this issue (continuous culture building and follow-up of its effectiveness);
- know what to effectively do when such EWS are detected for continuous knowledge building to ensure long-lasting imprinting.

Regular discussions and follow-ups of possible scenarios (e.g., known unknowns and unknown unknowns) is a useful exercise to make people aware that these types of events can occur and discussions such as these increase collective knowledge and awareness about the establishment. Thus, a positive environment for an "early warning sign detection" culture.

13.5 Conclusions

Authorities have a unique role in the governance of safety, which provides them with opportunities to generate and disseminate foresight. Authorities are able to identify trends and new threats to safety due to their ability to have: integrated view of the status of regulated activities; collect and review events that occur in a large number of companies; and to observe interactions with a multitude of other actors (such as research organisations, unions, citizens and other relevant authorities). Safety authorities have channels that help disseminate foresight and lead to changes in safety management both within companies and at the regulatory level.

Safety authorities can produce and disseminate foresight through their interactions with actors at different system levels and in different industry sectors, as a part of many different activities such as inspections, events' trend analysis, work on regulations, market surveillance, and more currently, horizon scanning and adversarial exercises.

From hindsight to insight to foresight: learning from the past, combined with multidimensional analyses assists in looking into the future and identifying the possible roads to follow. The possible obstacles on those roads will lead to new viewing angles to identify both existing and emerging risks. Here the role of the safety authority is that of a facilitator and enabler: when the viewpoint of foresight is included in the regulator's daily activities, the discussions between the authorities and the organisations it interacts with will generate new possibilities to maintain and improve safety.

13.6 References

Ayres, I. & Braithwaite, J. 1992. Responsive regulation: Transcending the deregulation debate. New York. Oxford University Press, ISBN: 978-0195093766.

Berg, S. Infrastructure regulation: risk, return, and performance. Global Utilities, 1 (May), 3-10, 2001

BEA. 2012. Rapport final. Accident survenu le 1er juin 2009 à l'Airbus A330-203 immatriculé F-GZCP exploité par Air France. Vol AF 447 Rio de Janeiro - Paris. French Bureau d'enquêtes et d'analyses (BEA).

Bloomfield, L. and Whaley, B. 1963. The political-military exercise: a progress report. Center for International Studies, MIT.

CSB 2007. Investigation Report. Refinery explosion and fire. https://www.csb.gov/file.aspx?DocumentId=5596 [Retrieved 27 May 2020]

Cracknell AP. UAVs: regulations and law enforcement. International Journal of Remote Sensing. 2017 May 19;38(8-10):3054-67.

EC 2008, DG Energy and Transport, Cross cutting comparison of regulation and operation of industries requiring specific safety rules, Workshop Summary, EUR 23203.

Grenfell Tower Inquiry. N.d. Available at https://www.grenfelltowerinquiry.org.uk/ [Retrieved 15 July 2020]

Habegger, B. 2010. Strategic foresight in public policy: Reviewing the experiences of the UK, Singapore, and the Netherlands, in Futures 42 (2010) 49–58, Elsevier.

IAEA 2020, Effective Management of Regulatory Experience for Safety, International Atomic Energy Agency, IAEA-TECDOC-1899.

IRGC 2017. Introduction to the IRGC risk governance framework, revised version 2017. Available at https://infoscience.epfl.ch/record/233739/files/IRGC.%20%282017%29.%20An%20introduction%20to%20the%20IRGC%20Risk%20Governance%20Framework.%20Revised%20version..pdf [Retrieved 15 July 2020]

Lee, G. 2019. What roles should the government play in fostering the advancement of the internet of things, Telecommunications Policy, Volume 43, Issue 5, June 2019, Pages 434-444

National Academies of Sciences, Engineering, and Medicine. (NASEM) 2018. Designing Safety Regulations for High-Hazard Industries. The National Academies Press.

NEA 2016. The Safety Culture of an Effective Nuclear Regulatory Body. NEA report number 7247, OECD Nuclear Energy Agency.

OECD 2012, Recommendation of the Council on Regulatory Policy and Governance.

OECD 2014, Regulatory Enforcement and Inspections, Best Practice Principles for Regulatory Policy.

Robinson L.A., Nriagu J. Assessing regulatory costs and benefits. 2008. Nriagu, J. (ed.). Encyclopaedia of Environmental Health.

Tukes 2019. Survey on the safety of exiting escape rooms in case of emergency completed. Available at https://tukes.fi/en/article/-/asset_publisher/pakohuoneiden-poistumisturvallisuuden-kartoitus-valmistui [Retrieved 7 Mar 2019]

Tukes 2018/1. At your own risk. Available at https://tukes.fi/en/at-your-own-risk [Retrieved 7 Mar 2019]

Tukes 2018/2. Tukes tested products from cheap online shops: nearly all of the toys, child care supplies and electrical appliances were dangerous. Available at https://tukes.fi/en/article/-/asset_publisher/tukes-testautti-halpisverkkokauppojen-tuotteita-lahes-kaikki-lelut-lastentarvikkeet-ja-sahkolaitteet-vaarallisia [Retrieved 7 Mar 2019]

UK HSE. Health and Safety Executive Foresight Center 2018, Foresight report: The future world of work and workplace health, available at www.hse.gov.uk/horizons/assets/documents/foresight-report-2018.pdf

Vaughan, D. 1996. The Challenger launch decision: Risky technology, culture and deviance at NASA. Chicago. University of Chicago Press, ISBN: 978-0226851754.

Weaver J. L. Offshore safety in the wake of the Macondo disaster: the role of the regulator. Houston Journal of International Law 379. 2014

Wood M. The mutual joint visit programme on inspections under Seveso II: exchanging lessons learned on inspections best practices. Institution of Chemical Engineers Symposium Series, Vol. 150, pp. 977-94. 2014.

Conclusion

ESReDA Project Group on 'Foresight in Safety'[121]

This conclusion sets out the key messages of the ESReDA Project Group on Foresight in Safety. Please note that each of the other thirteen chapters has its own conclusions.

Overview

The ESReDA Project Group on Foresight in Safety ("the Project Group") is a diverse team of researchers and practitioners. Safety is a multidisciplinary field, and works by exchanging different visions and approaches. The context of foresight in safety is ably summarised in Figure 2.1 of Rasmussen and Svedung (2000)[122]. Although multidisciplinary, some of the knowledge in the field of safety arises from the efforts of researchers working in their own discipline. That means that pretty well everyone else working in the field is either integrating this knowledge, or applying it, or both. Against this background, the Project Group has asked how the concept of foresight applies in safety and what challenges exist.

The word 'foresight' is not new in safety, but neither is it settled. In fact, the connotations of the word are evolving and contended. Since Roman times, the concept of foresight has been used in law to decide matters of blame and causation after harm has occurred. Those legal cases focus on whether the event was itself foreseeable and whether enough effort went into foreseeing it and avoiding it. Although commonplace, foresight resists exact definition or description as a function or capacity. Nevertheless, knowledge can still be shared about the conditions that govern safety foresight and the processes that achieve it.

Collating the Project Group's key messages brought to mind a quotation from Santayana[123]. The first sentence is very familiar: "*Those who cannot remember the past are condemned to repeat it*". The quotation continues, "*In a moving world readaptation is the price of longevity…*" and that our institutions must give "*birth to a generation plastic to the contemporary world and able to retain its lessons*" (Santayana, 1905).

Foresight: Dynamic, Not Static

Foresight is a projection based on our knowledge and beliefs at a given moment. But, new data or further reflection may well change the possibilities we foresee. It is usual to have less knowledge when committing to a particular design or policy, than later when the results of our decisions unfold. By shoring-up our provisional arrangements in the light of new information, we can mitigate the paradox of 'learning later' but having to 'commit now'. The principle for foresight is to remain skeptical and critical, permanently ready to update our models and challenge assumptions. But this principle and the related mindset is costly and it is not without its practical challenges.

Change is a basic concept in safety and unites all its branches (process safety, occupational safety, etc.). The concept can be found in textbooks and programme reports from the 1960s to the present day. At its simplest, safety sees any change as '*the mother of twins: progress and trouble*' (Johnson, 1980[124]). Foresight is used to keep an eye on, and head-off when necessary, the troublesome twin. However, by the 1970s, the rapid rate of technological change was recognised as a new phenomenon in its own right, and our old tools of foresight seemed inadequate. By the 1980s, it was recognised that when change is discontinuous, previous technological precedents may be irrelevant to foresight or even misleading. All these views remain current, and have implications for the practice of foresight in safety.

Foreseeing accidents and trouble from incremental change is more *retrospective*; whereas radical, discontinuous change relies more on creative, *prospective* foresight. Examples of the latter include 'gene driving' technology that forces genetically engineered changes in individual organisms to be expressed with high

---

[121] *The conclusion has been prepared by John Kingston, Ana-Lisa Vetere Arellano and Yves Dien on behalf of the project group.*

[122] *Rasmussen J. Svedung, I (2000); Proactive Risk Management in a Dynamic Society, Swedish Rescue Services Agency. (https://www.msb.se/RibData/Filer/pdf/16252.pdf#page=10 – Retrieved on July16, 2020)*

[123] *Santayana, G. (1905) The Life of Reason. [online: http://www.gutenberg.org/files/15000/15000-h/15000-h.htm]*

[124] *Johnson, W.G. (1980) MORT Safety Assurance Systems, Edited by Marcel Dekker, Inc.*

likelihood in subsequent generations. Another is the difference of chemical properties between regular and nano particle sizes. For instance, nano gold is a poison, whereas regular gold is biologically almost inactive. Foresight of radical or disruptive change is challenging, even more so when changes interact. In many areas, notably technology, it is increasingly unsafe to assume that the near future will be an extension of the past. There is a pressing need to enable foresight in such systems.

In areas characterised by rapid or discontinuous change, foresight can be blinded when technology and organisation are seen in isolation. Within safety, the term 'socio-technical' has almost become a cliché: often used, but superficial and patchy in its application. In practice, however, technology and organisation appear often to be managed, researched and educated, as two separate domains. This separation creates a void in foresight, which needs to be open to the safety consequences revealed by both perspectives and their interaction. The challenge of rapid and discontinuous change requires the 'sociotechnical view' to be refreshed, as Rasmussen and Svedung called for twenty years ago. An example of an approach that does this is the ESReDA Cube model[125].

In summary, foresight in safety:

- *is a continuous process, because knowledge and systems continually change;*
- *has difficulties seeing possibilities created by radical, discontinuous change;*
- *integrates the social and the technical knowledge of systems at every level;*
- *is applied skepticism.*

Foresight: A Multi-Actor Activity

We take it as axiomatic that foresight about safety improves when several perspectives are shared and debated by different actors. However, there are multiple challenges. Mostly, these stem from the messy reality of foresight activity in the practical world. Foresight is as much an issue of *agency, structure, power and influence*, as it is a function of expert knowledge, experience and method. This is true both of organisations and society in general. For all these challenges there are solutions, but only if we recognise that there is a need for them.

Within foresight in safety, as in risk management generally, the 'stakeholder' concept is increasingly recognised as relevant. A stakeholder is, according to Freeman (2010), '*…any group or individual who can affect, or is affected by, the achievement of a corporation's purpose'* [126]. Stakeholders can be within the risk-owning organisation or outside of it, but Freeman's definition implies that stakeholders are within the overall system.

Where the corporate capacity for foresight is at stake, appearances of consensus are deceptive. Within organisations it is usual for individuals and groups to have a range of different opinions about future possibilities. An organisation is not a "monolithic whole". Within an organisation different visions coexist concerning the way the "system" is working, its level of safety, and unsafe functioning or unsafe acts. By the same token, except in the simplest cases, foresight cannot be monolithic. However, organisations invariably adopt single positions on matters, albeit hedged with contingencies. This practice is pragmatic and expedient; it allows the organisation to make progress, to get on with providing its services. Foresight, in contrast, is more like a competition between individual visions of the future than a common denominator of these visions. Furthermore, the best informed vision is not necessarily always the winner, because influence and power also count in the competition of ideas.

Therefore, the apparent consensus in an organisation's risk assessments and policy documents needs to be treated with caution. Specific risk analyses and policies will

---

[125] *The ESReDA Cube is a conceptual model focused on the "learning from accident" process. It represents learning as a three dimensional space taking account of "what needs to be learned", "who should learn" and "how it is learned". The Cube was developed by ESReDA Project Group "Dynamic Learning as the Follow-up from Accident Investigation". This document can be found at*

https://esreda.org/wp-content/uploads/2016/03/ESReDA-dynamic-learning-case-studies-180315-1.pdf

[126] *Which is defined here as "…any group or individual who can affect, or is affected by, the achievement of a corporation's purpose". Freeman, R.E. (2010). Strategic Management: A Stakeholder Approach. Cambridge University Press.]*

be based on facts known with different levels of confidence. Some of the facts will be indisputable, but not all of the facts. Yet, decisions and policies have to speak with one voice. However, what they say is the product of compromise and uncertainty rather than a ringing consensus about the meaning of immutable facts. Therefore, decisions and policies are invariably simplifications, and may become *oversimplifications* unless reviewed. The danger in such 'faux consensus' is that it lures senior decision-makers toward complacency, because it suggests that a matter is settled. In contrast, for the sake of foresight, the matter is best treated as a momentary stock-take in the continuous, diligent search for future possibilities. Foresight is *now*, based on the best knowledge today; not as we saw the matter yesterday. As well as evolving knowledge, stakeholders and situations change. This means that a single position can only be tenable in the short term. Organisations need to support foresight as a continuous exchange of perspectives, even dissenting views, within their communities.

An assumption of the multi-actor view is that different actors can communicate and debate. If actors are to agree about foresight and early warning signs, there must be some measure of shared knowledge about how things work. In industrial safety, this usually means that actors share a level of technical knowledge of their organisation's operations. For example, in asset management, a field related to safety, the relevant standard[127] notes that shared technical knowledge helps top management make sound, well-supported decisions.

Open debates favour safety, but very few actors will coincide on all points. As mentioned, consensus about foresight and decisions is an ideal, but the path to it is paved as much by disagreement as agreement. In contrast, some organisations tend to treat disagreement as poor conduct. In general, to avoid being seen as trouble-makers, individuals will abandon defence of their viewpoint. Therefore, we should not be surprised when people speak out only *after* the accident they foresaw. It is not reasonable to expect heroism and self-sacrifice to be the safety backstop for cultures that discourage individuals from giving voice to foresight.

Historically, we should discriminate between two groups of individuals who open up debates: as those individuals who warn against mishaps based on their professional judgement. First, mostly engineers who understand the design assumptions and limitations. Their judgement is often based on evidence but not always... Secondly, people who criticise the appearance of phenomena they do not fully comprehend or are only partially informed, but base their judgement on social media and 'influencers'.The COVID 19 situation has demonstrated a public debate on social distancing, herd immunity and personal protective equipment.

Position and qualifications are, however, not an infallible guide to who has valid foresight (the response to the Covid-19 crisis illustrates this, too). The reliability of knowledge has never been easy to assess, and it is increasingly challenging. As a source of of facts social media deserves caution, but so do claims of any kind, including those made in peer-reviewed journal articles. Furthermore, the characteristics of the knowledge underlying foresight varies with the context. Sometimes, the situation allows testable answers to black-and-white questions. But, at other times, our knowledge is far less categorical and the problems open-ended. Describing the latter, Hardin (2002[128]) writes "*In an economic theory, it makes sense to say that you know one thing and I know a contrary thing in some context. I might eventually come to realize that my knowledge is mistaken and therefore correct it, especially after hearing your defense of your contrary knowledge. But there is no role for a super-knower who can judge the truth of our positions. We are our own judges. If we wish to seek better knowledge, it is we who must decide from what agency or source to seek it*". In safety, the situation often arises that risks created by one group are borne by a different group; pollution risks, for example. Often, the group exposed to the risk base their foresight on anecdotal observations and general knowledge. Through self-education and professional assistance, such groups arrive at a point where the content of their claims can no longer be dismissed[129].

Another issue is when views about what constitutes reliable knowledge are incommensurable. Although it can be said that *'Everyone is entitled to his own opinion, but not his own facts'*[130], dispute about facts—and even what constitutes

---

[127] *ISO 55000:2014, section 2.5.2.*

[128] *Hardin, R. (2002) Street-Level Epistemology and Democratic Participation. Estudio/Working Paper 2002/178.* *http://www.march.es/ceacs/ingles/publicaciones/working/archivos/2002_178.pdf*

[129] *A recent illustration from the field of patient safety can be found in the report of the UK Independent Medicines and Medical Devices Safety Review. HMSO (2020) First Do No Harm: The*

report of the Independent Medicines and Medical Devices Safety Review. [online at: *https://www.immdsreview.org.uk/downloads/IMMDSReview_Web.pdf]*

[130] *Attributed to Daniel Moynihan:* *https://en.wikiquote.org/wiki/Daniel_Patrick_Moynihan*

a fact—is a recurring feature in contested foresight. As stated earlier, influence and power also count in the competition of ideas—but who wields the power is not always obvious, neither are differences of epistemology.

We recognise that suppressing dissent and disagreement may sometimes blind the organisation's foresight of credible future accidents. Therefore, as well as shared technical understanding, we need also to encourage, and not suppress, the expression of different views. We need to be able to disagree well[131]. As Espejo[132] points out, *'a consensual domain is none other than the play of a particular set of interacting models'* (1989; 445-446). However, as suggested, in many organisations, the suppression of dissent (including self-censorship) is, unfortunately, normal.

Order in society is sustained by various forces. However, one of the effects can be the discounting of views, even data, that do not fit with the current orthodoxy. Well-investigated accidents show that this riddles foresight with blind spots. A question for the practice of foresight is how to better tolerate and enfranchise dissident voices within our organisations and social structures. At present, the public record contains an ever-expanding file of whistleblowing cases showing that many organisations are immature in this respect. And outside of industrial safety, in the wider realm of social goods and social ills, the question is just as relevant. Although beyond the scope of this work, new models are appearing to support constructive debate and the decision-making authority of institutions. We note, for example, the operation of citizens assemblies in Ireland[133] and elsewhere.

Foresight requires flexible approaches to anticipate the 'unthinkable'. Assumptions are constraints on the range of possibilities from which foresight proceeds. Constraints make foresight possible (in the unconstrained system, *everything* is possible!) but these assumptions will also rule-out some possibilities that may, in fact, be valid and worth thinking about. An investigator remarked of his own practice: *"I must think the unthinkable even if I dismiss it on the basis of evidence'. At least think about it."*[134] In all professions that contribute to safety,

this self-honesty appears to be fundamental to extending foresight. We must try hard not to fool ourselves. But there are various disincentives, such as our credibility in the eyes of one's peers, a desire to be seen as a team player, and a wish to avoid the discomfort of dissonance.

The Project Group noted that the visualisation of hidden or weak signals has an important role in predicting possible incidents and accidents. The etymology of the word invites us to think about *foresight* as *visions* of the future. However, as the foregoing discussion makes clear, it is helpful to consider foresight more as a *process* in which stakeholders strive to communicate, debate and make change happen. It brings to mind the advice: "As visual metaphors never perfectly fit the target domain, they also trigger sense making and discussions about the risks and the shortcomings of the chosen metaphor. In this way they help to clarify risk understandings in groups by sparking lively debates" (Eppler and Aeschimann, 2009; p82).[135]

The term 'multi-actor' suggests humans, but technology has reached the point where we need to recognise that some actors are non-human. Big Data analytics is a relatively new paradigm; dating back to about 2010. Big Data analytics can improve predictive ability and generate safety-related foresight in a number of ways, helping to detect emerging safety threats. Big Data may be a means to identify early warning signs that would be missed by human observers. The technology shows promise, but at the same time generates new risks, for example the opacity of algorithms for non-expert users. It is perhaps too early to reach conclusions about the contribution of Big Data analytics to safety foresight. That said, the development of autonomous vehicles is processing prodigious quantities of data to shape the algorithms necessary. This may well become the definitive case study of foresight in safety through Big Data. People working in safety need to keep an eye on developments in Big Data and machine learning.

Big Data holds the promise of extending safety foresight, but also of compromising it. The offer of powerful, objective prediction is a strong inducement to use the

[131] *Stephens, B. (2017). The dying art of disagreement. Keynote speech, 24 September 2017. The Lowy Institute. online: https://www.lowyinstitute.org/publications/dying-art-disagreement. Accessed, 4 June 2020.*

[132] *Espejo, R. (1989). A cybernetic method to study organisations. In: The viable system model: interpretations and applications of Stafford-Beer's VSM. Edited by Espejo. R., and Harnden. R., John Wiley & Sons, Chichester.*

[133] *https://www.citizensassembly.ie/en/previous-assemblies/citizens-assembly-2016-2018-/*

[134] *John Fitzgerald, quoted at https://www.basw.co.uk/resources/psw-magazine/psw-online/think-unthinkable Accessed 2nd June 2020.*

[135] *Eppler, M.J; Aeschimann, M. (2009). A systematic framework for risk visualization in risk management and communication. Risk Management, Vol. 11, Iss. 2, (Apr 2009): 67-89*

tools of Big Data. However, there is evidence[136] that without careful governance these tools can further entrench social inequality and bias. Furthermore, for all their power, these systems will not be omniscient. This, coupled to their opacity, creates a challenge to safety assurance. Therefore, embracing Big Data, like many new technologies before it, places high stakes on both sides of the balance.

Expertise is essential for foresight. Experts see warning signs in data, and foresee possibilities that are invisible to non-expert. However, how to qualify as an expert is an issue. Knowledge can be of different types, with some types being more often recognised as having the hallmark of expertise. For example, qualifications awarded by professional bodies and universities provide tangible evidence of expertise. In contrast, the extensive empirical knowledge of experienced individuals is less easily measured and may consequently be undervalued, or even discounted, as expertise. Furthermore, irrespective of their background, experts are unlikely to perform well in foresight tasks if they lack independence[137]. History is littered with examples of this kind of bias—scientific opinions about the link between tobacco smoking and cancer; and about the link between tetraethyl lead petrol additives and lead poisoning, to name just two. Foresight is a projection of expert knowledge, but expert knowledge is not an objective quantity.

As noted earlier, memory is a critical aspect of foresight. A significant example of this is the recall by decision-makers of the futures foreseen by experts in earlier life-cycle phases. Of particular significance is foresight by *designers*, which informs their assumptions and design choices. These are too easily not communicated to actors later in the life cycle. The B737 MAX case illustrates this point: pilots missing crucial knowledge about the behaviour of technical systems[138] that was well-understood by designers. Another point here is that the technical system in question was a radical departure from the expectations of pilots; an instance of disruptive rather than derivative design.

Experts are also needed to provide interpretive balance to safety metrics. There is a trend in many areas of safety towards monitoring through metrics. Measurement is to be applauded, but no matter how well-designed, the construction of metrics requires various assumptions and simplifications. Useful though these data may be, they cannot be the whole truth; and treating them as such will blind foresight to other, valid interpretations. We should be alert to spurious objectivity in safety as in any field, and experts can provide countervailing voices. This is especially important if a measure is used as a target[139] or key performance indicator.

This being said, non-experts are especially useful for providing 'out-of-the-box' ideas, and are able to ask questions which are less influenced by expertise and bureaucratic fragmentation and professional norms.

Within its multi-actor view of foresight in safety, the Project Group noted the importance of regulators. Who is a regulator and what is regulation, are both relevant questions? Regulators include statutory agencies: the enforcers of safety and environmental protection laws. However, *regulation* can also be seen more widely: as the operation of networked groups of stakeholders who support, or sanction, risk-owning entities in pursuit of safer products and activities (Braithwaite, 2017[140]). Foresight and regulation connect in many different ways: as a competence, as communication, and as an object for regulatory interventions.

Regulators can generate and disseminate foresight through their interactions with actors at different system levels. The privileged access of enforcement agencies allows them to inform their foresight, to communicate it widely in industry and to renew legislation when new knowledge is obtained (e.g. the precautionary principle). The Project Group noted that the value that regulatory agencies can bring to safety foresight depends on a number of factors. In particular: close cooperation between operating companies and regulatory inspectors, (ii) regular discussion and follow-ups of lessons learned, with a focus on near misses, (iii) regular discussions and follow-ups of possible scenarios (iv) specific skills and competencies for the inspectors and other regulatory personnel, and (v) follow-up of recommendations by regulators. In addition to these five points, a basic assumption is that regulators can properly engage with risk owner's models. This is not always possible. For example, a fundamental challenge to effective regulation, including self-regulation, is caused by the "black box" nature of many

---

[136] *An overview is provided by O'Neil, C. (2016) Weapons of Math Destruction. Pub. Crown.*

[137] *"It is difficult to get a man to understand something, when his salary depends on his not understanding it." Upton Sinclair (1994) "I, Candidate for Governor: And How I Got Licked". University of California Press.*

[138] *The Manoeuvring Characteristics Augmentation System (MCAS).*

[139] *Goodhart's law.*

[140] *Braithwaite, J. (2017) "Types of Responsiveness". In: Drahos, P. (Ed.). Regulatory Theory: Foundations and applications. Acton ACT, Australia: ANU Press. Retrieved June 12, 2020, from www.jstor.org/stable/j.ctt1q1crtm*

machine learning models. This inscrutability makes it difficult for risk owners to build and test mental models of system operation, for regulators to check the underlying assumptions and inner workings of the system, and for the legal system to inspect the logic underlying the model's predictions in case of an accident.

In summary, foresight in safety:

- *is most acute when several perspectives are shared in a community;*
- *is vulnerable to power imbalances between stakeholders;*
- *depends on ready willingness to review past decisions and commitments;*
- *is more efficient when stakeholders share operational knowledge;*
- *requires stakeholders to be able to disagree well;*
- *is more effective when dissenting voices are listened to—not necessarily agreed with—but taken into account and discussed;*
- *may be helped by Big Data and machine learning, but could be hindered by it;*
- *will vary between experts, even when all their views are valid;*
- *can be blinded by metrics, especially when the metrics are used as targets;*
- *is an example of the value that regulators can add to safety in cooperation with industry.*

Foresight: Memory and the Future

Foresight of future unwanted possibilities involves making associations between monitoring data, mechanisms of failure, and preventative and mitigating actions. This knowledge is partly discovered by experience, but also created by imagining, modelling and theorising. For example, causal models can be created using incident scenarios. This allows the systems modelled to be modified, detection set-up, and interventions planned.

Well-investigated accidents tell us that loss of memory is a recurring root cause of disasters, and a vulnerability in many organisations. What needs to be remembered are monitoring data, safety performance indicators, mechanisms of failure, preventative and mitigating actions, and the causal models in which all these elements cohere.

Organisational memory is likely to be vested in many different substrates, both human and non-human. Substrates include: the memories of the individuals who populate the organisation; the technology into which designers' have encoded their foresight, and documentation of various sorts, in particular on processes. There is almost always scope to improve the reliability and capacity of these substrates for the sake of safe operations.

When trying to avoid forgetting, it is tempting to equate memory with storage. We know a lot about storage and might prefer to put our effort into the things we understand best. However, all the storage in the world cannot deliver faultless memory or perfect foresight. Human memory is these days seen as a process rather than a store of facts. Similarly, foresight for safety assumes that organisational memory is a continuous process that integrates all the different substrates within the organisation. Therefore, as well as storage in databases, documents, people, and artefacts, we must attend to the whole process for the sake of foresight.

Early Warning Signs have been a recurring concept in the Project Group's deliberations. Foresight entails identifying the events and conditions that signal the increasing likelihood of an unwanted event. Before a thunderstorm, the gathering of black clouds and distant thunder are early warning signs. Seen this way, foresight links current knowledge to future possibilities. In the field of safety, early warning signs are crucial links in this chain.

Knowledge of early warning signs and associated actions are what needs to be remembered in the organisation. Memory needs to **store** this knowledge reliably. Moreover, to be remembered, knowledge must also be encoded in the first instance, and **retrieved** at point of need. Encoding, storage and retrieval of this information can be made the subject of assurance. Knowing how to use this knowledge in different contexts and situations is a competence that is not trivial.

Loss of memory is a critical failing in foresight. It means that early warning signs will go unheeded; we wait under the darkening sky and are surprised when lightning strikes. Theoretical models, such as the encoding-storage-retrieval model just mentioned, can inform ways to prevent this kind of forgetfulness. Moreover, to preserve its memory, industrial practice must recognise the effects of organisational 'macro' phenomena such as plant ageing and outsourcing. In addition, accidents themselves have value as stories. Stories are a means to revive memories of early warning signals and to remind about the seriousness of what

they presage. Rules alone seldom communicate the experiences that gave rise to them.

In summary, foresight in safety

- *is a process closely related to memory;*
- *depends on memory in general, and of early warning signs in particular;*
- *is precarious, because organisational memory does not look after itself.*

Foresight and Risk Assessment

In industries with complex operations, foresight has become almost synonymous with analytical risk assessment. However, foresight is also deeply implicated in the monitoring and review process that exist in parallel to risk assessment. The analytical approach to risk, developed in aerospace in the 1960s, was quickly adopted in the US military industrial complex, and spread globally and sectorally to most industries by the 1990s. Within that tradition, risk assessment is usually seen as comprising several sub-processes, including risk identification, risk analysis and risk evaluation.

Foresight is closely associated with the risk identification step of risk assessment, although how closely depends on how one defines these terms. To some extent, risk identification is actuarial; it is informed by past failures and successes. However, risk identification—the foreseeing of possible futures—is also creative and relies heavily on the knowledge of the people involved, their imagination and the models they create. For that reason, risk identification is sometimes singled-out as the least reliable part of the risk assessment process. *Least reliable* does not, however, equate to *bad*; it simply means that, all other things being equal, no two analyses of the same system will produce exactly the same risk model. This implies that there is room for discussion, and for humility, in even the most robust and meticulous risk analysis.

To better capture risks, risk analysis approaches are needed that are more open to different worldviews and opinions. However, the qualitative roots of a risk analysis can sometimes be obscured by the complexity of quantitative evaluations made. For example, Vesely et al. point out in their classic handbook[141] that a Fault Tree Analysis is *"a qualitative model that can be evaluated quantitatively"*. Quantification is often necessary, but it may create an impediment to the qualitative discussion and review that we have argued is essential to foresight in safety. We note the ISO standard on risk management[142] emphasises that communication and consultation are intimately connected to the risk assessment process. How to make this communication work between technical people and lay people is one of the questions that workers in the field of safety continue to grapple with. Another is how to make opaque risk models discussible, a point made earlier in respect of Big Data and visualisation.

In summary, foresight in safety:

- *is greatly informed through risk assessment, but not synonymous with it;*
- *involves combining qualitative and quantitative knowledge—a challenge for communication and debate in the multi-actor arena;*
- *has to be approached with humility, as despite all efforts there is always room for discussion and improvement.*

Foresight in Safety: A Wider Perspective

Most of this chapter has been in the context of an organisation or within a sector. However, foresight with a wider perspective is necessary to avoid the shocks and embrace the opportunities that originate from beyond those boundaries. The hallmarks of an international approach are, according to the Project Group, a global warning system (of early warning signs), a rapid and trustworthy information system, global decision-making procedures, necessary reserve capacity, and international cooperation to avoid global inequality in disaster management. The last COVID19 crisis can provide examples of successes and failures in that respect.

We mentioned earlier how the rate of socio-technological change was recognised in the 1970s. Within safety, this challenge to foresight has driven innovation in

---

[141] *Vesely, W.E., Goldberg, F.F., Roberts, N.H., and Haasl, D.F. (1981) Fault Tree Handbook. NUREG-0492, US Nuclear Regulatory Commission. [Online:* https://www.nrc.gov/docs/ML1007/ML100780465.pdf, *accessed 12 June 2020]*

[142] *British Standards Institute, 2018. BS ISO 31000:2018. Risk management — Guidelines. Geneva, Switzerland: International Organization for Standardization (ISO).*

modelling and risk analysis. However, the changing safety landscape extends beyond these. To keep pace with rapidly evolving technological advancement, globalisation and demographics, diversity of worldviews and stakes, safety foresight requires a greater embrace of governance principles.

Risk governance at all levels (in the sense of Rasmussen and Svedungs' model) is significantly challenging. Foresight in safety is a subset of foresight in general. According to an online EU foresight guide,[143] foresight is defined as "*a systematic, participatory, future-intelligence-gathering and medium-to-long-term vision-building process aimed at enabling present-day decisions and mobilizing joint actions. It can be envisaged as a triangle combining "Thinking the Future", "Debating the Future" and "Shaping the Future". Foresight is neither prophecy nor prediction. It does not aim to predict the future – to unveil it as if it were predetermined – but to help us build it. It invites us to consider the future*".

The Project Group sees a need to incorporate foresight thinking into the classical risk management approach. The aim of the change is to bring about a more integrated way of thinking, debating and shaping the future. Part of this change would be for stakeholders to consciously incorporate *megatrends* when designing processes and making decisions. Megatrends are "*large, social, economic, political, environmental or technological changes that are slow to form. Once in place, megatrends influence a wide range of activities, processes and perceptions, both in government and in society, possibly for decades*"[144].They are the underlying forces that drive trends that are observable now and will most likely have significant influence on the future[145].

The megatrends viewpoint allows foresight of the dynamic, unfolding nature of large, complex systems. At this scale, an iterative approach appears to be critical to foresight in safety. However, the field of safety has yet to rise to the methodological and sociotechnical challenges inherent in an iterative approach. This is starkly illustrated in major accidents, but also in everyday examples of the inflexible bureaucratic approach that characterises much of safety practice. The present authors endorse the value to foresight of managing details—such as by

sophisticated record-keeping and cost-control—but note that these practices do not really acknowledge that complexity creates its own patterns. Therefore, the Project Group recognises a need to develop know-how and supporting tools to address the dynamically complex and evolving safety landscape with foresight thinking at all governance levels. In rising to this challenge, it goes without saying, perhaps, that advantage should be taken of new technologies to complement conventional approaches.

The history of major accidents leads us to believe that vigilance for anomalies is critical to foresight. Once an anomaly is recognised as an early warning sign, and the connection made to future possibilities, there is usually time to act. Latent flaws can be uncovered and fixed. (And, on a good day, we'll also ask "if this was wrong, what else should we be looking for"?). This kind of vigilance has many enemies, among them, production pressure, a changing workforce, plant ageing and inadequate monitoring[146]. However, we also note Turner's point[147] that risk management is based on beliefs, not certain knowledge. Overestimating the reliability of knowledge can cause us to overestimate the reliability and safety of the systems we create. In foresight, a modicum of doubt and humility should always be welcome and, when decisions are taken under pressure, a modicum of forgiveness in hindsight. This mindset is far from easy to sustain and its added value can only be appreciated from time to time and in the long run.

Change management, education and learning offer opportunities to improve foresight in safety. This is in contrast to safety regimes based on compliance, control, deregulation and privatisation. The challenges are: to integrate change management in a broad, multidisciplinary management model; to stimulate the development of competence, flexibility, insight and responsibility instead of conventional education; and a culture of dynamic learning, instead of static, rule-based training among all actors with safety responsibility, including safety professionals.

---

[143] *http://www.foresight-platform.eu/community/forlearn/*
[144] *http://ssl.csg.org/Trends/Megatrends%20Definitions%20and%20Categories.pdf*
[145] *https://ec.europa.eu/knowledge4policy/foresight_en*
[146] *For example, see 'normalisation of deviance' as described by Vaughan, D. (1996). The Challenger Launch Decision. Risky Technology, Culture, and Deviance at NASA, The University of Chicago Press.*

[147] *Turner, B. (1978) Man-Made Disasters. Wykeham Publications.*

In summary. Foresight in safety needs:

- *approaches designed to cope with radical and discontinuous change;*
- *to find ways that recognise complexity in systems;*
- *a global approach to collating and sharing data and knowledge;*
- *to embrace governance principles;*
- *to include wider megatrends in its imagination of future possibilities;*
- *to be unceasingly vigilant.*

In closing, there is still a lot to learn within the *foresight in safety* landscape. This text is the continuation of a journey that started a few decades ago because of concerns about quality of accident databases, of accident investigations, of learning from accidents and foresight in safety. It is the Group's mission to get acquainted with this complex and evolving landscape. Against this backdrop, members of this Project Group will take stock of what it has learned and start a new ESReDA Project Group on *Risks, Knowledge and Management*. This will continue to look at activities and disciplines related to risk assessment, identification of early warning signs and emerging risks, foresight, investigation of events and lessons learning, management of barriers and lines of defence, reliability, and change of policies and culture; however, it will focus on knowledge management aspects of these. And the learning odyssey continues...

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub – Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office
of the European Union