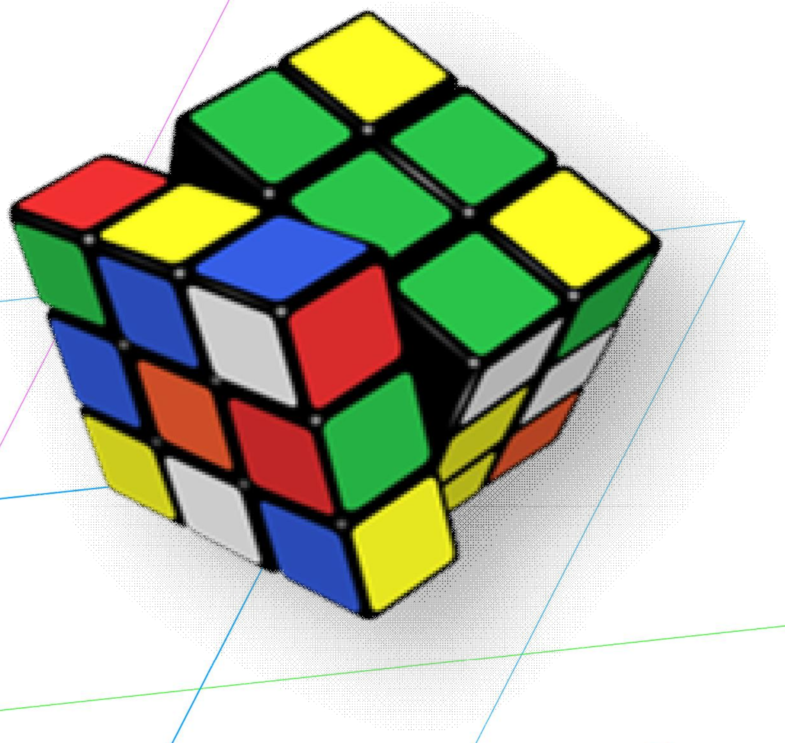
Kindunos Safety Consultancy Ltd, Gorinchem, the Netherlands

Challenges to the Investigation of Occurrences

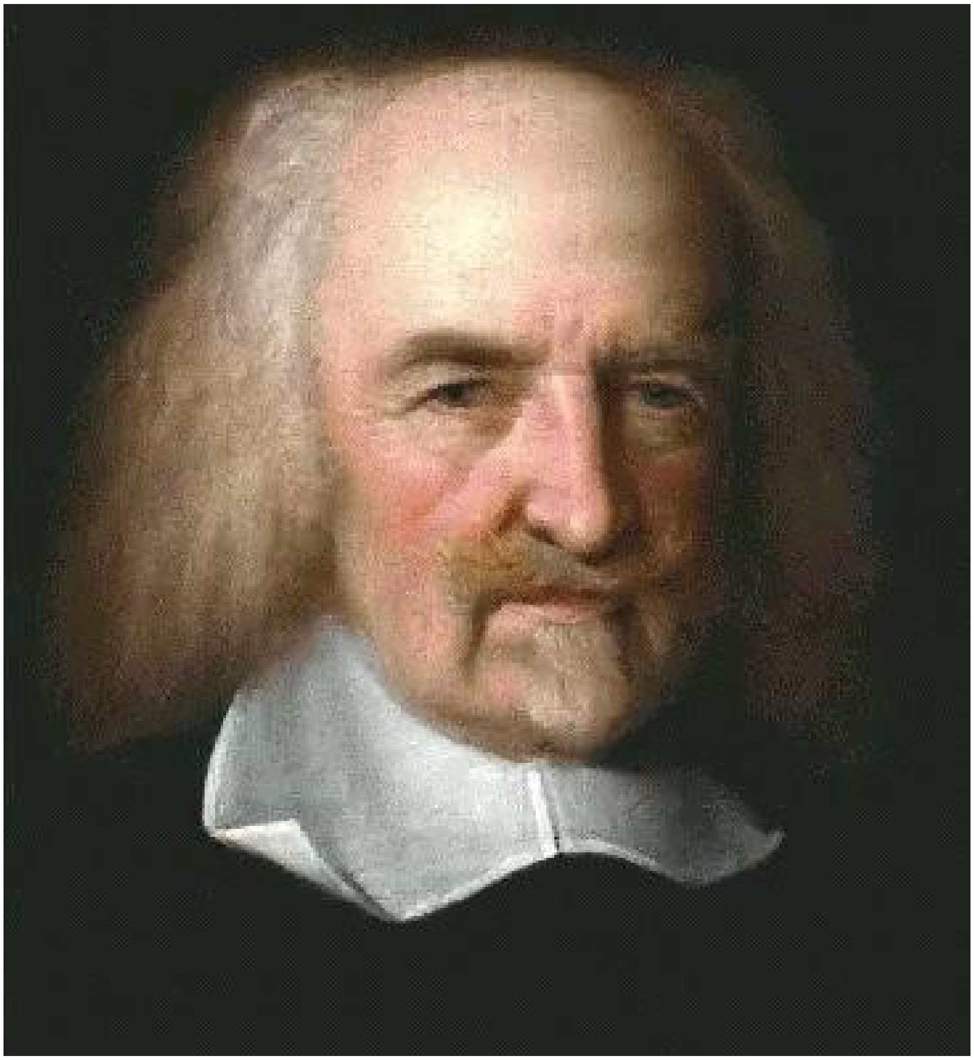
Concepts and Confusion, Metaphors, Models and Methods

John Stoop

Side Document of the
ESReDA Project Group
Dynamic Learning from Accident Investigation



ESReDA
European Safety, Reliability & Data Association



Ignorance of remote causes, disposeth men to attribute all events,
to the causes immediate, and instrumental:
For these are all the causes they perceive.

Thomas Hobbes. Leviathan, Chapter XI
1588-1679

Contents

Correspondence address:
Stoop@kindunos.nl
Gorinchem, the Netherlands
January 2015

During his presentation on precaution and accident investigation on a Conference on Ethical Aspects of Risk at Delft University of Technology (14-16 June 2006), an aerospace engineer and safety investigator overheard two social scientists on ethics in the audience saying:

"There is still one out there who has the nerve to defend himself"

Project Group Dynamic Learning from accident investigation

The ESReDA Project Group Dynamic Learning from Accident Investigation stands in a tradition of consecutive projects exploring several aspects of accident investigation. With the organization of the 45th ESReDA Seminar 'Bridging the Gap between Safety Recommendations and Learning' (2013) held in Porto (Portugal), which built on the 24th ESReDA Seminar (2003) held at the JRC-Institute for Energy in Petten (the Netherlands); on ‘Safety investigation of accidents’; the 33rd ESReDA Seminar on ‘Future challenges of accident investigations’ at the JRC-Institute for the Protection and Security of the Citizen in Ispra (Italy) and on the 36th ESReDA seminar on ‘Lessons learned from accident investigations’ at EDP in Coimbra (Portugal), it has prepared three deliverables which have been printed and published by ESReDA:

- Accident Investigation Practices - Results from a European Study (2003 - report);
- Shaping Public Safety Investigations of Accidents in Europe (2005 - ESReDA Safety series);
- Guidelines for safety investigation of accidents (2008) available for free download on the ESReDA website.

The main objective of the project group has been to work out recommendations on how to capture, document, disseminate and implement insights, recommendations and experiences obtained in investigations of high-risk events (accident and near-misses, and safety as well as security) to relevant stakeholders via:

- Proposing adaptation of investigation methods to specific features of each sector and aimed at facilitating more impact;
- Identifying barriers within companies, public authorities and other involved stakeholders that may hamper implementation of recommended preventive measures;
- Providing methods for dynamic learning from accidents;
- Highlighting good practices on how to develop recommendations from accident investigation findings and understanding relevant preconditions for future learning (resilience, learning culture);
- Giving advices and suggestions regarding Operational Feedback Systems for relevant decision makers.

Members of the project group are:

- Mr. Nicolas Dechy, Engineer In Organisational And Human Factors IRSN, France
- Mr. Yves Dien, Expert Researcher, Electricité De France, EDF R&D, France
- Mrs. Linda Drupsteen, Researcher , TNO Safe And Healthy Business, The Netherlands
- Mr. António Felício, Engineer In Generation Management EDP, Portugal
- Mr. Carlos Cunha, Optimization and Flexibility Management, EDP, Portugal
- Mr. Sverre Røed-Larsen, Project Manager, SRL HSE Consulting, Norway
- Mr. Eric Marsden, Department Recherche, Foncsi, France
- Mrs. Tuuli Tulonen, Senior Researcher, D.Sc.(Tech.), Tukes, FINLAND

- Mr. John Stoop, Managing Director Kindunos Safety Consultancy Ltd, The Netherlands
- Mr. Miodrag Stručić, European Commission, Joint Research Centre, Institute For Energy and Transport The Netherlands
- Mrs. Ana Lisa Vetere Arellano, Scientific Officer, European Commission, Joint Research Centre, Institute For The Protection and Security Of The Citizen Security Technology Assessment Unit, Italy
- Mr. Johan K. J. van der Vorm, Senior Technical Consultant TNO Safe And Healthy Business, The Netherlands.
- Mr. Ludwig Benner, Senior Consultant and corresponding member, United States of America. Mr Benner was awarded an Honorary Membership of the ESReDA Working Group for his dedication to the development of safety investigation methodologies.

ESReDA

ESReDA, The European Safety, Reliability and Data Association, is an European association that provides a forum for the exchange of information, data and current research in Safety and Reliability. The safety and reliability of processes and products are topics which are the focus of increasing European wide interest. Safety and reliability engineering is viewed as being an important component in the design of a system. However the discipline and its tools and methods are still evolving and expertise and knowledge are dispersed throughout Europe. There is a need to pool the resources and knowledge within Europe and ESReDA provides the means to achieve this.

Contact:

- ESReDA General Secretary, Mohammad Raza, ALSTOM Power,
  7, Brown Boveri Strasse, 5401, Baden, Switzerland.
  Mohammad.Raza@power.alstom.com
  http://www.esreda.org/

Preface

After publishing the book on *Shaping public safety investigations of accidents in Europe* in 2005 and the *Guidelines for safety investigations of accidents* in 2009, ESReDA now intends to present its achievements on learning from safety investigations on the ESReDA Seminar in Porto, Portugal, Autumn 2013.
Simultaneously, the Resilience Engineering Association organized its 5th conference in the Summer of 2013 in Soesterberg, the Netherlands, focusing on managing trade-offs. The conference has addressed current challenges organizations face, why resilience is needed to address those challenges, and how Resilience Engineering can be put into practice.

During the process of preparing the products of the working group and the conference, discussions have been exploring notions of resilience engineering, organizational learning, system dynamics, complexity, case based analysis and investigation models. Such discussions have been guided by the intention to provide a practical approach to safety investigations that could be applied across various domains and industrial sectors on minor as well as major events. It became clear that several schools of thought exist, related to specific technological and industrial domains and scientific disciplines, creating confusion and differences in interpretation, creating controversies between theoretical approaches and best practices.

Clarifying such confusion is a time and resource consuming activity that is well beyond the mission of the Working Group and its intended deliverables. The mission of the Working group focuses on three deliverables in the form of documents that should provide a practical guide to dynamic learning from accident investigations: Barriers for Learning, Guidelines for Learning and the Training Toolkit. The Working group decided to add a Side Document in the form of an essay on concepts, confusion, metaphors, models and methods, reflecting the more theoretical elements and the broader perspective in the work done over the past 4 years.

According to the purpose of identifying barriers for learning, this essay is divided into two parts:
Part 1 explores potential barriers for learning and elaborates in chapter 2-5 on several basic notions, scientific principles and theories that – mostly implicit- lay behind the conduct of safety investigations to clarify and resolve ambiguities that presently exists. It reflects a current state of the art in the debate on investigation methodology. Suggestions and notions by researchers from various scientific disciplines and application domains are assessed regarding their alternatives for safety investigations and safety and risk management approaches. Such alternatives challenge current retrospective investigation approaches by advocating new approaches with a paradigmatic shift that should make investigations obsolete such as:

- Design of socio-technical systems: operations management and safety management
- Sophistication of modeling: Bayesian belief networks and big data systems.
- Introduction of sociological issues: Safety Culture, Leadership and Governance.

This essay aims at describing and identifying trends and patterns in current debates and operational practices that might require deeper understanding in order to understand dynamics and complexity of systems. The essay explores notions, theories and concepts that might provide additional explanatory potential in providing structure and transparency in dealing with the safety implications of knowledge management, engineering design, system change and system state transition management.

Part 2 explores prospect options to overcome barriers for learning and focuses in chapter 6 on enhancing the diagnostic capability of safety investigations itself, expressed in a new framework for dealing with complexity and dynamics in safety investigation methodology. This framework is developed by the author, inspired by the Working Group discussions and emphasizes the paradigm and unique potential of single event investigations.

Part 2 addresses several critical notions for academics and practitioners in the investigation community:

- there are several dilemmas and pitfalls in dealing with current investigation notions and perspectives. Normative constructs are applied without being made explicit or clarified as valid and applicable to the investigation process: the phenomenon of complexity and unavoidability of emergent system behavior as unknown-unknown properties in operations, the value of a pro-active over a reactive approach as a feasible replacement of feedback learning, rejection of a technical perspective in favor of social and organizational perspectives. It is the author's observation that some social scientists might mistakenly be barking at the Newtonian design and engineering tree instead of the Tayloristic tree of scientific management and organizational failure. A more clear and explicit distinction between investigation goals of finding the truth and establishing trust in system performance versus allocation of blame and liability at a methodological level seems to be lacking.

- various scientific arenas debate the validity and theoretical correctness of notions and concepts across disciplines, the use of metaphors, methods and models. Such debates deal with the abolition of event analysis in favor of modeling and substantiation of systems behavior without the burden of proof of collecting evidence on a case basis. Instead, a refocusing is suggested from accident investigation towards optimization of primary production and control processes. Coexistence between various school of thinking is advocated by senior scientists from almost every discipline. Such a coexistence could favor a theoretical and methodological basis for the emergence of an investigation science, which by definition represents the ability to conduct investigations based on a familiarity with a broad range of disciplines and the ability to pursue several lines of investigation simultaneously. To this purpose, two relative unknown forms of logic reasoning are clarified: abduction and construction.

- the emergence of IT and software engineering made new scientific notions and technological expertise available for application to the notion of safety. Instead of the conventional mechanical notions regarding 'failure', 'load concept' and 'kinetic energy', notions of 'complexity' and 'emergent properties' appear. Such notions shift a focus from structure and architecture of systems to process control and performance. Such a shift can be traced back to inherent problems with software reliability, rate of automation of operator's tasks, certification and testability of embedded IT software and system designs across all phases of the decision support process as has been demonstrated by the Y2K bug, viruses, hacking, cybercrime and a series of automation issues in the aviation sector.

- the use of communication metaphors as analytic and solution generating tools in applying constraints on operator behavior from a managerial perspective, while such metaphors are solely based on a classic concept of energy transfer which is proven invalid in domains of cognition, decision making and governance. In order to communicate manipulating solutions for modern software interactions and decision making algorithms, the metaphor of the Rubik Cube is introduced.

- new developments and concepts, such as resilience engineering and forensic engineering, could be mutually supportive for enhancing safety and managing complex and dynamic systems. In addition to the existing technological diagnostic potential, an organizational and social diagnostic framework and toolkit will be welcomed in the community of safety investigations. Such a cooperation across disciplinary boundaries is in its early phase of development.

- how to convince change agents and game changers of the usefulness of integrating safety in their activities? Designers/manufacturers and operators/managers have specific potential to enhance safety. Investigators should speak their language, know their logic reasoning process, assumptions and tradeoffs in order to communicate their findings and recommendations. Linear and static analytic safety approaches do not longer match with their tools and techniques of concurrent engineering design principles, scientific optimization methods and computerized modeling and simulation algorithms. This essay therefore introduces notions of state-space system modeling and vectorial representations of such models.

- Finally, cooperation with operational experts and incorporating their experiences in the investigation process is indispensable in identification of systemic and knowledge deficiencies. All parties and perspectives should be incorporated in the investigation.

Involving operational expertise and experience is a requisite for change and learning:
In an article on High-Altitude Upset Recovery in Aviation Week, captain C. B. "Sully" Sullenberger described his ditching in the Hudson as a seminal accident. "*We need to look at it from a systems approach, a human/technology system that has to work together. This involves aircraft design and certification, training and human factors. If you look at the human factors alone, then you're missing half or two-thirds of the total system failure...*"
He also emphasizes the importance of the availability of primary flight parameters: "*accurate airspeed indications alone aren't the best data the crew needs to recover from an upset. That requires knowing the wing's angle of attack (AoA). We have to infer angle of attack indirectly by referencing speed. That makes stall recognition and recovery that much more difficult. For more than half a century, we've had the capability to display AoA (in the cockpits of most jet transports), one of the most critical parameters, yet we choose not to do it."*

Involving scientific modeling and human behavior theories:
On one hand, the chief investigator of BEA, concludes in "The final word: Air France flight 447" (Troadec 2013) that: *the combined use of ergonomics of warning design, training conditions,*

recurrence training process did not generate expected behaviour and showed the limits of current safety models.

On the other hand, the vigilance, proficiency and flexibility of qualified air crews proved of paramount importance in the ability to recover from unforeseen and unprecedented events. In case of the Qantas Flight QF32 the Airbus A380 loss of containment event was brought to a successful closure by deviating from procedural flight performance and rule based responses in a potentially unrecoverable failure (ATSB 2010). We should also learn from successful recovery from potentially disastrous events.

Involving certification and inspection authorities:

In the AA Flight 587 case, it took investigators almost 2 years to identify a flaw in the hydraulic system that had been certified as failsafe in the 1960's. Since the FAA and aviation authorities cannot match the resources of companies such as Boeing and Airbus, manufacturers engineers signed off most of the elements of the Dreamliner battery pack. Leaving a final approval to the governmental agencies, subtle conflicts of interest could influence the assumptions of manufacturers engineers in earlier phases of certification.

As stated by Tom Haueter, the NTSB chief investigator on AA flight 587: *"It's the assumptions that kill you, and if things do not work out the way you planned, things can go very bad, very fast".*

1. Introduction

Major transitions in socio-technical system developments have been argued, based on internal and external conditions. Internal factors are focusing on performance pressure within a system in order to control required growth, modal shift demands, intelligent operation and expansion of transport services. External factors should be integrated in a future systems development, dealing with land use, detrimental environmental effects, sustainability and safety. Such transitions have changed the operating environment of safety investigation theory and practices.

Technological innovation
Major issues in several modes of transportation have lead to such a system pressure that major changes should be introduced. A 'system leap' forwards may be inevitable. Rather than applying proven technology and pragmatic improvements on a detailing level, technological innovation may be necessary to overcome constraints in system development.

Conceptual change
A 'system leap' approach introducing conceptual changes of a non-technological nature in the architecture of complex systems dealing with business modeling, market development and globalization. External pressure exists with respect to spatial planning, land use and urban development coping with congestion, external safety and environmental constraints in densely populated areas and compact city concepts.

Integrating safety
Historically, safety has been submitted to a fragmented approach, scattered across industrial sectors, policy domains and scientific disciplines. To fulfill internal and external demands, a 'conceptual leap' in safety notions may be required. It may become necessary to transform safety from an operational cost into a strategic issue, integrated in each phase during the life cycle of such systems. The question is how safety should be integrated in these developments to achieve a pro-active and sustainable safe operation throughout their life cycles.

Reading advice
In this publication, some background information is provided on the use of the ESReDA safety investigation method. After this short introduction, chapter 2 provides a scoping and embedding of investigation in the various school of safety thinking. Such scoping indicates the success rate for learning potential by the feedback loops that are available.
Chapter 3 introduces the scope of intervention potential, dealing either with systems optimization within the operating envelope or systems adaptation by design.
Chapter 4 elaborates on a series of notions, such as complexity, uncertainty and problem orientation and knowledge management.
Chapter 5 discusses the claims for a paradigm shift and the transition from model to method. Selecting STAMP and FRAM as the most prominent representatives of these developments
Chapter 6 elaborates on the basics of the ESReDA method, comprising forensic sciences, time as a diagnostic dimension, the separation between recomposing the event and modeling the system and the notion of the ESReDA Cube as the link between analysis of events and changing systems.

Part 1

About concepts, confusion, metaphors, models and methods
Exploring the diversity of present thinking

2. History of investigations

Accident investigation as an analytic tool foremost has its origin in all modes of the transportation industry and found a rapid expansion to process industry and the energy sector in the 1960's. Gradually, investigation principles have been applied to sectors with a less prominent technological signature, such as medicine, firefighting, rescue and emergency. This long and dedicated history has caused a wide variety and divergence of notions, fundamentals, methods and perspectives.

Therefore, generalizing investigative notions have to be put in its historical perspective, tracing back their origins and motives, their ability to enhance safety by embracing a socio-technical systems perspective and the specific role of investigations in this approach (Lees 1960, Edwards 1972, Hendrickx 1991, Benner 1996, Rimson and Benner 1996, ETSC 2001, Leveson 2002, ESReDA 2005, Young et.al. 2005, Katsakiori 2008, Benner 2009, ESReDA 2009).

Over the decades, four safety Schools of Thought in high technology complex systems have emerged (McIntosh 2012). At the same time, new technological developments in high technology sectors such as aviation have forced the investigation community to reflect on the next century of investigation theory and practices, which create major challenges for the profession (Hersman 2012, McIntosh 2012).

At the same time, challenges are to be met in a competitive and open market on a global scale (EU 2011):

In 2050, the European air transport system is integrated in a complete logistical transport chain and part of a fully interconnected, global aviation system that is based on a multilateral regime rather than on a series of bilateral agreements. Interoperability between Europe and the other regional components of the global network is complete. Commercial air transport services are provided mainly by airlines organized as a few global alliances. Thanks to tight links between technological and regulatory approach, Europe has a global lead in the implementation of international standards covering all aviation issues, including interoperability, the environment, energy, security and safety. This leadership ensures that the global regulatory system enables market access and free, fair and open competition.

Within Europe the number of commercial flights is up to an expected 25 million in 2050 compared to 9.4 million in 2011. At the same time, these plans put high demands on maintaining the present safety level in the aviation sector (EU 2011):

Overall, the European aviation transport system has less than one accident per ten million commercial aircraft flights. For specific operations, such as search and rescue, the aim is to reduce the number of accidents by 80% compared to 2000 taking into account increasing traffic.

2.1 four schools of thought

Safety in modern transportation systems has been an issue for about 150 years. It evolved as a discipline from several different domains and disciplines and has a strong practical bias.

Consequently, various 'schools of thought' have been merging, of which the most important can be categorized as 'Tort Law School', 'Reliability Engineering School' and 'System Safety Engineering

School' (McIntyre 2000). In addition a fourth school is defined as 'System Deficiency and Change' (Stoop 2002).

Each of these schools represent a different pattern of thinking and can be considered as consecutive, representing the societal and scientific safety concepts of their times.

In general, four principal safety engineering design concepts can be derived from these schools of thought:

- *deterministic engineering design*. This concept is essentially reactive in its learning potential and focuses on failure modes, identification of failure causes and accident prevention strategies by developing technical design options. Failure modes are established from post-event investigations with a technical-analytical emphasis on the failure of hardware components and the acceptability of mechanical loads and margins (Carper 1989, Petroski 1992).
- *Probabilistic engineering design.* This engineering design school primarily focuses on the mathematical probability of failure and reliability of the system components performance during the system life cycle. This probabilistic concept was developed in hydraulic engineering, process industry and nuclear power supply. This school applies a wide diversity of techniques such as Reliability, Availability, Maintenance and Safety (RAMS), Probabilistic Risk Assessment (PRA), Failure Mode and Effect Analysis (FMEA), human engineering and High Reliability Organizations (HRO).
- *Systems engineering design*. This engineering design school emerged from aerospace and defense applications. Unpredictable interactions characterize complex systems and modern technology. This engineering design concept expands a strict technical intervention towards incorporating environmental, organizational, social and societal conditions in a socio-technical systems design concept (Rasmussen and Svedung 2000, Amalberti 2001, Hollnagel and Woods 2006). Methods and tools of this concept focus on modeling systems and dynamic interactions with their environment. An interest in disaster and emergency management occurs as a consequence to focus on the safety performance of the overall system during all phases of its life cycle.
- *Safety deficiency identification and system change.* This engineering design concept is dealing with the participative nature of major projects and systems change, taking into account safety requirements and interests from various groups of stakeholders during normal and deviant operation of transport systems. This concept transforms the closed nature of the engineering design process into a participative, open process in which the pressure to innovate technologically, organizationally as well as methodologically is clearly present. This concept establishes a procedural relation between the design phase and operational phase of socio-technical systems and identifies a specific role for single event investigation of major events. In order to provide feedback from disastrous events to the systems operational performance, a timely transparency, credibility of findings, public confidence is required (IDAIP 2001, Van Vollenhoven 2002).

2.2 elaborating on systems safety frameworks

Historically, safety investigations have seen a development in which a focus on technological failure has been complemented by a focus on human behavior, organizational failure and governance risk decision making (Fahlbruch and Wilpert 2002). Over times, a more integral safety notion has been created by broadening the scope from pre-accident causation contributing factors towards safety enhancing and recovery factors during rescue and emergency handling in the aftermath of serious events. Such developments have been triggered by serious events in the Channel Tunnel and tunnels in the European Alpine region (Eisner and Stoop 1992, Eurotunnel 1994, Stoop and Beukenkamp 2003). In aviation, the issue of external risk came under scrutiny after the ElAl crash near Amsterdam in 1992 (Stoop 1997). In the international aviation community, a gradual development took place in incorporating family assistance and victim support in dealing with the aftermath of aviation accident, first in the USA and Canada, later on in Europe (Troadec 2013). Informing the public about major accidents and incidents through independent investigations has become a Citizen's Right and Society's Duty (Van Vollenhoven 2002, Van Vollenhoven 2006). Merging these safety aspects into a notion of 'integral safety' has broadened the identification of system deficiencies, which may have their origin before, during and after an occurrence. Consequently, the population at risk involved in the investigation has expanded from crew and passengers to employees at airports, residents in the vicinity of an airport, rescue and emergency services, family and relatives that may encounter traumatic damage due to an air crash.

Such a broadening of the safety scope however, is not enough to assess the future potential societal impact and subsequent risk acceptance of air accidents and incidents. According to Flight Path 2050, air traffic volumes will increase by about a factor of 2.5, while maintaining the same safety level. In this concept, increasing the absolute number of aviation accidents is simply socially unacceptable. Such challenges put a stop to linear thinking in current definitions of risk and safety. Where initially risk (R) is defined as the product of probability (p) and consequence (c), such a definition restricts itself to a single event level. Such a definition lacks a systemic context and higher order, non-linear relations between probability, consequence and the population at risk. While R=p*c expresses the risk of a single event, it is necessary to multiply this individual risk by the exposure (E) to risk, to achieve the social impact (I) of all occurrences on society.
In short:

Risk (R)=probability (p) X consequences (c), while
Impact (I) = probability (p) X consequences (c) X exposure (E).

This exposure (E) is expressed in the overall the number of flight movements and will increase with a factor of about 2.5. According to Flight Path 2050, the social impact I= p*c*E has to be maintained at a constant level of safety performance. Since the aviation sector already has achieved the goal of one accident per 10 million flights and aircraft survivability rates are dependent on the size of aircraft, the risk R cannot be reduced with a factor of 2.5 without dramatic interventions. Simultaneously, a goal of 80% reduction of aircraft accidents has to be achieved. This non-linearity in risk acceptance cannot be achieved by interventions in the performance of the system. Changes in properties and processes are required. Eventually, innovation becomes inevitable. Maintaining the established safety levels will require an expansion of the systemic context, involving design, manufacturing, certification, operations and operating environment in safety investigations and analysis.

Safety investigations will have to take into account such non-linearity and systemic contexts. Inevitably, safety interventions will have to deal with redesign and update of the system, modification of conditions and constraints in the operating envelope and organizational or institutional conditions. The main focus shifts from product deficiencies to 'system deficiencies', replacing the conventional focus on deviation, liability or blame. Focusing on a multi-aspect approach, a variety of safety performance indicators becomes available, covering additional aspects beyond working conditions, internal and external safety such as rescue and emergency, public risk perception, incident handling or public order and security.

2.3 a role for accident investigations

Due to a series of major accidents and disasters, the focus of attention in public safety perception has shifted from complying with quantitative risk standards towards independent accident investigation of major events (Stoop 2012). At a European level, mandatory investigation agencies are recognized as indispensable safety instruments for all modes of transportation, for which various EU Directives have been developed (ETSC 2001). After the tunnel fires and explosion in Toulouse, several directives have been established, dealing with specific hazards in such situations.

Characteristics and properties of modern, open, complex systems can be identified and analyzed along the lines of:

- a preventive analysis of the primary processes and relevant actors during design and operation including their safety critical strategic decision making issues. However, such a preventive encompassing analysis is not always feasible in practice due to the complexity and dynamic nature of transportation systems and their permanent and incremental adaptations.

Therefore, a second reactive approach is indispensable:

- an in-depth and independent investigation into systemic incidents, accidents and disasters. Such independent investigations may provide a temporary and timely transparency as a starting point for removing inherent deficiencies in such systems before they manifest themselves as 'emergent' properties.

Accidents and incidents are the manifestations of such inherent and emergent properties and may provide evidence and explanation through investigations.

The importance of early detection of systemic deficiencies is pivotal in contributing to the concept of First Time Right and Zero Defects. Recent studies have indicated that failures to spot and anticipate safety flaws during certification of new aircraft have been linked to 70 % of US airline-crash death in the past 70 years (FSI 2013). The Dreamliner lithium-ion battery fires have renewed questions whether complexity and new technologies of new aircraft have outpaced a manufacturers' and regulators' ability to identify deficiencies during design and certification. Although certification standards have preventing fatal US airline crashes since 2001, occasions have occurred where assumptions were incorrect and not conservative enough. The use of 'special conditions' in the absence of regulations for new technologies, as applied in the Dreamliner case, have proven the need to modernize certification processes and standards (FSI 2013).

The history of the most deadly American airline accidents since 1993 is dominated by cases where manufacturers and regulators did not foresee how aircraft might fail. USAir flight 427, TWA 800 and American Airlines 587 killed 627 people, while over the past 20 years out of 1123 death, 5 of such accidents counted for 783 of these casualties (FSI 2013).

There is a specific role for accident investigations as a partner in a more institutionalized network as a prerequisite for a further professional development, sharing professional expertise and participating in knowledge management.

In such a network, safety investigations may serve as:

- a repository for information dissemination and common learning
- a problem provider for knowledge development and systems change
- a public safety assessor and public spokesman beyond and above parties involved
- a peace keeper in high military conflict risk regions in establishing exclusion of deliberate acts of violence between countries by attacks on civil aircrafts.

Safety investigations do not have the option but to render advisory opinions to assist the resolution of disputes affecting life or property.

Safety investigations represent a specific analytic instrument with its own characteristics:

- independent from blame and vested interests of third parties and stakeholders
- a cased based approach, based on a systems perspective
- evidence based with respect to its findings and recommendations
- pro-active learning by developing generic principles, notions and knowledge, combined with dissemination of findings and recommendations on domain and sector specific solutions and change strategies.

Safety investigations serve a triple goal:

- Vision Zero: prevention of fatalities and injuries among the population at risk
- First Time Right and Zero Defects: no socio-economical losses during the introduction of new products and processes that jeopardize business continuity
- A Citizens' Right and Society's' Duty: providing society a timely transparency on the factual functioning of systems.

2.4 ethical aspects in investigations

However, some are critical about the notion of Vision Zero as a absolute statement (Hale 2006):

Claiming zero accidents as a goal denies conflicts and tradeoffs with other goals.

Hale realizes that such an ethical position may raise discussion, because such decisions have to be put in context:

If we accept risk and say honestly to ourselves that some accidents are not worth the cost of avoiding, we have to face the conflict between perceptions before and after an accident. He notices a general dilemma for safety: is it about blame and punishment, *or* is it about learning to do better: '*We cannot have both'* (Hale 2006).

Other notice that it is possible to avoid this dilemma by establishing a School of Thinking in identifying safety deficiencies and system change (Stoop 2004). It is possible to reconcile conflicts and to overcome contradictions by establishing independent investigations as a Citizen's Right and Society's Duty (Van Vollenhoven 2006).

It is possible to achieve a safer aviation system than ever before, aiming at the goal of no fatalities and injuries in commercial aviation, where no fatal US air carrier accidents have occurred since 2009 (Hersman 2012). Systems such as aviation and railways represent a separate category of non-plus ultra-safe systems (Amalberti 2001).
The aviation investigation community considers accident investigation a unique incentive for safety changes through the release of reports and recommendations, especially those who deal with systemic and knowledge deficiencies (Arslanian 2012). *We can have both (Stoop, this essay).*

Are such disagreements purely personal, individual beliefs and values, or can they be traced back to underlying socio-cultural differences between world regions; are they imposed by higher order social, cultural or economical values? An exemplary discussion is provided by the debate on road safety developments in Europe, with the Vision Zero principle (ETSC 2005).
In a debate on how far a Vision Zero can be effectuated, differences of opinion on whether such a goal can be achieved, clarify underlying differences in a socio-economical perspective on safety as a social value. While some emphasize the ethical approach to human life in averting fatalities and injuries and addressing responsibilities at a societal level, others emphasize the inevitability to balance safety against other societal values. They emphasize the need to make cost-effective decisions in terms of a rational socio-economical policy and a human desire for fulfillment, where risk of death and serious injury is a matter of degree (ETSC 2005). At the operational level, such a balancing values dilemma is formulated as the ETTO principle: the Efficiency-Thoroughness-Trade-Off (Hollnagel 2009). Such differences can be traced back to differences in socio-economical models and the value of life in each of these models.

Three competing socio-economical models exist in the Western hemisphere, which are seldom made explicit in debates on safety culture and organizational culture:

- the Anglo-Saxon model of liberal values, dealing with self-relianceness, private entrepreneurial initiatives, freedom and limited social security, with a dominant position for market mechanisms. In this context, safety and risk are based on cost-effectiveness considerations, taking into account probabilities of occurrences and responsibilities of corporate management
- the Scandinavian model of humanitarian values, where social cohesion, common wealth, human rights, and stability of the economy leave more room for governmental control and participation. Safety and risk in such a model, deals with preserving the unprotected from hazards beyond their control. This includes a Vision zero regarding inflicting death and injury on road users.
- The Rhineland model, dealing with providing a human face to socio-economical, political and power relations. In this model a role for governmental control and guidance is foreseen, aiming at a welfare state, achieving consensus between social partners, providing stability on a medium and long terms. With respect to safety, continuity on the long term prevails over short term profit and cost-effectiveness and democratic participation in policy making decisions is stimulated.

Unfortunately, subsequent organizational structures and their functioning at an entrepreneurial as well as governmental level, have not yet been studied extensively by scientific research with respect to the safety performance and failure mechanisms of these models. In turn, such models are based on underlying ethical principles and moral values.

Theory of ethics

At the level of ethics, several theoretical concepts coexist. In the international literature a distinction is made in three consecutive theoretical schools of ethical thinking, each with their strengths and weaknesses (Hoppe 2011).

A first *teleological* school of ethical reasoning is referred to as consequentialist ethics: the moral value of actions is judged by the contribution of these actions to the common good for society. Based on the work of the Greek philosopher Epicurus on hedonism, Hume (1711-1776) developed the concept of utilitarianism, which was elaborated by Bentham and Stuart Mills in the 18th and 19th century. However, this teleological school leaves moral issues due to a normative interpretation of its hypotheses: what is good and for whom? Such a hypothetical good can be immoral; e.g. there is no objective answer to the question on the monetary value of human life.

In contrast to this ethical school of thinking, a second school of *deontological* reasoning states that assessing actions and decisions according to universal rules of morality dealing with being morally right or good, irrespective of the consequences. This school is linked to the name of Emmanuel Kant, the German philosopher who laid the basis for this school in his The groundwork of the Metaphysic of Morals. His Categorical Imperative should apply universally and without exception: "Act only on that maxim through which you can at the same time will that it should become a universal law''. This means that we should only adopt plans of action or treating other persons that we would be willing for anyone or everyone to adopt. While the strength of Kantian moral philosophy emphasizes the inherent dignity and worth of the moral agent, it has its limitations in leaving no room for exceptions and gives very little guidance for dealing with conflicting moral duties, excluding moral dilemmas (Hoppe 2011).

The third school of ethical reasoning deals with *individual virtue* and is linked to the names of Aristoteles, Thomas of Aquino, David Hume and Friedrich Nietzsche. This school is teleological because the reasoning is goal oriented, aiming at the highest good or happiness, achieving self-realization. This ethics is not based on action, but refers to character, in which virtuous persons can be exemplary for other individuals. Cardinal virtues can be defined as courage, prudence, compassion, trustworthiness, benevolence or justice. Because such virtues are indeterminately variable in individual cases, judgment concerning individual cases must be left to the wisdom of each person (Hoppe 2011)

Although each school has strengths and weaknesses, their theory or applicability are not flawed as such. Their usefulness is context dependent and can provide us with deeper insight in the morality of stakeholders in complex systems such as aviation. Their decisions and moral grounds for action should be made, providing transparency in the specific socio-technical system characteristics and properties.

Societal and temporal context

A major weakness of each of these ethical theories is the societal and temporal context in which they were developed: before the industrial revolution ethical philosophers were not able to identify the relative autonomous role of technological development and their inherent or emergent properties and inherent hazards.
Such weaknesses manifest themselves in particular in distributed networks with delegated responsibilities and a 7/24 character such as aviation and maritime, where the consequences may have a global and long term impact, such as with emissions of greenhouse gasses, nuclear power waste or environmental pollution.
In global transport networks such as aviation and maritime, specific principles, notions, tools and techniques are developed to manage and control unforeseen, unpredictable and unprecedented consequences. By creating a high level playing field for an high technological industrial sector, safety is regulated and managed on a sectorial level beyond the corporate level or individual moral responsibilities.

A most prominent principle is the Precaution Principle. In case of an uncontrollable hazard on lack of understanding of the actual failure mechanism, the motto is: First comprehend before control hazards and gain insight to create oversight over potential consequences. Grounding of aircraft in case of inexplicable deficiencies, banning poor performing carries from parts of the network by blacklisting and revoking of type certificates and pilot licenses are applied on a regular basis. Making a go-around is a standard operational procedure in case of an unstabilized approach.
In addition, delegated responsibilities are distributed over the operators in the network facilitating authoritive and qualified decision making on a local level and dedicated actions in specific operating conditions and environmental constraints. In the aviation and maritime community, individual operator capabilities are assessed along lines of Good Airmanship and Good Seamanship and professional judgment by disciplinary law and Codes of Conduct. During design and manufacturing, product safety standards and certification processes are developed to a very high international level. On a corporate level, a Just Culture has been established, facilitating confidential incident reporting, accompanied in most countries by a legal protection of Whistle Blowers. At a societal level, blame free and independent investigations are settled by international standards in the ICAO Annex 13 protocol. EU Directives and national legislation in the majority of nations are involved in the international aviation sector.

Such an approach is quite different from a conventional risk assessment in stationary, stand alone and linear systems, where economical cost benefit analyses are dominant and where risk is defined as the linear product of event frequency and event consequences and responsibilities are established at a corporate level and inherent socio-economical cost-benefit considerations.
Consequently, at a corporate level, in such sectors making so-called 'sacrificing' decisions becomes inevitable: sometimes, something should be sacrificed, in submitting the interest of one party to the interests of other parties. But can we bargain risk budgets similar to European milk and fish quota or reduce safety to negotiating a price for risk budgets? You may erode safety standards and think you can get away with it all the time, until accidents occur (Van Vollenhoven 2006). To the decision makers in charge, it is not only a decision about their own career perspective. It may have an impact on vital interests of other stakeholders as well. Protecting whistle blowers who object these decisions or issuing a Code of Conduct for professional behaviour is necessary, but not sufficient. We should respect the professional dignity of our crews and respect our passengers

(Ten Hove 2005). Like a most experienced accident investigator stated: treat the victims of an air crash as if you were among the victims yourself. Like a CEO of a major European railway company once stated: *if we start to kill our passengers and crew, we are on the wrong track.*

Road traffic victims only get relative little attention from society despite the size of this population at risk (WHO 2004). Individually they suffer from severe traumatic experiences and face long term consequences. They should not be forgotten, having our focus on major events in high technology and high performing systems. Unfortunately, this is no new message. Also at the societal level, sacrificial decisions are made.

In 1994 in his valedictory lecture, De Kroes – professor safety of transportation at Delft University of Technology- stated: *If we do not want to bring financial and cultural sacrifices for safety, we bring human sacrifices (De Kroes 1994).*

Such 'sacrificing' decision making however should not imply a moral judgment on these decision makers themselves on an individual level. There is no sense in shifting blame from pilots to individual managers or governance. It indicates the necessity to explore the non-linearity of complex risk decision making. Blame allocating investigation is the domain of judicial inquiries. Non-linearity deals with notions of 'affect' and 'context' which have to be taken into account in a dynamic decision making environment, representing multiple agent perspectives and multi actor values.

First, parallel to 'ratio', the social-psychological notion of 'affect' is important in risk perception and risk assessment. In the allocation of responsibilities in risk perception and risk acceptance, a dual process in reasoning is distinguished (Slovic 2004). On one hand there is a cognitive rationality (decision making based on rational arguments and validated knowledge) while on the other hand an emotional rationality exists (decision making based on ethical considerations, based on individual and social norms and values). These processes are equivalent because they each represent a distinct decision making process of the human mind. They complement each other rather than being contradictory (Kahneman 2011). Unfortunately, they are not equally distributed across all stakeholders (Van Ravenszwaaij 1994). Risk bearers think along lines of mental pictures, consequences and scenarios, while policy makers think along lines of frequencies, performance indicators and rational utility functions. There is a distinction between 'how' versus 'how often' (Hendrickx 1991). Scenarios are a powerful instrument to bridge this dual process in multi-actor decision making processes (Kahneman 2011).

Second, it is also a question how people come to their decisions in their socio-economical context. In a neo-classical economical approach, a rational assessment of arguments is given by logic decision making rules, which are valid irrespective of the individual and the environment in which they operate. Such an assessment has neither a relation with the psychology of the decision maker, nor with the decision making process. There is only one valid outcome possible.

In contrast with this substantive rationality, the institutional economical school in thinking adds the environment to this rationality by taking into account the process and context of the decision making. This procedural rationality is intendedly rational, but lividly so. It is dependent on deliberations and depends on the process. It is determined by search mechanisms and storage of patterns, transaction costs, norms and habits. This rationality focuses on a satisficing and optimizing instead of a maximizing strategy. Investigations into such decision making processes emphasize a detailed empirical exploration of complex decision making algorithms. This school of economic thinking has relations with ’naturalistic‘ decision making theories.

2.5 ethics in engineering and management

With the beginning of the Second Industrial Revolution during the 19th century, engineering became an independent, professional and above all, practical activity for a distinct group of technical craftsman. Civil engineering –distinguishing itself from military engineers dealing with weapon development and warfare- emerged as a discipline. They focused on relief for the working classes and inhabitants of rapidly expanding cities in their needs for housing, infrastructure, sewer systems, water supply and urban development. Simultaneously, mechanical engineers were employed in rapid developing major enterprises in mechanizing production and organisation of production processes.

Such developments became the forerunners of Taylorism and scientific management, aiming at control over working forces, rationalisation and distribution of labour and productivity increase. In addition, a series of major accidents and disasters indicated the need to provide more scientific support to the design and construction of major projects. Through accident investigations, the explanatory mechanisms for explosion of steam boilers, collapse of bridges and capsizing of vessels laid the basis for scientific domains such as thermodynamics, metallurgy, fracture mechanics and hydrodynamic stability. Today, engineers can be personally taken liable for the designs and products of their company by the corporate homicide and corporate manslaughter legislation in several western countries.

Also in the humanities, such ethical debates have emerged, discussing perspectives and commitment on safety versus business efficiency and cost reduction strategies (Harris 2011):

> With the continuous decline of the accident rate in aviation, some scientists emphasize the need to shift the focus from individual behaviour and human error towards a systems approach. With increasing technical reliability and structural integrity of aircraft, it was evident that the major cause of air accidents had now become human error and that this often resulted from the failure of the flight deck crew to act in a well co-ordinated manner. Going through their successive phases of ‘mechanical’, ‘electro-mechanical’ and ‘electro-optical’ research into the workload on the flight deck, disclosed that automation did not so much reduced the workload, but rather changed its nature. After a period of improving ‘clumsy’ automation’ and ’opaque’ flight deck designs, it became increasingly apparent that it was almost impossible to separate design from procedures and training in optimizing the whole system of operations, including the human element. With the general decline in accident rates in aviation, human error is now the primary risk to flight safety. Three quasi-antagonistic parameters can be applied to measure system functioning: safety, performance and costs. Airlines are required to balance the requirements for safety against both cost and performance considerations, *but the Human Factors discipline has until recently concentrated almost exclusively on the safety aspects of this organisational performance troika* (Harris 2011). From a manufacturer’s perspective, providing a ‘better’ human-system interface does not ‘add value’: a failure to provide a pilot-friendly interface merely detracts from its value. As a result, it is often difficult to make a convincing cost-based argument to for a manufacturer to heavily invest in Human

> Factor research. *All modern flight deck interfaces are more than adequate for their job and do not unduly promote error. Minor deficiencies become a training issue to be dealt with by the airline* (Harris 2011). The key to demonstrating the utility of Human Factors in aviation is not to count the cost of investing in it, but to calculate the savings that it makes on a through-life basis. Such efficiency gains are best accrued by taking a systems approach and avoiding the urge to promote uncoordinated, local solutions to problems. Therefore, *Human Factors as a discipline must avoid its natural inclination to rush to claim the moral high ground by marking its territory solely within the realm of aviation safety* (Harris 2011). While increasing levels of specialisation have served to develop the science base, it has also simultaneously mitigated against its coherent application in commercial aviation. However, the opportunity now exists to begin to capitalise on the developments made in the last half century. It is superficially attractive and a great temptation to illustrate many Human Factor principles with an accident illustrating the consequences of *not* applying Human Factor principles. It has a shock and awe value and every accident is usually in itself a good story (and everybody likes a good story*). It should be remembered however that such cases are a* failure *of Human Factors* (Harris 2011).

Consequently, Harris advocates an integral systems approach in which Human Factors should be applied properly by system engineers and designers. However, in view of the Air France 447 and Qantas Flight 32 cases, his assumption that *modern flight deck interfaces are more than adequate for their job and do not unduly promote error* is questionable*.* Implicitly but clearly, Harris acknowledges that safety may establish a scientific reputation for a specific discipline, but cannot exclusively dictate Human Factors paradigms at a systemic level. A collaborative effort between engineers, psychologists and practitioners becomes indispensable, recognizing multiple actor based values, experiences and perspectives, irrespective of disciplinary paradigms, moral standards or commitments to specific stakeholders interests.

Over the decades, on an individual level, individual engineers and designers have fulfilled a role of whistle blowers, advocating improvements for the prevention of disaster, such as Lorenzo Coffin in the railways, Ralph Nader in the automobile industry and Mary Schiavo in the aviation industry. Frequently issues occur which have been preceded by alarming signals, such as with the DC-10 cargo door, the Space Shuttle program and recently, the Lithium-ion battery issues with the B787. Engineers and designers have taken various positions in the debate on ethical behaviour and have practised each of the schools of ethical thinking from both a deontological, theological or virtue perspective. At a collective level, at the turn of the 19th century, a ‘social engineering’ discipline emerged among civil and mechanical engineers, advocating protection of the working conditions, housing and hygiene of industrial workers (Lintsen 1980).

To protect their profession integrity and ethical standards, in several industrialized countries formalized Codes of Ethics and Conduct have been established for the engineering communities. In October 1983, the International Society of Air Safety Investigators issued a similar Code of Ethics and Conduct for their members.

In conclusion

Safety investigators inevitably are aware of such variety in decision making processes, socio-economical contexts and moral framing. Such a scoping of investigations focusing on motives and moral has not yet been applied on a large scale. It could provide transparency in decision making processes and actions irrespective of a judgment on ethical conduct or blame.

3. System change management

Learning lessons from investigations aims at change. In this chapter the various opportunities to transfer learning into change are explored from a systems engineering perspective. Communicating interventions in order to achieve sustainable change is supported by metaphors. A new metaphor might be necessary to comply with changing mental representations of how to achieve a safe systems performance.

3.1 safety interventions

In order to distinguish between complex and safety critical systems and other less complex and critical systems with respect to their required change capability and adaptivity, safety enhancing interventions can be categorized in two main classes:

- Linear interventions and first order solutions. Simple problems allow restricting the solution design space. This is valid only if the number of solutions is small, the number of design variables is small, their values have limited ranges and optimizing within these values deals with sacrificing of aspects among the limited set of variables. Such interventions reinforce the design space *in the detailed design phase* by reallocation of factors, more stringent compliance with operating rules and regulations, elimination of deviations, applicable to simple, stand alone systems
- Complex interventions and second order solutions. Complex dynamic problems demands expansion of the solution design space. Such solutions focus on concepts and morphology, reallocation of functions to components, reconfiguration and synthesizing of sub-solutions, involvement of actors, aspects, teamwork, communication, testing and simulation. Such an expansion of the design space occurs *in the functional design phase* by developing conceptual alternatives and prototypes, applicable to complex and embedded systems.

When first order solutions have failed and do not prevent a critical event, a redesign of the system becomes necessary. This means that in terms of where to start the intervention, a clear distinction has to be made during the analysis between the event as a phenomenon and the system in which the phenomenon occurs.

Consequently, *recomposition of the event should be separated from modeling the system.*

In making the transition from a linear safety intervention towards a complex safety intervention, the concept of critical load is applied. Accident scenarios can be considered critical loads on a system: once the critical load is applied, the system will fail if the loads is increased, exceeding the load capacity under the given operational conditions.

In complex interventions, the focus is on events in a systems context rather than on isolated factors and generic aspects, as is the case with linear interventions.

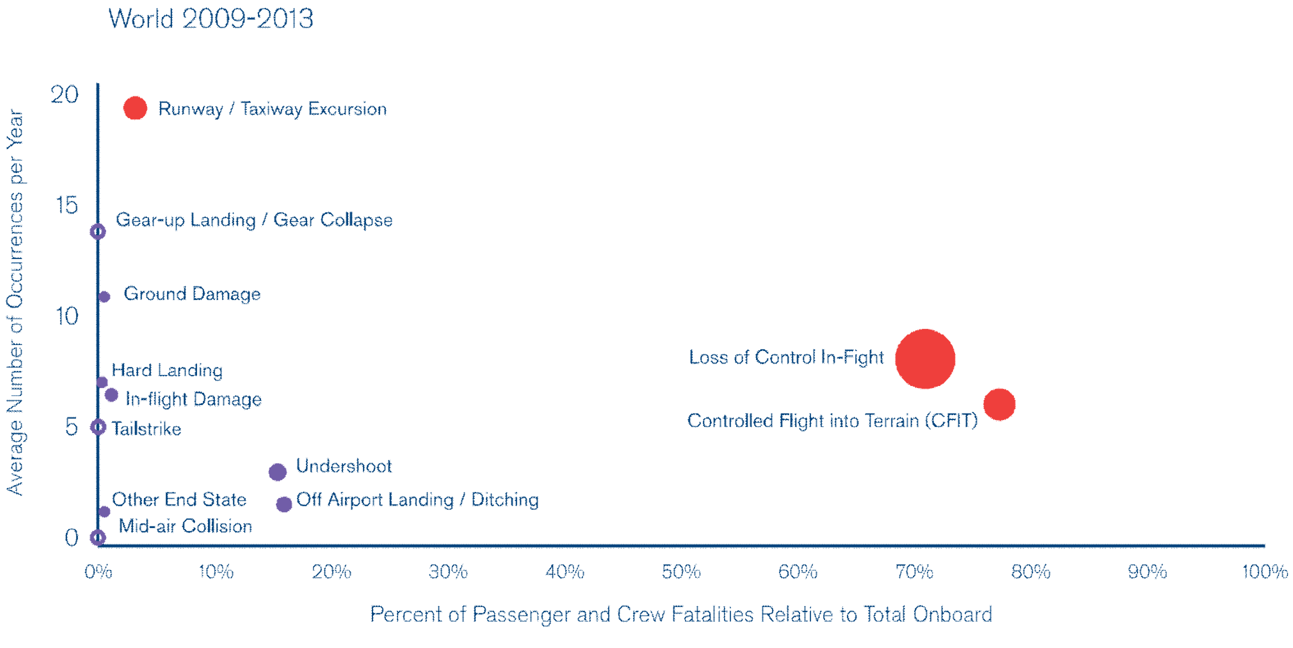
The re-composition of events takes place by identifying and synthesizing descriptive variables into scenarios in their specific operating environment and constraints. Such synthesizing is primarily evidence based. Analysis of the event discloses explanatory variables, providing a basis for knowledge based interventions and identification of change variables.

The redesign of the systems is based on such change variables and is conducted along the lines of engineering principles by generating design alternatives in the enlarged solution design space into the form of a limited set of prototypes. These prototypes contain a relocation and addition of functions, changing the morphology and configuration and incorporate additional actors and aspects. The testing of these prototypes is conducted by running scenario tests, definition of limit state loads and simulation of complex and dynamic systems behavior in virtual reality. Analyzing system responses, before they are put into practice, are based on First Time Right and Zero Defect strategies. The responses of systems can be determined by a gradual enlargement of the disruptions which are inflicted upon the system, until oscillation and instability occur. Responses of systems may become visible by a gradual or sudden transition to another system state by passing a bifurcation point. After such a transition, the safety of the systems can be assessed according to the acceptability of the new safety integrity level, also in a technological sense.

In discussing safety of socio-technical systems, the relative autonomous nature of technological development should be carefully taken into account because of its engineering design capabilities.
In a debate on the differences between scientific research and engineering design, it has to be carefully acknowledged that traditionally, research as the objective of ‘understanding’ as sufficient in itself. Engineering design in contrast, also has the objective to change and design, based on insight and oversight over complexity and dynamics of socio-technical systems. Safety science therefore, belongs to the engineering disciplines with a multidisciplinary perspective. To learn across industries, new approaches should be developed by psychologists and social scientists while more research should be needed to bridge the gap between industrial sectors and sharing of information across design domains, taking into account differences in culture between researchers and engineering designers (Drogoul et.al. 2007, Hale et.al. 2007).
Others refer to engineering design as a methodology with its own inherent forms of logic reasoning, decision making processes, scientific and creative principles (Stoop 1990, Roozenburg and Eekels 1995). Within the design disciplines, a distinction is made between architects and engineers because of their different roles in the creation of artifacts. Engineering design methodologies differ among each other due to their distinct intervention in either transformation, transportation or conversion processes of matter, form and function in a specific application domain (Stoop 2004). The role of safety can be made explicit in various phases and substantive steps in the design process (Stoop 1990).
Physical phenomena occurring in technological processes are fundamentally different from social processes and interrelations between actors, entities and organizations. Technology in itself contains many forms, incorporating invisible knowledge, notions, principles and decisions from previous life cycle phases. The physical appearance of a product and process does not disclose inherent properties, principles or interactions to end-users in their operational environment. Design decisions are frequently made under conditions of high uncertainty and incorporate numerous trade-offs (Anderson 1999, Raymer 1999, Roskam 2007, Obert 2009, Torenbeek 2013). Safety margins and design standards, identification of failure mechanisms, probability assessment, consequence analysis and identification of a design envelope should reduce the uncertainty again to an accepted level. Designers deal with optimizing intended performance and are not in a position to gain oversight into all uncertainties and unforeseen behavior of their designs in practice. Such behavior variance however can be designed into their processes such as with the Japanese design philosophy of Limit State or Critical State Design methodologies.

Finally, engineering design faces challenges that are inflicted upon the industrial sector by the characteristics of the operating environment. As stated in January 2014 by Deborah Hersman, chair of the NTSB: aviation is all about defeating gravity. While maintaining control over each of the physical flight parameters and functions in the conduct of a safe flight, a primary concern of each pilot is to aviate, navigate, communicate and manage the flight in this hierarchical order. In the task performance, a pilot has to maintain strategic oversight, to plan ahead of the aircraft and keep control over the dynamic balance between kinetic an potential energy of the aircraft during all flight phases and aircraft configurations (Faleiro and Lambregts 1999). A pilot should mandatory deviate from procedures and regulations in order to cope with unanticipated variance and conditions in order to comply with principles of Good Airmanship and his delegated responsibility in a distributed network (Stoop 2004).

Some characteristics of flying are inherently dangerous and cannot be designed out of the system or fully controlled during operations. According to EASA, such characteristics bring inevitably types of events: mid-air collisions, loss of control and controlled flight into terrain remain potentially disastrous events (See fig 1).



Note: Circle size increases as total fatalities increase; circles with white centers indicate no fatalities.

Fig 1 Inherent risks of flying

Designers need an intellectual counterpart in assessing the safety and operational performance of their designs. Such a role is historically provided by accident investigators and safety managers. To fulfill their role, their expertise should be involved in the design process. Consequently, such a collaborative engineering design methodology may provide a perspective for improving the safety performance of complex systems at a socio-technical level.

The potential for systems engineering design in providing safer solutions requires to:

- Identify inherent properties before they manifest themselves as emergent properties
- Deal with complexity and dynamics by focusing on functions rather than on factors
- Focus on design principles and properties rather than optimizing performance
- Introduce systems dynamics by synthesizing interrelations into accident scenarios
- Apply a proof of concept by testing solutions in a dynamic simulation environment

Therefore, it is necessary to:

- develop event scenarios separated from systems models
- develop prototypes of safer solutions
- create dedicated virtual systems models, representing their specific properties
- facilitate testing and validation of safety solutions in these virtual models, parallel to the real system.

Instead of identifying causes in order to establish the involvement of factors, actors, their motives and interrelations during the event, the operational performance of the *system as such* becomes relevant in the potential change towards a safer performance and the ability to learn from undesirable disruptions. Historically, safety oriented interventions have been focusing on elimination or mitigation of factors, actors or aspects, breaking up the sequence of events in order to prevent its recurrence. Instead of the event and the causal relation to the mishap, identifying systemic deficiencies and knowledge deficiencies become the critical issue in system change and knowledge development.

A conceptual shift from control to intervention

If we shift from managerial control strategies towards applying an engineering design approach to safety at the socio-technical level, what does this mean for the accident investigation process? How do we substantiate such an engineering design approach in the accident investigation methodology? How do we substantiate the concept of resilience engineering in socio-technical systems practice (Hollnagel et al, 2008)? In order to do so, a communication tool is required to discuss such integration.

Safety interventions are gradually shifting from optimization of production processes, towards adaptation of system properties into innovation of concepts and technology by redesign.

In order to achieve such redesign, the event must be redefined in the first place by applying an engineering design methodology (Carper 1989, Stoop 1990; Petroski 1992, Dym & Little 2004):

- *decompose the event* to identify contributing variables and their causal relations
- recompose the event by *synthesizing safety critical variables* into credible scenarios
- provide *analytical rigor* to the scenarios by identifying their explanatory variables, based on undisputed empirical and statistical evidence and scientific research
- make the *transition from explanatory variables towards control and change variables* in order to adapt the systems properties and functioning
- develop *prototypes* of new solutions
- test the prototypes by *exposure of the prototype to the accident scenarios* in a virtual simulation environment

- represent the intervention process and outcomes in a *communication metaphor.*

3.2 the use of communication metaphors

In discussing an investigative process with outsiders, the rationale and results frequently are made transparent to the lay public by applying metaphors. Over decades, the Domino theory, Iceberg principle and Swiss Cheese metaphors have been popular representations. However, their powerful communication capabilities are also frequently mistaken for theories, principles and models with descriptive, analytic and even explanatory authority (Hollnagel and Dekker 2004, Stoop and Dekker 2012). Such metaphors are based on traditional, mechanistic representations of occurrences and suggest a logic in presenting solutions that can not be validated, verified or falsified. Their existence is beyond scientific proof or denial.
Taking away a domino stone to eliminate a sequence of events, toppling icebergs to disclose multiple incidents, shifting slices of cheese or removing rotten apples from the basket have symbolic value. They provide no solution to safety problems or do not avert disaster and bad luck, similar to avoiding walking under ladders, encounters with black cats in the night, hanging horse shoes over a door or painting a vigilant eye on the bow of a vessel.

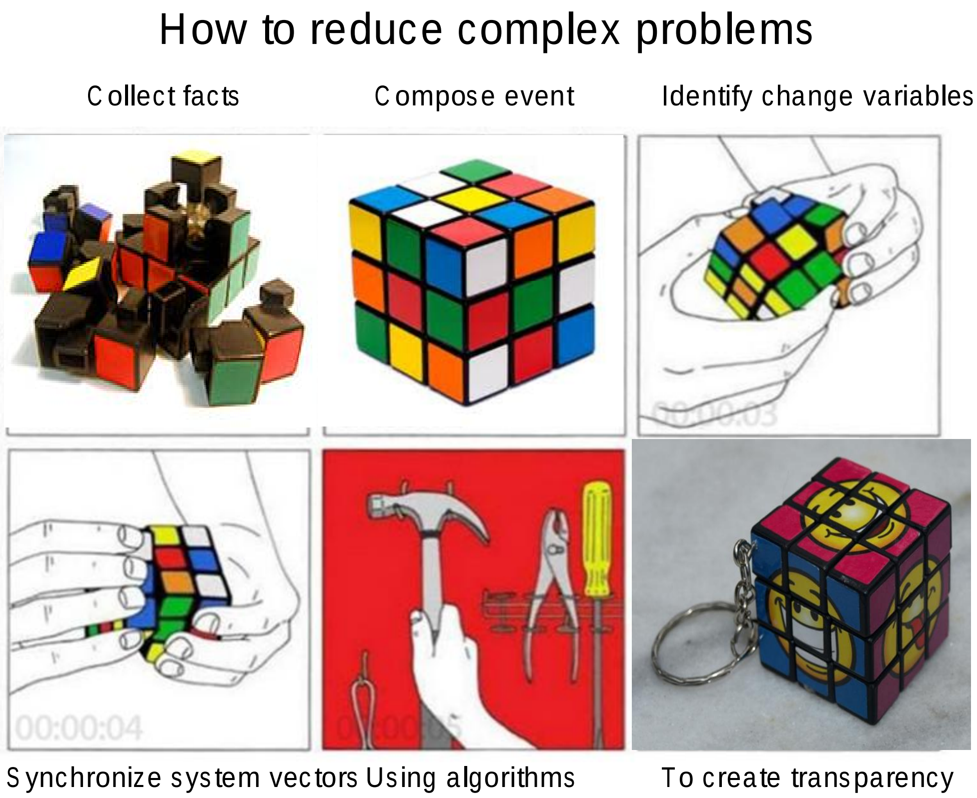


Fig 2 The Rubik Cube

Their strong communication potential does not yet exist for IT systems and mathematical modeling in representing safety issues. Such IT representations should create similar opportunities to explain and discuss IT issue with lay people and practitioners.

Visualizing the illusion of solving complex problems by rational manipulation of elementary building blocks through applying mathematical algorithms is demonstrated by the Rubik Cube. In following such a metaphor, rationally rotating attributes of safety vectors, irrespective of the starting positions, should lead to consistent configurations and homogeneous solution spaces. Such solution spaces could be considered vectorial Eigen Values of complex systems, each representing specific values, attributes and metrics (Stoop and Van der Burg 2012).

The Rubik Cube provides an interesting metaphor for mathematical modeling and multidimensional optimization of safety issues, representing dynamic events in their systemic context. Such a metaphor can be translated into a three dimensional solution cube as depicted in figure 3. The dimensions of the cube will be explained and elaborated upon in chapter 6.

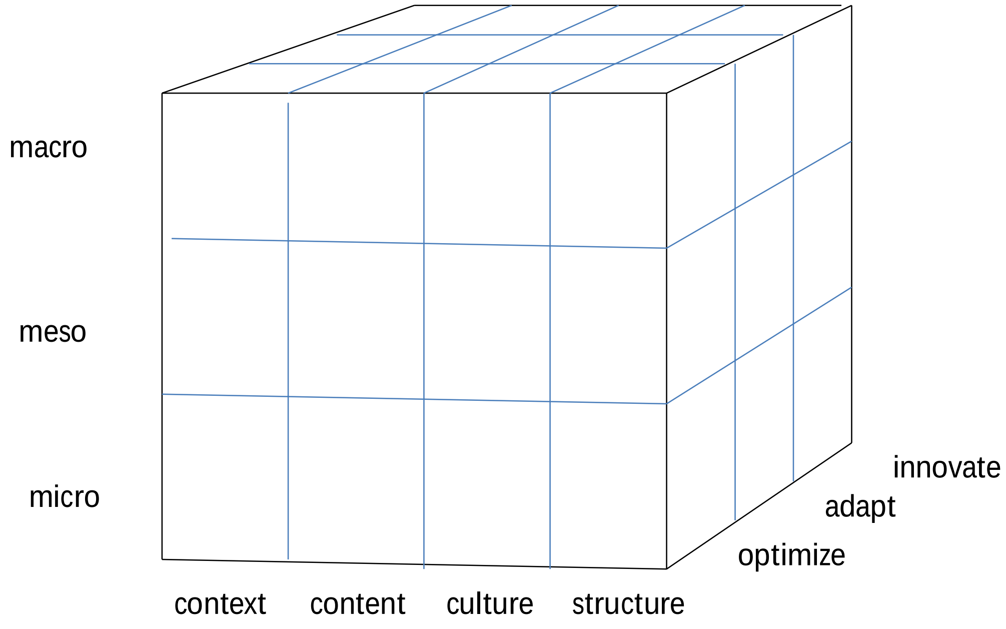


Fig 3: The ESReDA Cube

3.3 barriers for learning

In this chapter five barriers for learning are identified:

- the event and the system are not separated. Diagnosing the event becomes subordinated to unravelling the dynamic behavior of its operating environment.
- there is no stop-rule in an ever expanding scope towards ‘complexity’. Eventually complexity readdresses the focus from causation to controlling consequences.

- there is no classification of socio-technical systems, based on their inherent hazard characteristics as systems with specific high energy densities. Releasing such energy contents may create potentially catastrophic consequences, in particular in the transport and energy sector and process industry.
- separating design from operations removes the opportunity to relate related failures and deficiencies to managerial and governance control strategies.
- there is no adequate communication metaphor for achieving consensus across actors and stakeholders in dealing with non-linear system properties.

4. Confusion on notions

Dealing with complexity from an IT and mathematical communication perspective may provide a new and more appropriate metaphor. Such a metaphor however does not provide a scientific notion that can be applied during an investigation. Complexity as a scientific phenomenon has emerged during safety discussions over the past two decades. Complexity can be discriminated along lines of mathematical and social lines, each with its own interpretations, applications and potential in providing transparency. Complexity is a social construct in itself.

Learning from events is based on knowledge acquisition, management and innovation. Each of such notions too has its own history and interpretations.

4.1 dealing with complexity as a mathematical issue

Throughout the design of artifacts, products and processes, leading sectors in industry and technology have been exemplary for monitoring and diagnosis of opaque systems, unspecified problems and intractable interrelations (Petroski 1992, Carper 2001, Hollnagel and Woods 2006). Their metaphors have been used to communicate across disciplinary boundaries, with operators, management, governance and with the lay public. As such, the terminology has adapted to the notions behind these metaphors, based on the scientific notions and principles in such leading technological and disciplinary domains. Applying modern computer metaphors in diagnosing and understanding complex and dynamic systems, the notions of 'complexity' is widely used in safety debates. It is considered of primary interest in tackling the issues of relative ignorance about the full behavior of the system during design and operations under underspecified operational conditions and incremental adaptation to a changing environment (Meech 1992).

A question is: *where does this notion of 'complexity' come from and what are the underlying notions and developments?*

Complexity as a notion emerged with the introduction of information technology and software design. By general definition, complex systems present problems in mathematical modeling, where relations between parts of a system create collective behavior through which the system reacts and relates with its environment (Wikipedia). The equations from which complex systems are developed are derived from statistical physics, information theory and non-linear dynamics and represent organized but unpredictable behavior. Since physical manifestations are not defined in a virtual reality, references are made to the mathematical information model, without reference to the underlying physical subject the model represents. Such references are in contrast with modeling of physical systems, represented by the artifacts, products and processes as designed by engineering design processes manifesting themselves by their observable outcomes as disruptions, incidents and accidents.

Complexity as a notion is derived from the combined domain of artificial intelligence, cognitive psychology and neurosciences which have provided the base for computer technology, information processing and system dynamics. In order to comprehend unforeseen situations and to avoid disaster, IT design methods make use of network theory, software learning algorithms and simulation and gaming tools. Predicting the performance focuses on identification of performance variability and systemic drift due to long term change to the environment. The focus is on prediction of the actual output –either safe or unsafe- rather than on deviation from normative and designed and testable performance due to an intractable variance in potential system states, configurations and modes (Meech 1992, Hollnagel 2012, Stoop and Dekker 2012).

In addition, development of open access global networks create opportunities for a wide variety of communities to establish networks without formal authorization from a regulator who manages the development and access to such networks. Such virtual networks more and more have characteristics of organic growth and expansion, disclosing opportunities for services and communication beyond any regulatory or governance oversight.

Such an approach is in contrast with design methods for physical systems which focus on decomposition, components, structures and predictable performance assessments by testing in physical reality. In compliance with the scientific reductionism and logic positivism of its times, a conventional mechanistic approach dissects a system into its components in order to understand the behavior of complex systems, to identify the task elements of an operator and the necessary skills to perform a normative task. In avoiding damage and disaster, the kinetic energy concept is applied in understanding the potential damage a technological hazard may inflict on the system, the operator and its environment. Transfer of kinetic energy in such systems may inflict damage by laws of physics and kinematics. Damage is a direct observable output of a system mishap and a safety critical performance indicator for its functioning.

Equations for modeling physical behavior of complex systems are derived from physical laws, dealing with Newtonian principles, energy conversion and state transitions. Such systems are designed top-down, applying an architecture based on engineering design principles. In such systems design, physical principles cannot be disconnected from control over these principles and require a basic operating understanding of the physical control skills and compliance with control laws. Dealing with operator compliance with predefined standards fits perfectly well with a Tayloristic division of labor and allocation of responsibilities within and across organizations. In such a Tayloristic approach, elimination of the operator from the system as a costly, unsafe and unreliable component often is an implicit, ultimate goal. If such an unreliable component cannot be removed from the system, rule based compliance behavior should be forced upon operators, monitored and controlled by safety management systems. Such a structuralistic and normative approach does not fit into demands with respect to flexibility, timely responsiveness, accurate process control and adaptive approach as encountered in IT systems development. A conventional 'Newtonian' approach to IT systems and cognition in its decision making processes does not comply with the specific nature of its technology and dynamic systems behavior (Dekker 2005).

There is a fundamental difference between engineering design processes with a top-down architecture and bottom-up construct of IT software design, operating process and network configurations (Leveson 2013).

A more appropriate approach seems necessary for IT systems with respect to the man-machine-interface relations in monitoring, controlling and understanding of complex and dynamic systems (Meech 1992). Such an approach can be described along lines of knowledge based, adaptive systems, describing the system in terms of understandable (based on Knowledge Based engineering principles), predictable (based on substantive and professional judgment of designed properties and actual performance), accurate (in terms of required and available resources) and timely (in terms of time required versus time available). In coping with anomalies and emergencies, the role of operators shifts from monitoring to diagnosis, based on a timely exhibit of faults, their diagnosis and remediate actions. In such a perspective, investigating human behavior should also shift its perspective by leaving out the normative component in allocating blame and liability and restricting its assessment of its behavior to the formal logic in their decision making processes. Identification of explanatory variables becomes important by asking the question: why was it reasonable for operators and other stakeholders to act as they did given the circumstances? Such a change in perspective can be formulated as a new view on human error, expanding the Skill-Rule-Knowledge based model of Rasmussen to levels beyond rational decision making, formal algorithms and procedures (Dekker 2005). In such an assessment, on the intrasubjective level, intuitive and subconscious motives may play a -yet not fully explored but important- role on the reflex level of human decision making and responding to critical situations (Gorter and Jaeger 2014). On the intersubjective level, rational decision making in operator performance can be expanded to the level of teamwork, automation attitude and flight crew recovery actions (Mohrmann 2013).

According to Pritchett (2009), three prevalent perspectives on applying automation can be identified: a focus on the technology itself, on automation within the operating environment, and on automation as a team member as depicted in fig 4. For two prevalent forms of aviation automation significant research and design guidance has been developed: automation based on modal behaviors, such as the flight management system, and alerting functions, which direct attention or motivate action as warranted by events (Pritchett 2009).

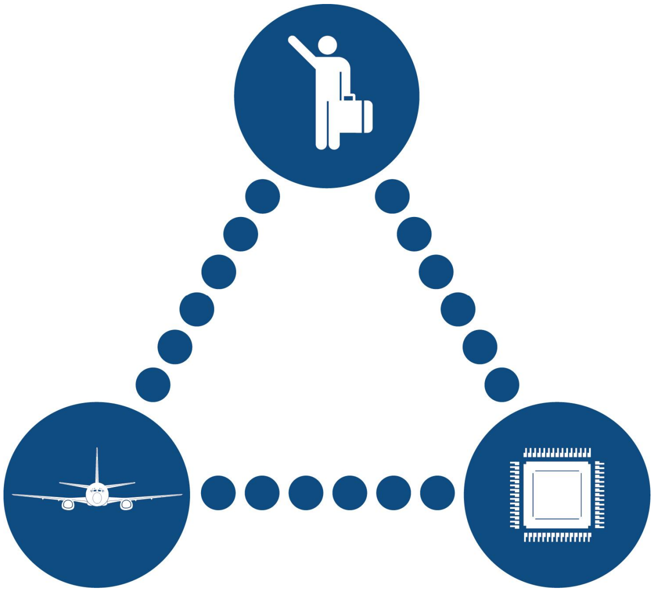


Fig 4: Man-Machine-Interface

To this purpose an operator should have both an accurate a-priori and actual mental model of the system behavior, where feedback is provided by interface management between man and machine (Schneider and Rosa 2011). Since the dynamic behavior of the system directly influences the operator workload, a move to user centered design should be applied for highly technological, complex and dynamic systems with distributed and delegated operator responsibilities in an open network operating environment, such as with aviation (Stoop and Dekker 2012). The discretionary competence of operators in such systems require a careful tuning of the Man-Machine-Interfacing with respect to the machine (related to the physical processes and their recovery potential), the man (related to cognition and decision making processes) and interface (related to maintaining coherence in the operator conception of the systems functioning and creative problem solving) (Bird and Di Paolo 2008). Such an integration however, is not yet fully developed (Den Hertog 1999, Mohrmann 2013, Gorter and Jaeger 2014).

Human error at the in M-M-I systems level

Several major events in aviation such as AF447 (BEA 2012) and QF32 (ATSB 2010) indicate how thin the line is between successful skilled professional responses and a catastrophic outcome in a highly automated environment. The classic notion of 'human error' as undesirable deviation from a normative concept of flight control is predominant among human factor specialists. Human error is commonly accepted among psychologists as the leading artifact in causing accidents. In their perspective, human error should degrade the system from its optimal performance, creating mishap that could be prevented by safety management interventions. As stated by Harris: *For psychologists, the rejection of the concept of 'human error' is difficult to rationalize with the perspective of the system designer employing a formal prediction methodology to help avoid actions that will degrade the system.* Harris concludes: "*When considering human error, first of all pick your perspective then choose your label"* (Harris 2011, pp 100).

Consequently, in this perspective automation would be the solution to human error, resulting in full automated flight. In this concept, there is no space for a critical reflection on the design of a supervisory role and discretionary responsibility of the pilot (Meech 1992, Martin and Soares 2012). Leaving aviation, navigation and communication to automated systems, pilots should restrict themselves to a managerial responsibility, balancing safety against efficiency and costs (Harris 2011). However, such concepts rely on almost flawless automation and extreme low –and therefore negligible- failure probabilities, irrespective of technological imperfections and harsh operating conditions. In practice, such an approach might not be the most appropriate perspective to analyze and understand complex and dynamic interactions between flight management systems and operators (Stoop and Dekker 2012, Stoop 2012). It is a question whether it is possible to incorporate the know-how and know-why of operator experience into the design of safer systems (Morel, Amalberti and Chauvin 2008). Such a design could preserve craftsmanship and 'native' resilience of such systems, relying on a high level of adaptability and professional expertise, experience and tacit knowledge of the operators. These studies indicate potential adverse effects of classical safety interventions in terms of professional reluctance to accept further automation or through the emergence of new risks (Morel, Amalberti and Chauvin 2008).

Constraining operator behavior to the assumptions and limitations of the man-machine-interaction models of designers, trainers and managers in order to improve safety, makes systems more rigid to the detriment of self-managed safety. Supplementing the SRK Rasmussen model with a reflex and crew resource management dimension acknowledges the role of the pilot as supervisor with oversight and control over the aircraft. This role fits in well with the delegated responsibility of operators in a global network with distributed control over the primary production processes based on the principle of Good Airmanship or Good Seamanship (Stoop 2012).

4.2 dealing with complexity as a social issue

Tacit design knowledge transfer

The length of the systems life-span may erode knowledge about engineering design decisions made in the early phases of the system life cycle. Due to the long life-span of transport systems, design considerations, expertise and experience across various generations of designers may be lost due to its tacit nature. If a design philosophy, conceptual design considerations, intended and foreseen use, safety critical decisions and allocation of operational responsibilities are not documented during the early phases of a design, this knowledge will be lost after some time. Within the first engineering design school, criteria and standards on a detailing engineering design have been preserved and documented in order to facilitate a testing and certification of systems components. Its tacit design knowledge about the architecture of railway designs however has been lost with the decease of the chief design engineer and principal members of his design team (Kletz 1991).

Loss of tacit design knowledge may reduce learning from accidents. Accident investigations not only provide 'lessons learned' from new deficiencies, but also may disclose 'lessons forgotten' by a lack of a long term feed-forward loop between design and operations. Such a loop can be established by introducing a formal safety impact assessment procedure, indicating the constraints imposed by the design on operational applications (Stoop and Beukenkamp 2003).

After an accident it may become difficult to reconstruct such undocumented tacit design decisions due to the fact that designers and actors in the operational phase have different rationalities. It should be realized that actors involved in safety issues may have fundamentally different notions of risk and may apply completely different rationalities (McIntyre 2000, Strauch 2002).

Different rationalities

During the conceptual design phase, projects and products are defined by a systemic rationality derived from physics, mechanics, engineering design principles and construction. This phase is linear and confined to specialists. The results of these design decisions are firstly and only exposed to an outsider view and judgment after the detailing phase during prototyping, testing or operations. Risk perception of operators and users is based on a political and societal rationality. Such rationality is defined by interactions with other actors, negotiating and defining social reality in which operators have to cope with the complex and dynamic operational environment. Decisions made by commissioner and designer have led to a product which can be perceived by its physical appearances without revealing the inherent decisions of the earlier phases. Its operational performance can only be reconstructed by its physical appearance and behavior as exposed to operators and users. The technology which is applied is therefore 'to be discovered' for actors during the operational phase, taking the earlier design decisions as incontestable facts, concealed as 'inherent' properties.

Characteristics of the design may manifest themselves during the operational phase by incidents, accidents or disaster as 'emergent' properties. Transparency of safety aspects in both rationalities is a crucial issue since safety may be outbalanced and obscured by other interests of a higher order. Such aspects may manifest themselves only after an independent investigation into major accidents.

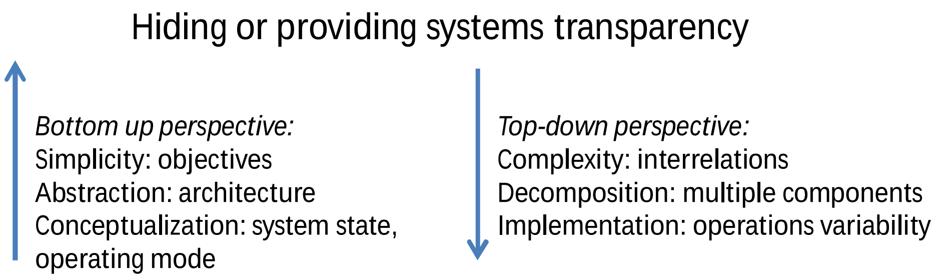
Rationality of a designer and engineer focuses on realization and covers a process of reasoning from goal and concept towards function and form. It follows a synthesizing and decision oriented line of reasoning. Rationality of an operator and user focuses on perception and knowledge. It follows a line of reasoning from observation, perception, towards structure, function and goal. It is analytic and conclusion oriented.

Different lines of reasoning

To understand risks and safety issues two different lines of reasoning are available:

- an ‘inside-out’ vision of commissioners, designers, engineers and other actors which have an oversight of structure and contents of complex systems during their design, development and manufacturing. They are capable of defining complex interactions, couplings and causal relations within the system, risk management, mitigation and control included. They are less capable of dealing with the actual behavior of the system in its dynamic social environment in terms of risk perception and risk acceptance issues.
- An ‘outside-in’ vision of operators, users, risk bearers, regulators, administrators and other stakeholders which have to cope with the system characteristics in its operational environment. They are capable of dealing with global risk notions and causal relations at an aggregated level, but lack a profound insight into the functioning of complex systems. They may concentrate on perception and acceptance rather than on reconstructing event sequences and controlling risks.

Control level	Change rate	Control variables	Control context
philosophy	200 years	context	paradigm
methodology	20 years	systems	framework
technology	2 years	properties	processes
application	0.2 years	tools	data



Hiding or providing systems transparency

Bottom up perspective:
Simplicity: objectives
Abstraction: architecture
Conceptualization: system state, operating mode

Top-down perspective:
Complexity: interrelations
Decomposition: multiple components
Implementation: operations variability

Fig 5: The complexity and simplicity dilemma

An ‘inside-out’ vision is likely to define risk in terms of a program of requirements and standards, as a consensus document for the actual design and manufacturing. An ‘outside-in’ vision is likely to define risk in terms of a defined reality among actors, negotiating risk as a ‘social construct’ to achieve consensus on perception and acceptance between stakeholders. If such a consensus is lacking during events with a high social impact such as disasters, a ‘battleground of subjective safety opinions’ situation may occur (Rosenthal 1999).

In conclusion:
The discourse on ‘complexity’ proves to be an issue of design domain and actor dependent perspectives and has its counterpart in a yet undefined notion of ‘simplicity’.
Complexity is not an objective property of a system, but represents a disciplinary and actor dependent perspective on the actual system functioning.

4.3 dealing with uncertainty

Mathematical modeling
In the academic and societal debate on accident investigations, the issue of dealing with uncertainty frequently shows up as a part of the discussion. While some are dedicated to statistical analysis and probabilistic modeling of accidents, others advocate a case study based approach of single events. Some claim a unique position for probabilistic modeling as the only satisfactory description of uncertainty, based on a long historical reputation, while others reject the old fashioned Newtonian and Euclidian axioms and laws (Klir 1994, Sheridan 2008). Several schools of thinking on how to deal with uncertainty feed the discussion from a mathematical, social and engineering perspective. Uncertainty as an academic notion has its background in mathematics since the late 16th, early the 17th century with the concept of numerical probability, related to names such as Pascal, Huygens, Bernouilli, Laplace, Bayes and Gauss (Klir 1994). In achieving a satisfactory description of uncertainty, a variety of theorems was developed, dealing with various forms of mathematical modeling, conditional probability and imprecise and subjective probability. Klir, in a series of academic debates on various concepts of uncertainty, argues that probability theory is capable of representing only certain types of uncertainty; to capture the full scope of uncertainty, one has to go beyond probability theory. In dealing with disjoint events, probability becomes problematic when we leave the world of mathematics and enter the real world, where scientific measurements are not without error (klir 1994). In creating paths of influences between variables, in the late 70’s and early 80’s graphical representations appeared in the form of influence diagrams. Such diagrams can be considered early precursors of the Bayesian Belief Networks, bridging qualitative descriptions and quantitative specifications (Morales Napoles 2010). According to Klir, scoping of probability theory opens up domains such as fuzzy logic, where natural language and vagueness deal with perceptual and metal constructs and imprecisions (Klir 1994). Simultaneously, a full information paradigm is established by closing the loops between feed forward and feed back learning cycles (Klir 1987).

Hierarchical ordered control loops

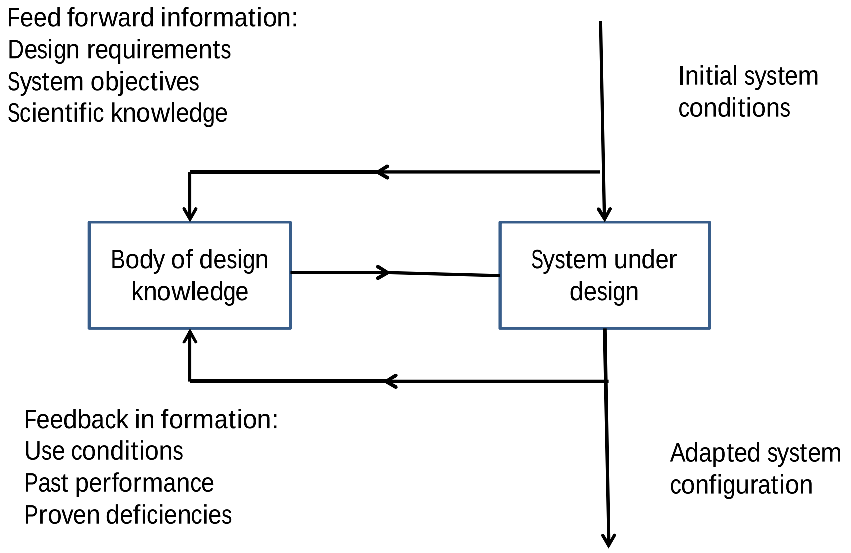


Fig 6: Klir's full information paradigm

Such a full information paradigm can be applied in Knowledge Based Engineering (Stoop 1990, Van Tooren 2003).

Convincing arguments for imprecise probabilities and the inability to provide accurate numbers, are given by uncertainties about the amount of information, ignorance about probability precision, accessibility of imprecise estimates, lack of time and computational resources, classification exemptions and limited consistency in information sources (Klir 1994). In dealing with imprecise probabilities, an additional type of uncertainty has been developed over the past three decades, referred to as fuzzy logic or vagueness. Such uncertainties are context dependent, providing a descriptive and qualitative representation of reality. Dealing with notions of vagueness is substantially different from probability theory, each representing distinct types of uncertainty. In conceptualizing uncertainty in a broader framework, each notions represents a different way of knowledge acquisition. In particular with the development of neural networks, fuzzy set theory has been successful in applications of exemplification; abstracting from examples. In conceptualizing uncertainty, Klir rejects a struggle between Probability Theory and Fuzzy Set Theory, to be replace by a coexistence between the two approaches (Klir 1994). Sheridan concludes deficiencies and difficulties in both established and innovative approaches, each dealing with difficulties in translating empirical observations in mathematical forms and the nature of aggregated versus anecdotal data.

In conclusion:
While in early phases of development, scientific notions are of a qualitative nature, eventually a quantification should take place in order to facilitate engineers to apply such notion in their design processes (Sheridan 2008).

Case study research
In the academic debate on accident investigation, doubts are raised about the validity of single case investigations. Common criticisms on case study research indicate that such investigations should not be generalizable, should incorrectly focus on anomalies of the system instead of the systemic characteristics itself, should have no explanatory potential and, from a theoretical perspective, be intellectually poor by the lack of a theoretical foundation (Yin 1994). In a survey on scientific research methods, Yin emphasizes the differences between the various research strategies; surveys and audits, experiments and case studies. There are various types of case studies; investigations into single event or multiple events, holistic or embedded analysis. In particular in the social sciences, case study research has become very popular. Single event research may preclude quantitative evidence and be of an qualitative and evaluative nature, may reveal substantive and decision making processes and temporal reasoning issues. Case study research enables conducting an empirical inquiry into a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident (Yin 1994). Designing such research requires identification of a specific method, enabling the construct of a preliminary theory on accident occurrences. Such research requires a predefined search strategy, the application of an investigation protocol and manual for practicing such research under operational conditions in the field. The rationale for selecting a single case study approach is given by the holistic approach required in case of unique and rare events, the safety criticality of the case and revelation of previously inaccessible phenomena. Coupling such a single case investigation to trends and patterns that may occur more frequently, leads to a second type of case study research: embedded multiple event investigations. Such investigations enable a quantification of events, by identifying commonalities and patterns across events. A theoretical replication of events may reveal underlying properties such as with Signals Passed at Danger in railways, Loss of Control and Controlled Flight into Terrain occurrences in aviation. It may reveal underlying cognitive or perception processes dealing with situation awareness and automation complacency. Such commonalities facilitate a design intervention into these properties.

Case study research discriminates six primary sources of information: documents/ archives, surveys and audits, interviews, field observations, examination of physical artifacts and participatory observations, of which the last source is not accessible in case of accident investigations beyond the level on on-board computer automated reconstructions. Analytical references for studying evidence is referred to as case descriptions -the event as such-, and theoretical propositions, -the systemic context in which the event occurs-. Such a context requires the building of a systems model in which the event occurs. Analytical techniques are pattern matching, explanation building, time series composition and logic modeling by creating cause-effect patterns between dependent and independent variables (Yin 1994).

In conclusion:
case study research principles provide a scientific theoretical foundation for accident investigations.

Engineering design
In the academic debate on accident investigations, doubts are raised on the validity of ‘old fashioned, mechanical Newtonian perspectives’. Some declare a need for a new paradigm and even an ‘Einsteinian’ perspective, based on quantum mechanics (Hudson 2010). It is questionable whether such a transition to relativity theory based paradigms is necessary. In aviation, Newtonian laws of gravity operate reliable, well within the limits of their conceptual assumptions and boundaries of their engineering design principles. They provide sufficient certainty on the predicted performance of the designed artifacts: aircraft travel at the speed of sound, not at the speed of light. Engineering performance prediction complies with Newtonian laws and principles of physics and vehicle dynamics. Physical representation and prediction of performance has gained an unprecedented high level of certainty due to a sophisticated system of engineering design methods, design specifications, performance envelope and performance limit identification, modeling, simulation, flight testing procedures and standardization of operating procedures.

After such design approaches, a sophisticated system of operational procedures has been developed by introducing safety management systems, incident reporting , introducing safety culture and institutional feedback by independent and objective accident investigation agencies. Due to changes in the technical, economic, social and cultural environment, the engineering design community in aviation is confronted with challenges in coping with new technologies, new concepts in human performance modeling, man-machine interface design principles, certification of artifacts, business modeling, production process optimization and market development issues, demanding compliance with high level playing field standards (EU 2011).
The aviation community faces maintaining the acquired level of safety performance of Non-Plus Ultra-Safe System during design, development and operations (Amalberti 2001, EU 2010, Stoop and Dekker 2012). Conducting safety investigation is a critical component in a systemic feedback loop from reality to design, certification and operation in such systems (Arslanian 2011).

4.4 interdisciplinary or problem oriented?

In order to create scientific attention to safety problems, a systems approach is advocated, characterized by an interdisciplinary problem orientation (Stoop 1990). The concept behind the systems approach is to consider various domains on a similar basis. By applying a systems approach, a complex modeling on the level of socio-technical systems comes within reach of scientific analysis and diagnosis. The approach describes the properties of systems, distinguishing the life cycle and the dynamic characteristics of systems.
However, a systems approach in itself is not sufficient to produce a scientific approach. Normative aspects play a role, dividing lines between scientific and engineering design methodologies which cannot be transgressed without problems.
In practice, debates about multi-disciplinarity have created confusion and competition between technical and social disciplines. Such a debate may create a fruitful basis for a new interdisciplinary

approach –such as bio-physiological mechanics and neurofeedback- it also create controversies between disciplines (Mathews 1978, Sheridan 2008). While mathematical sciences adhere probabilistic principles, social sciences adhere resilience concepts. While 'human error' has been prevalent as a basis for consensus among psychologists for about decades (Harris 2012), it is a questioned notion in engineering design and among pilot organizations in a better understanding of human perception, cognition and information processing. The process of risk decision making is much more complex, incorporating several mental functions and parts of the human brain. Developments in neuro-psychological research have revealed a complex interaction between affect and ratio, between slow and fast thinking in addition to already well known phenomena regarding skill, rule and knowledge based decision making (Slovic 1999, Slovic 2004, Kahneman 2011).

Such scientific debates are not productive with respect to problem solving. In addition to the concept of systems theory, the concept of inter-disciplinarity and problem orientation was developed, primarily inspired by engineering design methodology and environmental impact assessment methodologies (Koningsveld 1989, Cross 1989, Roozenburg and Eekels 1995).

Interdisciplinarity makes it possible to exchange needs and creates focus for knowledge acquisition and development. It structures the decision making process and clarifies the objectives to which this process is submitted.

Problem orientation makes it possible to predict and control problems more efficiently by an a-priori consideration of normative components in the problem definition.

A scientific approach makes it possible to clearly define the object of research and selection of the research methods and techniques. In this process of interdisciplinary problem orientation, cooperation between disciplines, actors and stakeholders plays an important role (Roozenburg and Eekels 1995). Interdisciplinarity does not refer to an finite and definable set of sciences that form a new scientific specialism, but refers to a guiding principle which mobilizes knowledge relevant to the problem definition. The need for information is defined by the bodies that constitute the objectives of the research or investigation and is structured in a decision making process as a part of the investigation process.

Consequently, the process of decision making indicates how a decision is to be taken, while the content of the decision is substantiated by expertise, operational experience and scientific knowledge. A problem orientation is particularly sensitive to misdirection by incorporation of normative concepts which are not made explicit (Koningsveld 1989). Input from disciplines and stakeholders requires precision in identification of the decision making points where such input is provided and interactions take place. Such input on problem definition refers to the context and dynamics of the socio-technical system in which the problem is identified and solutions are to be integrated (Robinson 1982). Such a problem orientation approach also requires a problem solving cycle of recognition-analysis-solution-implementation.

In conclusion:

Integration of solutions does not operate at a technical or procedural level, but represents a normative problem at the conceptual and cognitive level of system design (Stoop 1990). Consequently, identification of the phases and decision points in the decision making processes and a substantiation of the content of each of these decision making points, are inseparable linked.

4.5 knowledge management

Reality checks

Safety investigations also serve the goal of knowledge deficiency identification. Safety investigations are the problem providers for knowledge development.

Historically, on a case basis, investigations have disclosed failure phenomenon that had not been understood before. Examples in various high-tech industrial sectors provide show cases that have triggered scientific developments, establishing new disciplinarian domains. The De Havilland Comet is associated with metal fatigue in jet engines with pressurized cabins, Tenerife and Harrisburg are related to human error and human resource management issues, the Challenger is linked to organizational learning, while Fukushima indicated the limits of statistical data analysis on natural phenomena, Deep Water Horizon demonstrated the consequences of missing a well blow out as a top event in offshore rig risk assessments while AF447 revealed limitations of present generations of human performance modelling.

In their criticism on current practices in accident investigation and risk assessment modeling, several scientists link the criticism to models such as FTA, FME, Event trees and others to the conduct of investigating accidents itself, in particular to the investigation of events. The simplicity of analysis, the linear causality, loss of time as an analytic dimension and limited focus on the operational level and role of the operator should make event models inappropriate during investigations. Their descriptive nature and limitation in the number of failure mechanisms they encompass, reduces the explanatory usefulness and quality of event investigations (Johnson 2003, Sklet 2004, Dekker 2005, Dekker 2006, Hollnagel 2012). The lack of coupling to a systems approach reduces their solution potential (Sklet 2004, Stoop and Van der Burg 2012).

This shift from event investigation to event modeling however, is disputed by investigators: accident investigators do not apply models in the fact finding phase of an investigation, they are applied during design of systems and in the analysis phase of investigations (Benner 1980).

During the design, probabilistic models such as FTA are used in a generic and context free manner to describe a limited set of ‘top events’ which are allocated a certain frequency of occurrence. The eventual risk should stay within acceptable limits or risk levels. If such a frequency is very low, such failure mechanisms are considered acceptable and are not designed out of the system. This is based on the assumption that their occurrence will be fed back to designs and certification processes to enable further mitigation. In practice however, such feedback may be absent by a lack of feedback mechanisms, or fade as weak signals in an increasing information noisy environment. Such failure mechanisms may go unnoticed, until an accident occurs. Several accident such as Turkish Airlines TK1951 and Air France AF447 have demonstrated that there is no guarantee that pilots will detect and correct failure modes in a timely manner that have been overlooked or accepted as negligible during the design process.

Due to minor differences in the operating context and variety in software versions and software migration in a seemingly physical unchanged working environment, such signals may remain unnoticed. FTA and other probabilistic models have proven their value in safety assessment during

design, testing and certification of technical components, but have not provided a failsafe situation for all of the operating environments and contexts in which complex socio-technical systems have to operate.
Consequently, in attempts to overcome these deficiencies, three consecutive generations of models have been developed, evolving from simple to complex, from static to dynamic, from focusing on technology to social contextual dimensions (Hollnagel 2012).

Incorporating mathematical uncertainty in investigating event models shifts the attention from the occurrence towards an assessment of the likelihood of events and their sequential order. Such allocation of a probability in models is irrelevant for the investigation process: after the event the probability is ONE (Benner 1980, 2013).
Shifting from observations in the field by a focus on the event itself, towards creating a model of the event, deprives investigations from a powerful communication rationale: the evidence and context based interpretation of the event and the communication with the outside world. A focus on mathematical models should facilitate a more generic understanding of the structure of events, enabling prediction of similar occurrences in an encompassing accident typology. However, it is not possible to catch new situations, new challenges and changes in operations and context in a predefined model.

As stated by Arslanian of the French BEA: *it is not possible to rely only on a predictive approach. Prediction is not a replacement for correction, but prediction and correction are in fact two sides of the same coin. A permanent screening of available data to identify unforeseen hazards or to better assess risks needs feedback data, sometimes from the unpredictable (Arslanian 2013).*

In the investigative community, critiques focuses on the practical use of models for investigation purposes, discriminating between their application in the fact-finding phase and analytical phase (Benner 1980). As Benner concludes from a survey of investigation practices, for the benefit of structuring information collected it is required to apply a specific methodology. An investigation should take into account each of the events as building blocks, sequencing in a temporal and spatial order to create mental representations for investigators in an advancing time frame. This should facilitate a quality control over the reasoning process and inferring logic in the relations between events (Benner 1980). Such a method transforms experience data into performance improvement (Benner 2010). Using predefined models in accident investigation deprive an investigator from verifying and falsifying models that have been used in design and certification phases and have proved not to be fail safe in practice (Troadec 2013).

Safety investigations bear the element of *serendipity*; finding something out by accident through an open-minded, systemic and in-depth investigation of unpredicted events. Safety investigations are a reality check since preceding modeling, simulation and systemic decomposition during design, development, testing and certification all have their assumptions and limitations. It is necessary to make capital out of experience, to get feedback from the unpredictable, to learn from what we encounter in the field (Arslanian 2011). Even CATS, claiming to be the ultimate causal model for the air transport safety global aviation system, acknowledges that it cannot match reality: *CATS is the second best representation of reality, reality itself being the best* (Hudson 2009).

Safety investigations are reality checks which go beyond the level of ultimate modeling and consequently, represent a unique category of diagnostic approaches.

Serendipity

Such disclosure of knowledge deficiencies by serendipity is in particular of interest for a specific category of complex and dynamic systems, which emerged with the introduction of cybernetics, information technology and computer sciences: non-stationary systems. Such systems have dynamic properties that are time dependent and context sensitive, changing properties and relations over time. Since human behavior is non-stationary due to learning capabilities and the functional relations between man, machine and operating environment are so tightly coupled, that their functionality is developing and adapting over time rather than being specified by design. However, conventional engineering design and construction practices and paradigmatic boundaries between engineers and psychologists have made it rather easier to build systems bottom-up, looking very complex and hard to understand. Braitenberg proposes to call this the law of 'downhill synthesis and uphill analysis'. We may be successful in constructing artifacts, but make them hard to understand.

By giving the design synthesis process too much freedom, the result now can become intractably complex. If we specify lower level mechanistic building blocks and leave higher level and interactive aspect unspecified, we can present the latter as easily observable variables, leaving the system intractably complex. The constraints of conventional computer architectures impose on machines leaves higher levels of interactions 'to be discovered'. Although digital computers have an invaluable 'number crunching' role, they are not necessarily the best medium for building controllers that have to interact with dynamic environments. Similarly, conventional computers architectures might not be the best models of adaptive systems (Bird and DiPaolo 2008)

In exploring increasingly complex systems, Pask proposes that we should base our understanding of complex systems on our interactions with them and the regularities that emerge from such interactions. Again, such a concept is close to serendipity and safety investigations as a practical approach to 'complexity'. Hollnagel, Dekker, Stoop and Benner doubt the usefulness of pre-emptive modeling of accidents, because they provide an a-priori interpretive structure of events, while such modeling depends on the assumptions which should be open to verification and validation. Modeling complex and dynamic systems beyond the level of description is not useful because they contain too much variability, are incomplete and underspecified (Hollnagel 2012). Such systems operate in relative ignorance and knowledge deficiencies due to technological complexity, social dynamics and adaptation to their operating environment.

By detailed decomposition, analyzing the functionalities, operational processes and actual performance in reality, such uncertainties and unanticipated behavior can be described in terms of observable outcomes. The safety criticality of such outcomes can be assessed in a multi actor decision making environment. Consequently, emergent properties can be transformed into inherent properties, reducing the bandwidth of uncertainties. By describing functional and physical properties, relative ignorance can be reduced and system integrity can be maintained under emergent operating conditions.

4.6 knowledge innovation

Dissemination of the findings of safety investigations through reports and recommendations does not only mean public sharing of existing knowledge by sharing best practices, data collection and data management systems. There is also the challenge of providing *new* and *pre-competitive* knowledge to enhance the common body of knowledge in the industry. In case of innovation, such a knowledge development and management is not the exclusive responsibility and role of the industrial sector in which the accident occurred, but has a wider scope. Industry, academia, research institutes, society, each participate in cycles of knowledge and business innovation processes. Innovation of knowledge is submitted to similar processes as systemic innovation processes (Berkhout 2000). Such cycles of knowledge development covers four cycles in the innovation process: science, technology, engineering design and market and service development (see fig 7). Feedback loops exist between disciplinary sciences, technological research, product development and service building. Knowledge innovation processes cover the array of decomposition of knowledge deficiencies in identifying critical disciplinary and domain specific deficiencies, development of new disciplinary knowledge, integration of knowledge in dedicated solutions and the prediction of their properties in normal and deviating operating conditions and various system states. Emerging socio-economic changes in society may lead to existing solutions becoming less effective, and may even initiate generation of new problems. However, emerging developments in science and technology provide society with new solutions, offering new opportunities for the future (Berkhout 2000).

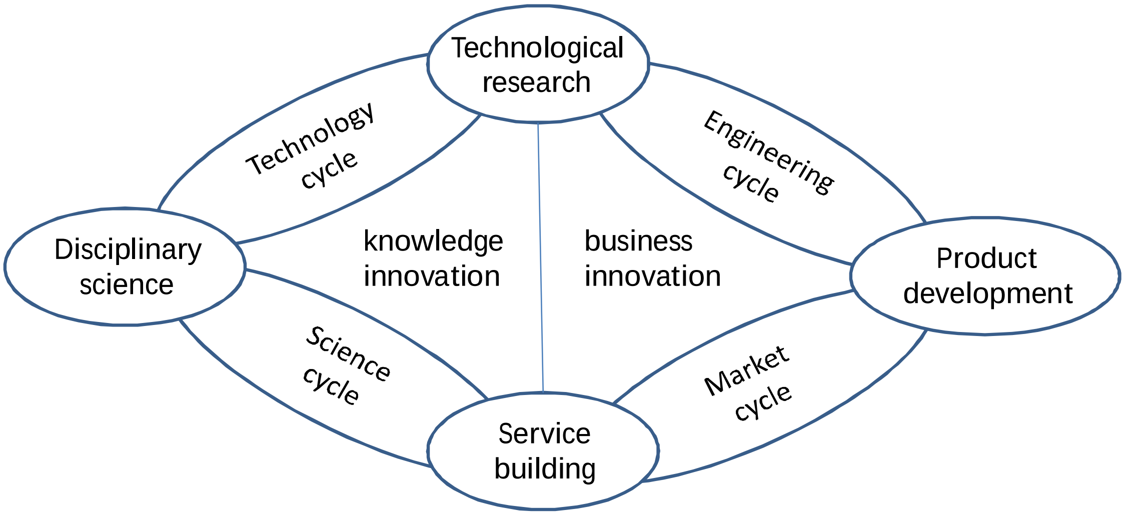


Fig 7: Innovation: interaction between knowledge and business

Since interrelations exist between these four loops of knowledge development, similar conditions exist for knowledge innovation as for business innovation:

- multi-partnerships in and across each cycle
- investing in all four cycles regarding resources and
- communication and coordination across the cycles.

4.7 barriers for learning

In this chapter several barriers for learning have been identified:

- there has been a shift between design disciplines from physical engineering towards information and communication technology and business modelling. Each of these disciplines has brought inherent failure mechanisms and reliability issues to the design based approaches that are not accessible by conventional safety analyses and investigation methods
- during analysis of complex systems at a social level, a diversity of rationalities across actors and agents exist, that require careful cooperation and coordination to achieve a common mental representation of the events and phenomena under scrutiny to avoid discrepancies on assessing safety as a 'technological' or a 'social construct'
- in the academic debate on uncertainties and probabilities, internal scientific discourses prevail on the correctness and scoping of notions and theories rather than the applicability and usefulness of such notions for other disciplines and practitioners. Such a dialectic process of conflicting notions easily creates controversies and does not lead to a synthesis based on a shared problem orientation, a feedback from reality and knowledge development.

5. About Models and Methods

In criticizing current investigation approaches, several researchers and safety analysts advocate the need for a new paradigm. Old fashioned 'linear' event descriptions and a 'fix and fly' remedy in dealing with causal factors should not comply with demands of modern technology, increased complexity of dynamic systems and coverage of the 'whole' system, incorporating all levels, life phases and technical, organizational and social disciplinary perspectives (Leveson 2002).

From the origin of the industrial revolutions on, adding fixes and applying best practices have been unsuccessful in achieving a sustainable improvement in safety performance, due to a lack of understanding of the functioning of technology. Only after developing supporting scientific knowledge such as metallurgy, thermodynamics, vehicle dynamics and stability, aero- and hydrodynamics, the failure mechanisms could be understood. Such supporting scientific knowledge has been developed in the aftermath of series of major events, varying from collapsing bridges, exploding steam boilers, capsizing vessels and crashing aircraft. Such events have created landmarks in physical accident investigations by linking a name to events, creating iconic value for such accidents: Tay Bridge, Titanic, Hindenburg, Tacoma Narrows, Derbyshire or the flight numbers TWA 800, SW 111, AF 447. Others, that almost failed became famous, such as the Tower of Pisa.

Over the past decades, it became apparent that human behavior could not be modeled similar to these physical failure mechanisms. A wide variance in human performance exits, depending on the skills, expertise and experience within the population of operators, dependent on the specifics of the operating context and technology supporting the operator task performance. Quantification of a generic failure probability of human behavior, similar to technical failure rates, has been deemed unsuccessful due to conditional behavior and individual learning experiences (Duffey and Saull 2008). In addition, major events originating from organizational failure emerged, which could be traced back to human performance, crew coordination, corporate management, governance and institutional oversight. Such events triggered an interest from behavioral and social sciences, with a link to names such as Tenerife, Harrisburg, Piper Alpha, Seveso, Clapham Junction, BP Texas Refinery, the Dreamliner Lithium–ion battery issue and many others in each of the transport modes, process industry and energy sector.

The question is not so much whether there is a need for new paradigms –obviously there is a need to expand the scope from a technological to other domains-, but whether such paradigms should replace or supplement existing paradigms. Such a paradigm shift may be accompanied by transitions in scoping and framing of the investigations by a shift from:

- the fact finding mission and accident reconstruction as the basis for further analysis into building accident models and system models, based on a predefined framework
- single event investigation into trend and pattern analysis on a statistical basis
- technological scoping towards behavioral and sociological scoping
- the engineering design phase towards the operational phase
- a focus on the operator performance – the hot seat- towards other actors at higher systems levels.
- establishing the causes of an event sequence towards mitigation of the consequences.

In the next paragraphs, these shift are explored by assessing the STAMP and FRAM as two distinguished representatives of new thinking in accident investigations.

5.1 paradigmatic shift in modeling

The STAMP model: a paradigmatic shift in modeling?

In order to enhance the analysis of human behavior in accident investigations, Leveson proposed an shift in paradigm by developing an improved model, the STAMP model (System-Theoretic Accident Model and Processes) (Leveson 2002).

Does this STAMP model actually represent a paradigm shift?

In her plea for a new model, Leveson rejects event based models and the notion of causality. In contrast to physical events, in event based models, selection of typical events, the selection of conditions to explain the events are subjective. Such selectivity and subjectivity provides room for allocation of blame.

The new model contains several new elements:

- a management commitment in occurrences need to shift from ‘cause’ to ‘reasons why’. Safety degradation may occur as the result of a series of decisions, instead of a single ‘root cause’ that could have been eliminated, preventing the accident to happen. In accordance with Rasmussen, eliminating such a root cause is not sufficient in preventing the accident, because another root cause could easily release the accident in another point in time or place. Explaining accidents in terms of events, acts, errors is not considered very useful for design of improved systems (Rasmussen 1997).
- the model should guide a *comprehensive analysis of the system as a whole, including social systems, technology and their underlying science*s. Without understanding the purpose, goal and decision criteria to construct and operate a system, understanding and prevention of accidents is impossible.
- Stepping up to the systems level. In contrast with component related accidents, systemic accidents are to be considered relevant for understanding mishap. Instead of component failure accidents, shifting attention towards dysfunctional interactions would address systemic accidents. Modern systems are considered too complex to analyze, while the lack of concern about interactions between components may stem partly from simplicity in the past, where analysis and testing allowed detection of all potential undesired interactions. In the past, changing the systems design would suffice in eliminating such deficiencies.

Accidents are no longer considered a failure to comply with specified requirements, but are unplanned and misunderstood effects of interactions of components with the systems as a whole. Since such errors are embedded in the systems design, a systems engineering approach should provide solutions.

In this argument for a new approach, the emergence of computer sciences and software design becomes visible: while individual components work as specified, together they can create a hazardous system state, producing emergent properties that have nor been anticipated during design and specification. According to Leveson, most software related accidents have been system accidents, not stemming from a lack of operation, but from operation exactly according to the intentions of their designers. Software design is abstracted from its physical realization, where the logic of interactions remain unnoticed until they are activated in order to fulfill a specific function (Leveson 2002).

Human error

A second notion in the new STAMP model is provided by the concept of deviation: how to identify the notion of human error in the new model? Leveson identifies human error as a deviation from a predefined standard, a non-compliance with pre-described sequences of actions. Such compliance should be forced on operators by management pressure, while a strict compliance with regulations leads to a breakdown in productivity and chaos if operators work 'by the rule'. In practice however, violations of rules is quite rational, given the workload and time constraints imposed upon task performance (Leveson 2002). Tensions between normative and effective procedures are a normal phenomenon in operational practices because of the learning potential of any operator: adaptation to changes in conditions and environment are natural. Errors are an natural part of their search for optimal performance. This decision making and learning is the domain of naturalistic decision making incorporating goals and motives of individuals and the collective operating environment of the professional community in which they participate. Adaptation is based on motives of cost effectiveness and increased productivity. Adaptation may lead to erosion of barriers, degeneration of redundancy and is an optimization process that should be predicted and controlled (Rasmussen 1997). Adaptation may lead to a drift into failure. Remedies for such a drift are introduction of a safety culture, fighting to resist functional pressure of the environment. In an attempt to understand and control such a gradual adaptation and potential drift into failure, Rasmussen and Svedung have build their models for a control based intervention in these processes (Rasmussen and Svedung 2000).

In the attempt to gain oversight over the whole system dynamics, Leveson proposes to redefine safety as an emergent property arising from interactions of components. Safety should be controlled by *reducing the degrees of freedom in design and operation* by introducing constraints in the control process over components behavior and their mutual interactions.

This proposal is a crucial shift in thinking from cause to interaction. Consequently, the interaction models should not specify single cause variables or factors. Safety should not be controlled by eliminating causal factors, but by constraining behavior through management, enforcing constraints on development of systems and by prevention through designing controls into systems structure. Safety analysis should focus on understanding why control was ineffective in preventing undesirable deviations.

Constraints in software intensive systems

In the STAMP model, a change in information technology conceptual thinking is incorporated (Leveson 2002). Computers are so powerful and useful because they eliminate many physical constraints. They eliminate physical laws that limit the complexity of designs. According to Leveson, this is called the *curse of flexibility*. Physical constraints enforce discipline on and control over complexity. By applying software, limiting factors of what is possible, successful and safe, change from structural integrity and physical constraints to intellectual capabilities. It is possible and relative easy to build software we cannot understand in its behavior under various conditions. We can construct software beyond human intellectual limits, becoming intellectual uncontrollable.

According to Leveson, the primary safety problem in computer-controlled systems is not software 'failure', but a lack of appropriate constraints.

This claim of dissociating design from control of software has to be rejected for several reasons:

- Such a claim is hard to validate in perspective of software reliability issues as shown by the K2Y Millennium bug, continuous hacking threats and virus contaminations. As a student at DUT formulated: '*As long as we do not have perfect reliable software, we need human operators to rely on'.* It is questionable whether we can ever achieve perfect software, taking into account permanent version migration and software upgrading, new computer hardware, expansion of service provisions and network developments. There is no such technological stability in IT systems design, development, certification and testing that we can restrict ourselves to a focus on 'constraints'. Engineering complex software is evolved from a craft to an engineering discipline. Where conventional engineering principles were used for the design and construction of computer hardware, software development was largely haphazard, undocumented and highly idiosyncratic (Leveson in: Launius R.D., Krige J. and Craig J. 2013).
- Relaxation of physical constraints also impacts human supervision and control of automated systems. Disconnecting operators from the proximity of the systems under control deprives them from the sensory perception of the system state and process progression by the lack of physical feedback such as haptic and audio-visual and motoric feedback. Such a disconnection has created the phenomenon of automation and situation awareness and automation complacency. Since operators no longer sense the process and system state directly, designers have to synthesize an image of the process and state to generate a mental model for the operators. In particular situations that are not anticipated by designers, operators needs for adequate information and correct feedback are not fulfilled (Leveson 2002). Consequently, accidents may occur. Recent cases such as with AF447 and QF32 have indicated the thin line between disaster or successful recovery. We seem to lack transition algorithms between virtual reality and physical reality.
- Incremental design and development of IT systems –and in particular operational flight control modes- has resulted in inefficient system architectures (Faleiro and Lambregts 1999). A compelling example is provide by the development of flight control systems. The implementation of Total Energy Control Systems control laws for longitudinal dynamic flight control for autopilot operational control modes and Fly By Wire command augmentation could solve most of the problems of traditional autopilots and auto throttles. Such traditional systems suffer from inefficient systems architectures with considerable functional overlap, too much hardware/software and less than optimum system performance (Faleiro and Lambregts 1999). Application of an integrated system architecture could force the total energy rate and energy distribution rate of the airplane to become more in-phase, providing a more systematic tuning of control laws and provide the ability to decouple airplane modes from airplane outputs. Similar arguments for a user centered and integrated approach of safety critical flight performance indicators can be made to the introduction of a direct angle of attack representation to commercials pilots as an indicator for the lift generating capacity in critical flight conditions (NLR 2013). It took

the AF447 case to re-open the debate on the usefulness and feasibility to incorporate the AoA indicator in the flight displays of commercial and general aviation aircraft.

The basics of the new STAMP model are:

- a shift from event to constraint
- cause is a lack of constraints or inadequate enforcement of constraints
- emergent properties should be submitted to behavioral control laws.
- accident occur as a result of lack of appropriate constraints.

In her search for a new paradigm, Leveson acknowledges that her ideas are a new application of existing safety engineering theory as developed in the 1950s by aerospace engineers. In her new model, she combined hierarchical levels in a socio-technical system with the theory of Rasmussen in the field of human-computer interactions (Leveson 2002).

In conclusion:

Leveson claims that design can be left out of the equation, because the engineering design and the technology can be incorporated in underlying scientific knowledge in the new control model has not been substantiated. However, due to the inherent properties of designed technological systems, it is questionable whether an encompassing control model should suffice in *replacing* the existing paradigm of technological decomposition. STAMP is to be considered a necessary and useful *addition* to existing paradigms for safe operation of systems from a technological perspective.

5.2 paradigmatic claims

In general, objections can be made against *paradigmatic* claims which are made from within one scientific discipline or design domain. In general such objections deal with:

- A separation between social, physical and virtual reality events is not possible: although the man machine interaction is ill defined, it is inseparable related to the overall dynamics in its appearances in socio-technical systems. Several recent major events unmistakably have a major component of failure in high technology systems beyond the component level, as demonstrated by AF447, AA587, Fukushima and Deep Water Horizon, while the Man-Machine-Interaction has seen very successful emergent recovery potential contradictory to adding control constraints during emergency operations such as with QF32 and UA1549. In particular Fukushima, Deep Water Horizon and AF447 have revealed knowledge deficiencies in scientific domains such as probabilistic risk assessment, FMEA and cognitive modeling of operator behavior. Does the STAMP model really *guide a comprehensive analysis of the system as a whole, including social systems, technology and their underlying sciences* as claimed by Leveson? Such a guidance cannot be managed by adding managerial constraints on a company level.
- Simplicity of explanatory models may be valid for existing sociological models but not for technological models. There is no simplicity in available explanatory models for technological development in view of over 100 years of technological development, modeling, simulation, certification and testing. Such a development is based on the concept of physical and mental load and limit load capacity and selection and training of operators as being ‘the right stuff’ type of human performers. Increasing the complexity of

models is not useful beyond their specific goals and applications, while stop rules and predefined modeling assumptions should be taken into account. In aviation, generic 'system as a whole' definitions already are in place on a global level, considering the ICAO treaty with its annexes structure and procedural framework. As there might be a need to create new generic managerial control models, a potential flaw in creating comprehensive models of the 'whole' system lies in their complexity and encompassing nature. In absence of a problem definition clarifying the goal of the modeling, such a modeling may become so complex that nobody can have oversight any longer over its architecture, or is able to substantiate the model with accurate and sufficient data, maintain control over its dynamic behavior and let alone, predict future behavior.

- all encompassing models have been questioned throughout times, reflected by the 1968 Stanley Kubrick's movie, 2001 A Space Odyssey, where the HAL 9000 computer as a representation of the ultimate state of perfection had to be eliminated by its operator to prevent the computer to take over. However, by definition, models are an abstraction and reduction of reality, developed for particular functions and with specific goals and specific problem definitions. In creating a model for understanding and assessing safety in the aviation industry, the CATS model (Causal model for Air Transport Safety) has been developed with the ambition to enable control over the aviation safety at an industrial level. As stated before, Hudson described the ambition as: *CATS is the second best representation of reality, reality itself being the best* (Hudson 2009). Such a claim is beyond any scientific verification and validation, similar to previous notions of 'Risk homeostasis', 'Accident proneness', or an almost inextinguishable statement that about 80% of the contributing factors to accident causation invariably can be explained by 'human error'.
- A shift in focus from events towards constraints denies the forensic phase of physical fact finding and data collection in an investigation. Such a data collection and information building phase cannot be replaced by interpretation of findings and information during the analytical phase of an accident investigation. Application of STAMP may demonstrate a potentially valuable contribution for structuring and analyzing complexity and dynamics in social systems and assessment of the adequacy of control structures on a case basis (Nelson 2008). However, such modeling cannot incorporate or replace the technological, engineering and scientific components of complex systems. Such concepts are based on physical laws, mechanical, medical and mental load concepts, performance limits, operating envelopes, certification and testing procedures. By disengaging the technological and scientific components from sociological and managerial components in the system 'as a whole', the STAMP analysis becomes a socio-organizational issue, which is legitimate, but a reductionist approach in itself. It is impossible to eliminate the physical and technological components from an accident analysis in high technology sectors such as aviation, process industry and energy supply.
- A balance between technological and social issues in high tech sectors may be different from to medical and financial sectors, but *replacing* paradigms and excluding established technological concepts seems presumptuous. Major investigations in aviation have demonstrated a continuum of technical, behavioral and social aspects in events such as AF447 (Troadec 2012). Such a replacement claim seems to reflect the desire to achieve scientific respect, comparable to the Weberian and Durkheimian debate in the 19th century and the plea for incorporating human factors in the design of new technologies (Mathews

1978, Edwards 1972). Rather than criticizing engineering design by labelling it as linear and simplistic thinking and rejecting technological-analytical principles, problems in understanding human behavior and organizational failure could be in the deficiencies identified in the human factor and computer technology interaction domain itself.

The BEA report on AF447 concluded that: *combination of ergonomics of warning design, training conditions and recurrence training processes, did not generate expected behavior and showed limits of current safety models* (ISASI Forum 2012).

- paradigmatic shifts might be inevitable in other scientific disciplines as well. In aviation mental health, a lack of expertise with respect to the actual performance indicators and the nature, severity and extend of mental health is noticed (Bor and Hubbard 2006). A distinct difference in perception between pilots and medical experts exists: it is not uncommon for pilots to be diagnosed with narcissistic and paranoid personality. Such a difference origins from the fact that experienced and qualified pilots are able to gain oversight over a situation and pay attention to safety critical details simultaneously, while being very self confident in their capabilities and assessment of the situation. Such abilities are explored more in detail by Kahneman, Slovic and others dealing with fast and slow thinking, affect and ratio in cognitive decision making. In aviation mental health, also changes in perspective take place from a medical and clinical approach towards a more psychological, psychosocial and neuropsychological approach, taking into account the specific mental states and preoccupations of pilots as highly skilled professionals, equipped with achievement motivation, emotional stability, introversion and high conscientiousness (Bor and Hubbard 2006). Such an approach is open to new research domains such as physio-psychological and neuro feedback (Mohrmann 2013, Gorter and Jaeger 2014)
- At the same time a silent transition in the role of pilots is taking place from flight performance standards in handling aircraft towards crisis managers in handling social dynamics during flights. While stressors have been changing over the past two decades, some human factor scientists believe that automation has relieved pilots from a high mental work load, reducing their function to host on board. They plea for a change in attitude avoiding a 'natural' inclination to define themselves solely with the domain of safety, towards a focus of cost savings and increased performance (Harris 2011). The domain should have matured to such an extend that it should possible to capitalize on the developments in the human factor research discipline for the benefit of operational efficiency. Events should be investigated only on the basis of their potential, not on the outcome, while safety could be considered part of the Safety Management System of a company (Harris 2011). Other scientists have criticized the notion that high levels of automation reduce labor costs and mental work load and reject the dualism of human error in the interaction between man and machine (Dekker 2006). Descriptive abstractions of 'loss of situation awareness', 'complacency', or 'ineffective communication' may be common notions among human factor scientists, but they do not provide behavioral variables that can be observed or measured. Preferably, they should be motives for further investigations into a hardly explored territory of norms, standards and values of working conditions and labor ethos.
- Finally, while identification of the inadequate control is not a stopping point, the system control structure should be examined why such control structures were inadequate. According to Dekker (2006), such examination should be submitted to the 'local rationality

principal', because each operator in the system acts according to what makes sense to them at the time. Such a local rationality is an inseparable interrelation between man, machine and interface within their operating environment as indicated by Nelson (Nelson 2008). A turbojet takeoff sequence is submitted to exceptionally tight parameters and contains the most consistent 30 dynamic seconds in aviation procedure regarding the aircraft's energy state and the crew's cognitive processing, passing through several safety critical system states.

Nelson states:

For those who have never had the privilege of performing a turbojet takeoff, it is impossible for words to convey the intensity with which the mind processes in this dynamic environment. Pilot's actions at that time, quite literally, result in life or death for all those aboard the aircraft (Nelson 2008).

From personal communication, with a senior B747 captain, the author received information on differences between passengers and pilots in assessing the risk of flying:

Passengers are aware of the risk of flying, expressed by their concern of falling out of the sky from 10 km altitude, flying in a vehicle with almost the speed of sound that is as thin as an eggshell, filled with fuel for about 40% of its weight, with an outside temperature of minus 40 degrees Celsius. Despite this awareness, they mostly sleep on the way in from the USA to Europe after a good meal and fine drinks throughout their en-route flight phase.

Pilots however, are more concerned about the takeoff phase, where they take off on a finite runway, where between V1 (no takeoff rejection anymore) and V2 (minimal flying speed), they may encounter an engine shutdown in no 1 or 4 (the outer engines), pulling their aircraft to the left or right with full power setting of about 260-300 kN thrust per engine.

Apparently, safety is a matter of perception as well as professional judgment.

Rasmussen's paradigm shift

Several prominent scientists in ICT and social sciences have advocated a paradigm shift in safety in responding to the challenge of dealing with risk in dynamic and complex systems. Such a paradigm shift frequently emphasizes a transition from safety engineering design principles into a risk management perspective at a corporate level in the operational phase (Rasmussen 1997, Rasmussen and Svedung 2000, Harris 2011). Such a paradigm shift should reallocate the emphasis from an old-fashioned 'Newtonian' cause and failure towards applying uniform control strategies in controlling deviations to prevent unintended performance. Such strategies should emphasize learning from success as well as from failure.

As formulated by Rasmussen and others, relying on empirical evidence, based on design and production process assumptions, should lead to a 'natural' migration towards boundaries of acceptable performance. Such a migration should lead to a 'drift into failure' due to efficiency versus thoroughness trade-offs during operations. A perverting of the existing 'Newtonian' paradigm by over-automation and hyper-Taylorisation in a New Economy context is noted, but not yet extensive investigated. Through a focusing on a selective number of variables –time and money- and trade-offs between operational constraints, over optimization occurs, forcing a trend into cheaper and faster performance. Such over optimization however, is not necessarily also better or safer.

Instead of responding to rare, unacceptable major events without statistical significance due to scarceness of reliable data, a top-down modelling from a control perspective of the whole system is advocated by Rasmussen (1997).

A first implicit assumption in this paradigm shift is the availability of sufficient reliable data derived from operational practices. Such an assumption however, is only valid for large numbers of small events in a constant and normal operating environment. A predictive value of such data collection is assumed to exist due to establishing a viable relation between normal performance, recoverable incidents and accidents.

A generic control structure to mitigate risk and undesirable deviations from normal operations would eliminate the necessity to assess safety in the design phase by time consuming, elaborated and expensive experiments, simulations, testing and certification standards and procedures.
Protection against major events would be based on termination of the flow of events after the release of the hazard. Under particular circumstances, the basis for protection should be on the elimination of the *causes of the release of the hazard* (Rasmussen 1997).
Data on human performance in operations, maintenance and management should be collected during operations and used for a 'life' risk analysis, based on an elaborated system of safety performance indicators. Such a predictive approach should be much simpler and cheaper than the analysis of a-priori acceptance of a design. Such performance data could be collected through other sources than investigations by incident analysis and expert opinion extraction, compensating for the absence of abundant event data.
According to Rasmussen, such effective risk management should be acquired through vertical studies of the control structure (Rasmussen and Svedung 2000).
In an evaluation of this paradigm shift, Stoop and Dekker point out that leaving behind the paradigm of cause-effect relations and a design based safety assessment strategy, has a major consequence with respect to identification of the inherent technological hazards in a design and on the safety assessment tools and tool integration in a safety assessment strategy (Stoop and Dekker 2010).
This raises the question whether a merging of both paradigms is feasible or that each high technological sector requires a distinct approach, based on inherent technological and systemic characteristics (Leveson 2003).

A second implicit assumption in the paradigm shift of Rasmussen deals with the shift in perspective from a design based approach towards a managerial based approach. Such an approach almost naturally addresses managerial responsibilities within a business unit or at the entrepreneurial level in a private company. In the transport sector and in general in complex open systems, safety responsibilities are distributed across a network of actors, agents and stakeholders which are assumed to communicate and coordinate such responsibilities. In addition, major infrastructural projects are organised along lines of public-private partnerships, embedded in national policy domains and governmental budget constraints. A partial, gradual deployment and testing of such projects has proven to be very complicated. Governmental reviews or parliamentary hearings frequently expose causes of delay and budget problems in developing high speed railway networks, next generation of air traffic management systems or communication satellite networks. Such projects have a considerable transition time between replacement of

obsolete systems and introduction and deployment of their successors. During such a transition period, frequent upgrades, modifications, successive product versions and reconfiguration of project organisations are taking place under conditions of continuous adaptation of requirements.
Safety is not always considered a strategic value, similar to environment, economy, sustainability or land use planning, covered by adequate tools in a network organisation such as an Integrated Safety Management System, a Strategic Safety Impact Assessment procedure or Critical Size Event assessment (RAND 1993, TCI 2004, FAA 2014).
In the Rasmussen paradigm, a downgrading takes place of the safety aspect as a strategic and societal value in a network configuration to a managerial constraint in a business operations environment with quantified safety performance indicators.

While the Rasmussen paradigm shift may be valid for responding to minor and medium sized events, validation of the paradigm for application on major accidents has not been achieved.
On the contrary, three examples in each of the complex socio-technical system domains have induced arguments to falsify the assumption:

- Deep Water Horizon showed that occupational risk management did not prevent a major well blow out in extrapolation of normal drilling experiences into deep water conditions
- Fukushima demonstrated deficiencies in applying PRA in accepting very low probabilities of water height exceedance beyond design based occurrences, based on limited data sets
- AF 447 demonstrated the limitations of human performance models for assessing and understanding flight performance under abnormal conditions.

In the Rasmussen paradigm, the physical hazards in these systems and their inherent technological uncertainties were not taken into account.

High Energy Density systems

Given the assumptions and limitations of the Rasmussen paradigm, identification of a new, distinct class of complex and dynamic systems becomes inevitable: a class of High Energy Density systems.
This class is characterized by the physical properties which rule the energy content of such systems and which obey laws of physics and chemistry, aerodynamics, thermodynamics, etc.
Their disastrous potential that is released in case of failure of the energy containment, deals with their specific physical parameters and by-products of runaway consequences:

- in the transport sector: *Control of potential and kinetic energy, $E = mgh + 1/2mv^2$*
- in energy production: *Conversion between mass and energy, $E = mc^2$*
- in the process industry: *Kinetic gas theory, E=PV equals nRT.*

Such physical, chemical and nuclear production processes and properties require control over the stability of the system, dealing with design assumptions and trade-offs in creating an architecture, functional structures, system configurations and performance optimization requirements. Such trade-offs are not observable for operators, who are confronted in practice with ‘emergent’ behaviour.
Such a class of High Energy Density systems challenges some of the assumptions of Rasmussen’s paradigm:

- knowledge deficiencies in the expert assessment of system behaviour are not taken into account
- Human behaviour discriminates a Skill-Rule-Knowledge based level of decision making, but restrict this framework to normal operations and rational decision making processes.

Unanticipated system behavior or unconscious mental processes are not taken into account

- The paradigm does not take into account a hierarchy in control functionality: physical realities should be controlled first in order to guarantee system integrity. In aviation the order of priorities is: aviate, navigate, communicate
- Precautionary strategies to control hazards are not identified as hierarchically dominant over mitigating consequences. Prevention of consequences only starts after release of hazards
- In High Energy Density systems, consequences of events can be unacceptable large. Since such systems are untestable in reality and cannot be isolated from their operating environment, a design based assessment of safety properties and consequences remains indispensable. Otherwise, acceptance of such systems should fall back on 'trial and error' approaches or premature release of components.

In conclusion:

Rasmussen's paradigm shift is a necessary condition to overcome scoping biases, knowledge deficiencies and design limitations in the transition that is taking place from a physical design reality towards a virtual design reality. New failure modes, complexity and variance in behaviour require a separate paradigm. However, such a paradigm should *not replace*, but should *supplement* existing paradigms. Several assumptions of Rasmussen's paradigm can be disputed and are open to further falsification and verification.

The validity of applying parallel paradigms as such is undisputed: the phenomenon of light can be described in terms of a particles theory, but also in terms of an electromagnetic wave theory. Both serve the goal of purposeful explanation.

With the introduction of ICT technology and ICT design methods, the traditional role of architecture and conceptual oversight over physical designs has been abandoned. Instead, in software design, bottom-up developments and idiosyncratic applications have prevailed. Such an approach has created notions of 'complexity' and 'emergent properties' and unanticipated operational behaviour. ICT and software designers emphasize the necessity to regain control and oversight over the complexity of such designs to mitigate the 'curse of flexibility'. Reinstalling a hierarchy of levels of organisation for development efforts, tool development and conceptual integration seems necessary.

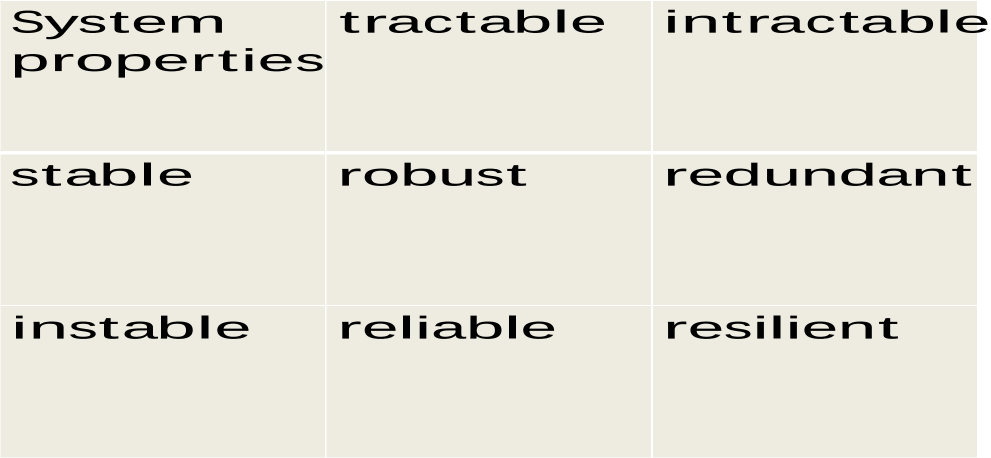
For learning purposes, feedback loops between design and operations are indispensable, due to the inherent characteristics and hazards of physical processes and virtual properties of ICT technologies and their interrelations with human performance.

In recognizing a new class of socio-technical systems as complex and dynamic, with unanticipated interrelations and emergent properties, the physical nature and properties should be taken into account. A specific class of systems should be identified as High Energy Density systems.

5.3 transitioning from model to method

Resilience engineering

Resilience Engineering pictures itself as a recognized alternative to traditional approaches to safety management. This approach does not adhere to safety as the product of compliance with normative procedures, but identifies safety as a systems property, in which performance variability may combine in unexpected ways and give rise to unwanted outcomes. Resilience Engineering pertains to all complex systems, but is of particular interest to high-hazard sectors such as aviation, ground transportation, the military, energy production and distribution and health care. A new concept of safety management emerges, proposing a shift from traditional rule based safety management in a static and deterministic working environment at a corporate level, towards a systems approach which takes into account systems dynamics and the existence of various –either safe or unsafe- system states at a sectorial level. Transitions towards unsafe system states may arise due to inappropriate and insufficient adjustments to change, production pressure or outside pressure. In order to cope with failure, the system should be able to recover from such deviation from intended performance. In this concept of resilience, failure and success are the flipsides of the same coin. Rather than compliance with performance standards, systems should possess the ability to adjust their functioning, *prior to* or *following* changes and disturbances. In this systems approach, the existence of performance variability should be acknowledged, being *controlled rather than constrained*. Such a resilience concept also resolves a traditional controversy between feedback and feed forward control. It resolves the controversy between safety management and accident investigations and instead, integrates these notions.
Resilience Engineering pictures itself as a recognized alternative to traditional approaches to safety management (See fig 8). This approach does not adhere to safety as the product of compliance with normative procedures, but identifies safety as a systems property, in which performance variability may combine in unexpected ways and give rise to unwanted outcomes. Resilience Engineering pertains to all complex systems, but is of particular interest to high-hazard sectors such as aviation, ground transportation, the military, energy production, distribution networks and health care (Van Kleef and Stoop 2014).



System properties	tractable	intractable
stable	robust	redundant
instable	reliable	resilient

Fig 8: Categorising systems

To this extent, resilience engineering adds to the toolbox of safety analysis and systems diagnosis for a specific class of intractable and instable systems by discussing the role of accident investigations and learning from experience. Rather than taking a Durkheimian approach –copying technological notions to social sciences- Hollnagel emphasizes the need for adaptivity, due to either changes in the systems or their operating environment. He acknowledges the importance of technology as a prime mover in systems evolution and introduce resilience as a new concept in the safety management of organisations.

A most prominent issue in resilience engineering is the recognition of the existence of various system states and specific application domains. State transitions are a most frequent phenomenon in socio-technical systems, while the capability of a system to adequately respond to such transitions is not always incorporated in the technological and organisational design. In assessing the value of resilience engineering, Sheridan discusses some shortcomings of both PRA/HRA and resilience engineering approaches. The former have difficulties in translating empirical observations into a required mathematical format, while the latter relies on qualitative judgement on the higher levels of behaviour where data are anecdotical. Future developments of resilience engineering should clarify whether these deficiencies can be overcome (Sheridan 2008).

FRAM; the Functional Resonance Analysis Method.

In dealing with resilience engineering issue, Hollnagel developed FRAM, the Functional Resonance Analysis Method because of the dissatisfaction with the way safety issues were addressed, especially how accident and incidents were explained. The concept of FRAM has been inspired by several scientific sources, in particular dealing with structuring events, second cybernetics, Structured Analysis and Design Techniques (SADT) and stochastic resonance.

Most interesting in the concept is the change in definition of a system: he abandons the notion of structure and decomposition into components in favor of a definition based on functions and processes. As a consequence, the focus of attention shifts from probability towards variability and from relative simple tractable systems towards incomplete and underspecified systems. The inherent relative ignorance that comes with such systems, is reduced by focusing on what actually happens. In this respect, Hollnagel looks at ''as is'' instead of ''as ought'' and thereby eliminates implicit normative judgments which are present in so many other accident models that are widely applied. He re-introduces an empirical cycle in the analysis.

FRAM explicitly focuses on differences between linear thinking and quantification of behavior with direct cause-effect relations versus patterns of events, the relations and unanticipated couplings between events. By doing so, Hollnagel no longer is dependent on the classic notion of ''cause'' and a linear relation in time in terms of ''if-then'' statements.

Instead of interpreting the event in terms of a predefined model, the model is based on four principles:

- failure and success are equivalent
- everyday performance is adjusted to match operating conditions
- outcomes are emergent rather than resultant from causal relations
- relations and dependencies develop in a specific situation rather than predetermined.

Such relations are described by introducing functional resonance which may occur in the system.

A investigative FRAM approach to events does not look for a cause, but tries to understand what should have happened in order to explain why it did not happen. An risk assessment oriented FRAM approach does not start by looking for failure probabilities, but develops a description what should happen in a everyday case. Identifying performance variability should indicate whether the outcome is affected positively or negatively. The FRAM approach differs from traditional HAZOP, FMEA and FTA approaches. In a search for failure probabilities based on these approaches, each component in a mechanical or technological system remains stable or fixed. Such an approach he states, is completely different from socio-technical systems, where such assumptions are not fulfilled. In socio-technical systems, relations and interdependencies among functions are not stable or fixed. Consequently an alternative approach to risk assessment should be found in order not to understand failure, but to understand how several functions may become coupled or how they can combine in unintended ways.

He identifies four steps in modeling such systemic dynamics:

- identification and description of functions, constituting the FRAM model
- instantiation of the model by identification of the potential and actual variability
- analyzing specific instantiations to understand coupling and unexpected outcomes
- propose ways of managing uncontrolled performance variability by monitoring performance indicators and dampening of variability.

In operationalizing the FRAM method, Hollnagel applies the Structured Analysis and Design technique (SADT), Hierarchical Task Analysis techniques and influence diagrams. Since the FRAM method is an analytic approach, it is unclear how the substantiation of the models in each of these cases is achieved. Instantiation of a model requires data, substantive expertise, professional judgment, event descriptions, etc. Where and why to choose a starting point will depend on the goal of an investigation or risk assessment procedure. Such a starting point will rely on either forensic information about incidents and accidents or design and certification documentation. It is questionable however whether relatively simple SADT, HTA and influence diagrams are the most appropriate tools for an encompassing substantiation of a complex and dynamic system.

Hollnagel emphasizes the need to reflect on safety issues from a methodological perspective. He criticizes several current accident models, in particular Accimap, Tripod and STAMP. Each of these models impose an a priori interpretative structure to the event. The value of the results of an analysis therefore depends on the correctness of the model, particularly whether it can be verified or validated.

As stated by Hollnagel, one might hope that models with incorrect assumptions do disappear over time. The main purpose of FRAM is to *build* a model, while the method of *developing the model* must stand on its own. It is questionable whether these current STAMP, Accimap or Tripod accident models can be falsified for modeling complex phenomena beyond the level of descriptive variables.

Hollnagels' FRAM method as a descriptive approach of what happens in a complex and dynamic system by adhering to a method rather than a model can be a valuable contribution to the accident investigation tool kit. It provides a new opportunity to understand socio-organizational dynamics in addition to a technological investigation of complex events instead of replacing technological investigations by sociological models.

In conclusion:
FRAM introduces a paradigmatic shift in socio-organizational analysis of complex systems. It eliminates the discourse on linear cause-effect relations, adds focus on functions and processes to a decomposition in structures and components and facilitates modeling of systems apart from events. FRAM also challenges a conventional linearization of complex phenomena, criticizing simplistic tools and models with predefined event categorizations where conclusions are based on General Failure Types and trivial, context-free outcomes. FRAM focuses on specific engineering design principles as derived from software design and application in the analytic phase of the investigation. Substantiation of FRAM and a case based application however, is still under development.

5.4 towards control of dynamics

In discussing the role and practices of the discipline of 'safety science', the focus of research is defined as the anticipation, recognition, evaluation and control of hazards (McCay 2007). Frequently a debate is raised about the troublesome role of the notion of 'cause' in this context, in particular in the area of incident and accident causation. According to McCay, a critical gap seem to exist where well established, accepted, generalized principles useful in explaining the accident as a phenomenon are extremely limited. Related disciplines such as medical, legal and engineering professions have such principles available for examining, describing and studying their phenomena of interest. Missing such a vital and central core concept within the 'metaphysics' of safety science, may undermine the efficacy and validity of the discipline (McCay 2007). In applying preventive strategies across a broad expanse of the hazard control spectra, a 'Rosetta Stone' might be missing, needed to decipher the phenomena of incident causation (McCay 2007). A Universal Model should be developed to harmonize and synchronize accident analysis, putting all the pieces of the puzzle together. Such a model should take the form of diagramming the incident in temporal and spatial dimensions, adding these dimensions to existing models in order to refine the collection of factors across the various types of evidence (McCay 2007).
Several surveys of the wide variety of data analysis tools that have been developed over the past decades indicate that certain requirements should be satisfied. Unless such tools have a commonly prescribed structure, grammar and content, they perform poor in understanding the interrelations and dynamics of complex systems and deliver poor performance enhancing capabilities (Sklet 2004, Hollnagel, Pieri and Rigaud 2008).

The discussion on safety investigation methodology has demonstrated a diversity in scientific interests in new approaches. According to Van Meer, this diversity has shown that each approach has its own merits and potential to structure the event as well as the system in which it occurs (Van Meer 2010). Forensic engineering and fact finding, structuring and modeling as well as applying a method or process each cover a specific part of the exploration process that is inherent with safety investigation and analysis. A 'one fits all' or 'best practice' approach does neither cover the diversity of aspects, nor does it discriminate between a 'reactive' versus a 'proactive' approach. As in most cases in exploring complexity and dynamic system behavior, a combination of these approaches will compensate weaknesses and deficiencies of each of the approaches. Consequently, each of the approaches is open for improvements and modifications to comply with new requirements and conditions.

There is room for both developing event descriptions, performance and systems models and investigation methods based on a systems approach.

Two transitions are required to facilitate such an approach:

- recognition of a variance in mental modes in dealing with human performance
- recognition of the necessity to deal with system state transition management.

5.5 barriers for learning

In this chapter several barriers for learning have been identified:

- in academic and safety science communities, the ability to model events and systems in their operational environment has gained high confidence levels. Such modeling however, is not related to engineering design methodologies where modeling, simulation and testing already have pre-empted the scope of failure and success.
- in such academic debates, a paradigm shift is advocated in shifting the focus from cause and sequencing based on a fact finding mission, towards control and modeling and statistical correlation between variables. Rather than taking into consideration design assumptions, knowledge limitations and inherent systemic deficiencies, a focus is on control variance by preventing a drift into the margins of a safe performance with unacceptable deviations and a high risk system state.
- applying reduction and decomposition principles, inductive and deductive forms of logic reasoning have enabled an unprecedented improvement of the insights of scientific disciplines in the safety performance of increasingly complex, dynamic socio-technical systems. However, their paradigmatic assumptions have inherently drawn the attention to observable and quantifiable variables, parameters and performance indicators, paying less attention to higher system order variables and properties such as culture, ethics and social structures. Such superstructures however, remain unnoticed but pose constraints on lower systems levels, achievable safety performance levels and the operational flexibility that remains.
- In the academic debate a stagnation of the dialectic process occurs by a repeated redefining of controversies such as cause versus control, technical versus social safety, probabilistic versus deterministic approaches. Instead of achieving a synthesis on a higher level of aggregation, paradigms, principles and properties are redefined, generating new definitions of safety without making the transition from problem definitions to problems solving capacities. Such a stagnation hinders achieving consensus, communication and cooperation between actors in the safety domain and change agents in the system and limits safety performance enhancements.

Synopsis:
The barriers that have been identified in the previous chapters are assessed with respect to the potential of present and new perspectives to overcome such barriers by developing a new school of thinking in dealing with complex systems.

A paradigm shift as advocated by several scientists is a valid plea for their domain -in particular ICT and social sciences- to overcome the inherent assumptions and limitations that are present in the engineering design and technological domains. Such a shift however, proves to create controversies and antagonisms, and frustrates the dialectic process in achieving a synthesis between disciplines, domains and technologies. Such controversies can be considered barriers for learning at a methodological level.

A further elaboration of complex and dynamic systems indicate a need for a specific class of High Energy Density systems, integrating properties of socio-technical systems in an encompassing safety enhancement school of thinking.

Such a school of safety thinking is necessary to serve the goals of qualified and independent safety investigations. Safety investigators should have multiple notions, tools and techniques at their disposal for their efforts to reflect on various school of thinking and challenge their assumptions, notions and limitations. They should have their own school of thinking. They should be able to remain independent, unbiased, impartial and refrain from taking sides in discourses between other school of thinking or scientific paradigms.

In the next part, fundamentals and building blocks are presented and structured for such a school of safety thinking.

Part 2

Towards a new school of thinking in safety : dealing with complex systems

6. Unraveling complexity in practice: tools and techniques

6.1 system change and adaptation

Accident investigation in transport has a long history and can be positioned in each of the four consecutive schools of thinking on safety and risk:

- deterministic: investigating failure in railways, shipping and aviation by on-site collection of facts. The emphasis is on technical causes, operator behavior and design improvements
- probabilistic: investigating failure probability in process industry and nuclear energy production. The focus is on modeling and failure probabilities, quantifiable into costs and benefits
- constructivist: accidents occur due to dynamics in organizations, depending on reliability, culture, learning and are a normal aspect of managing systems operations
- functionalistic: a focus on accidents is widened to safety as an activity in complex and dynamic systems with emergent behavior. The focus is on the process, dealing with various system states, operating modes, trade-offs across functions, resilience and change.

Throughout the discussions between these schools, several controversies have emerged that may have an implicit impact during the conduct of accident investigations:

- historical, systemic differences between sectors in industry, governance and world regions
- blame and liability versus independent, blame-free learning and prevention
- single, unique and unpredictable events or systemic deficiencies oriented
- generic accident modeling versus investigation processes and protocols
- simplistic, linear causality versus complex, emergent dynamic behavior
- technological versus behavioral, organizational and social perspectives
- solutions focusing on the event by eliminating factors, adding barriers and absorbing energy versus addressing systemic deficiencies, resilience, conceptual change and technological innovation.

For accident investigations, the starting point for providing a timely transparency in the factual functioning of complex and dynamic systems is the occurrence that triggers the investigation. Such an occurrence can be either a single accident or incident, a class of occurrences derived from statistical analysis and trend analysis, a recurrent phenomenon within a class of occurrences or a safety critical theme that frequently occurs within a certain time span (Kahan 1998).
In the next paragraphs, a further elaboration takes place on the event, system modeling by state/space vector representations, the usefulness of a dedicated methodology and formal forms of logic reasoning for investigating events, forensic sciences, decomposition of systems, the introduction of time as a dimension in such a decomposition and the introduction of safety as a system value. These paragraphs clarify the relation between theoretical notions of safety investigations and a practical applicability use of these notions in the ESReDA Cube approach. For further clarification and practical applicability, a reference is made to the document 'Case Study' on the ESReDA website.

6.2 the event

Accidents as a process

In experiencing difficulties during the investigation of accidents and incidents, already in 1975 Benner discusses the theoretical basis for accident investigation (Benner 1975). In understanding accidents, the concept of homeostasis is an essential theory. During their task performance, operators aim at maintaining a state of equilibrium in balancing interrelated functions and capabilities in response to varying influences arising as the activity progresses towards its intended outcome. From practical experiences as an investigator, he concludes that an accident is not a single event, but rather a transformation process by which a homeostatic activity is interrupted with accompanying unintentional harm. An accident is a process involving interacting elements and certain necessary or sufficient conditions (Benner 1975). Maintaining of homeostasis requires a continuing series of adaptive responses to perturbations which arise as the activity progresses. He identifies a series of consecutive steps which are subject to the investigation of the sequence of events. During his task execution the operator accomplishes a process of detecting the perturbation or indications of its presence or occurrence, predicting the significance of the data detected, identifying the adaptive action choices that would maintain homeostasis, selecting the best adaptive action, implementing of the action selected, monitoring the effects of the action implemented and deciding whether or not the adaptive response countered the perturbation sufficiently to maintain homeostasis without further adaptive response. He developed this approach into his MES concept: Multilinear Event Sequencing.

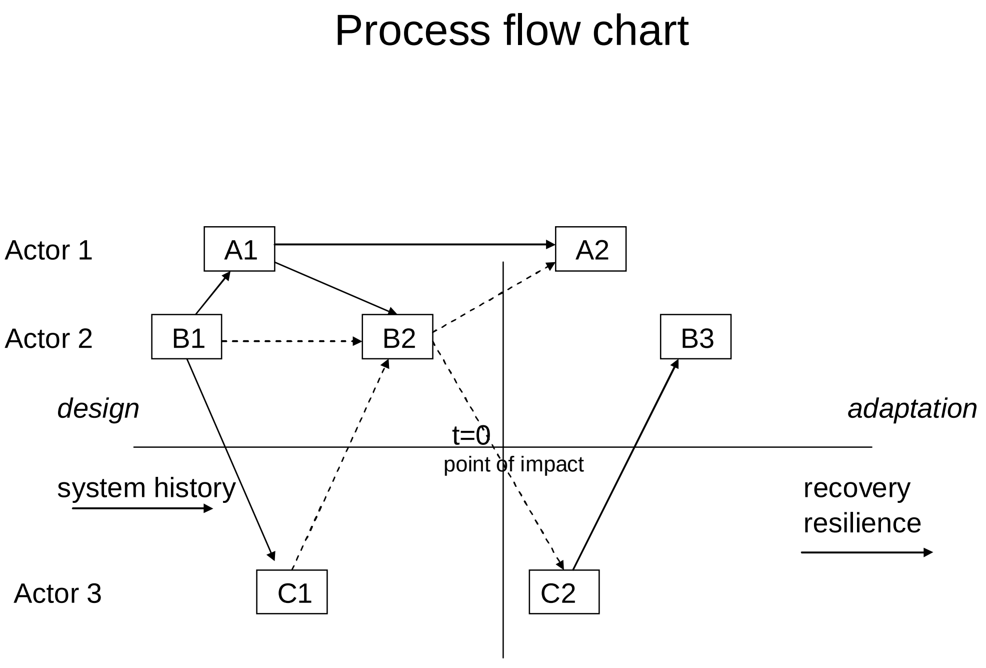


Fig 9: The MES concept

A chart flow diagram represents all the actions by all actors in the specific operating conditions. Instead of a narrative representation of the accident sequence in natural language, such a flow chart representation facilitates a testing of chronological temporal and spatial relationships in a logical and structured order.

The progress of an event is dependent of a natural progression based on physical laws, dealing with transfer of energy, kinematics and physical process dynamics. Benner extends such physical process progression to activities and decisions in responding to perturbations and interrelations that occur during the progress of the event.

During the process of investigation, the challenge for an investigator is to make the transformation from descriptive variables derived from the fact finding phase, to explanatory variables of the analytic phase into change variables in the recommendation phase in order to achieve performance improvement of the system (Stoop 2004). The key element is to transform source data into investigation inputs, analyses and lessons learned. To identify performance improvement opportunities from incidents, a crucial task is to transform data generated by an incident into documented ''building blocks'' to recompose what happened (Benner 2010, 2013). Such a transformation process is not possible by applying accident modeling approaches because such models apply a predetermined structure, causal hierarchy and sequencing of events. If incident source data and their documentation during an investigation are flawed, the entire improvement structure and learning ability is compromised (Benner 2010). Observed incident source data should be transformed into actor/action ''building blocks'' with a common and standardized structure to provide consistent documented data inputs from all sources. By applying a flow chart description of the sequence of events, the investigator should produce a depiction of what happened in a form that others can understand. In addition to data definitions, there is a need to identify data structure definitions, specifying the grammar, format, attributes of each data element (Benner 2010, 2013).

Reasoning logic

In practicing research, engineering design and safety investigations, various forms of logic are applied in exploring, analyzing and solving problems. Such forms are commonly applied without explicitly questioning their validity and applicability for the specific issue under scrutiny. They may however, have decisive impact on the learning potential and actions taken to intervene in the phenomenon.

Logic reasoning forms the basis for any scientific inference that has to lead to an informed conclusion, based on reliable data, relevant evidence and verifiable and objective explanations.

Based on observations, premises are verified or falsified, supported by evidence that provides proof of the truth of the conclusion. Such truth may be absolute or probable, depending on the evidence given in a specific situation. Throughout the centuries, logic reasoning has received great interest and attention by philosophical and mathematical disciplines, developing fundamentals for methods, theories and tools in understanding notions of cause and effect and their relations and the logic rules that must be followed to successfully take certain actions.

Depending on the true or speculative causal relation an intransient or probable relationship between cause and effect should be established. Understanding of the actual relations facilitates learning about the phenomena under scrutiny. Understanding and explaining relations between cause and effect also identifies the potential for adaptation and change by providing convincing proof.

Several principle forms of reasoning are available, categorized in two fundamental categories as deductionist (deriving top down at conclusions under certainty) and reductionist (deriving bottom up at conclusions under a certain level of uncertainty). Reductionist reasoning can take three forms of induction, abduction and construction (Roozenburg and Eekels 1995). These forms of logic reasoning have their preferential applications in different scientific disciplines and domains. They are represented in the following scheme (Roozenburg and Eekels 1995).

	Deduction	Reduction		
		Induction	Abduction	Construction
Form of reasoning	From general to specific	From specific to general	From specific to specific	From general to general
Characteristic for	Mathematics Logic	Physics Social sciences	Judicial sciences Historical sciences Medicine	Technological design Pedagogical sciences

Fig 10: forms of logic reasoning

Deductive reasoning reaches a conclusion by applying general rules that are valid over the entire domain of discourse, narrows down the range under consideration until a certain conclusion is reached. Deduction guarantees conclusions on the occurrence of events without ambiguity on the logic of the conclusions.

Inductive reasoning reaches conclusions by generalization or extrapolation from initial information under epistemic uncertainty. Conclusions are likely, but not guaranteed and leave room for speculation about the causality. In collecting and interpreting the available evidence, discussions may rise over the necessary and sufficient conditions to make the event possible. Inductive reasoning introduces the dimension of time in the interpretation of an occurrence because cause and effect are sequential related in time.

Abduction reasoning goes from observation to a hypothesis that accounts for the reliable data and seeks to explain the relevant evidence. In abductive reasoning the premises do not guarantee the conclusions but infers to the best available explanation. This form of reasoning is introduced by Peirce (1839-1914) and traditionally has gained interest from the fields of law. Modern developments in computer sciences and artificial intelligence have renewed interest in this form of reasoning, where diagnostic experts frequently employ abduction. Despite infinite possible explanations for any physical process that can be observed, a natural tendency exists in common use of the principle to abduct a single explanation for such a process. Reduction is frequently preferred in order to achieve a more transparent and comprehensive orientation to the environment and context in which the process occurs. Abduction is commonly applied in social sciences and artificial intelligence and applies principles of verification and falsification to reduce the number of potential explanations and to exclude contradicting evidence.

A best possible explanation is often defined as simple and elegant, supported by validated data sets and physical evidence. Such a tendency is frequently observed and disputed as linear thinking in discussions on cause and effect relationships and has been known as Occam's razor. Occam's razor is used as a heuristic and is based on the preference for simplicity in the scientific method to explain phenomena. Simple theories are preferred to more complex explanations because they are better testable and falsifiable. However, Occam's razor shift the burden of proof because for each accepted explanation, there is always a number of possible and more complex alternatives. Such an approach opens up the possibility to generate ad hoc hypotheses and speculations where no evidence is available. In accident investigations, applying Occam's razor may be counterproductive: allocating ultimate responsibilities to the operator is a strong and convincing argument which restricts proof to a linear relation between operator and equipment. Interrelations with higher systems levels may be underestimated or neglected. Simple explanations may be easily understandable and communicable, but do not necessarily have to be true: the mantra that 80% of all accidents is caused by operator failure and human error is beyond scientific proof. It can neither be verified nor falsified, but may easily become a social construct among stakeholders. Applying Occam's razor with rhetorical debating capabilities may provide convincing arguments in a public discourse or in court, but does not have to cover a full understanding and explanation of what happened in a dynamic and complex operating environment.
Such linearization however has a strong and valid tradition in judicial applications of abductive reasoning. Unravelling complexity in unlawful actions and establishing blame and accountability beyond reasonable doubt by eliminating suspects to facilitate a fair punishment by judicial court procedures is a cornerstone of the legal system. In the judicial system, admissibility of evidence and credibility of witnesses are subjected to careful rules and procedures. Abductive reasoning can be applied top down by zooming in on achieving undisputable verdicts and accountability of individuals or zooming out by a bottom up recomposition and unraveling of systems complexity and dynamics.

Constructive reasoning support the transition of a description from goal to function to form. Constructive reasoning can be considered the opposite of inductive reasoning. Induction aims at achieving knowledge, while construction aims at realizing an artifact, a functioning material object. Induction is a process of abstraction, construction is a process of concretization and materialization. In the transition from function to form, a wide variety of hypothetical forms is possible. While the intended functionality, use and operating environment can be described in abstract terms on a functional level, no indication has to be made yet on the actual use of configuration, geometry, materials or mode of operandi. In identifying the functionality, predictions on the actual achievement of the goals by the artifact has to be postponed to a testing phase where the actual form and performance are assessed. Constructive reasoning is an open process where many correct solutions are possible. There are no formal algorithms which define a correct and exclusive solution; creativity is required to come to a good solution. This also implies that a variety of system states and operating modes may exist that comply with the design requirements. The end state of a constructive reasoning process is a singular, physical artifact, a process, procedure, organization, system state or operating mode.

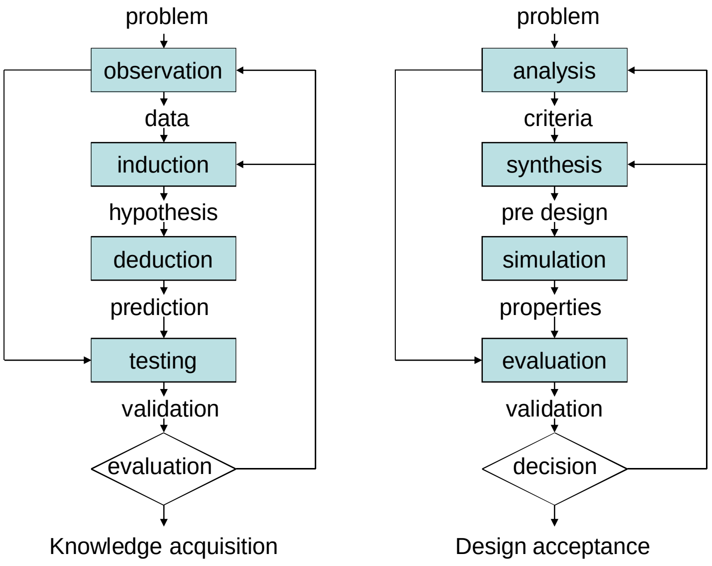
In contrast with deductive and inductive reasoning, abductive and constructive reasoning do not achieve a single absolute or probable 'truth'. They intent to achieve a 'credible and plausible explanation' and a 'goodness of fit' in the transformation from function to form (Stoop 1990). Discrepancies between explanations and imperfect fit depend on the assumptions, available knowledge, expertise, experience and operating conditions between intended use and actual use. This discrepancy introduces the notion of variance, since in operating practices, various operational practices may exist as legitimate, which all can be justified based on the goals, perspectives, tradeoffs, skills, experiences and resources which are available to the end users and operators. Such a variance reflects the trust that end users have in a correct and safe operation of the system under their control. In terms of human machine interactions, the question is to ask:

why it was reasonable for operators to decide and act as they did under the given circumstances.

Posing such questions on the legitimacy of human performance eliminate the issue of blame and liability from the M-M-I problem definition.

By doing so, a difference between investigation for judicial applications and for learning purposes is demonstrated by the notion of admissible evidence in the process of abductive reasoning. Blame and liability focus on the proof for admission of intentional harm and inflicting on willful damage to persons and property. In a judicial process, a conviction is based on direct and circumstantial evidence that is brought up by both parties where admissibility of evidence and credibility of witnesses and experts is contested and a balancing of the pros and cons of the evidence is based on reasonable grounds. Learning focuses on understanding the process where harmful consequences emerge unintentionally during regular conduct of a task, a process or as a side effect in an event beyond the control of any of the actors.
Allocating blame and liability is exclusive in funding the 'truth' while learning is inclusive in the creation of 'trust' by providing proof for an explanation that can create consensus and confidence among actors and stakeholders on the actual sequence of the event. In learning from accidents and incidents, a reasonable, plausible and credible explanation should be provided for the event under scrutiny.
The basic cycles of logic reasoning can be expressed in a flow diagram of successive decisions, representing the thought processes as depicted in fig 11 (Roozenburg and Eekels 1995).

After an initial phase of development in physical product design methodology, constructive reasoning has seen a growing interest among designers due to the application in computer and software hardware and the design of virtual reality systems. Design domains such as aerospace, maritime, mechanical and civil engineering differ from software engineering design. Similarly, in the software design domain functionality inevitably follows from goal definitions and design specifications. However, the final result does not materialize in an concrete form but resides in a virtual reality.



The basic cycles of empirical scientific research and engineering design according to Roozenburg and Eekels 1995

Δ

These two thought processes have their equivalent in the basic cycle of reasoning for forensic investigation processes:

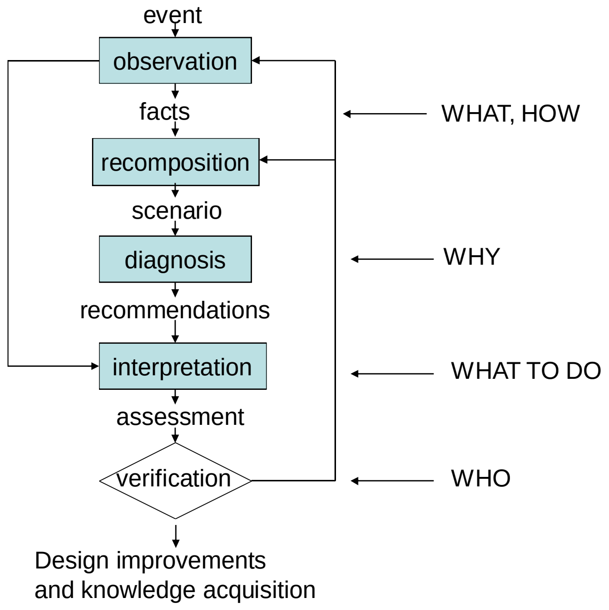


Fig 11: Three forms of thought processes

For the transition from function into form, abductive reasoning has to be applied to create preferential relations, facilitating dynamic allocation of functions to the interface design with optimized tradeoffs between multiple performance requirements and abstract system mode configurations. In most advanced engineering design environments -such as space technology and aerospace engineering design and manufacturing- such tradeoffs are supported by sophisticated methods, software tools and simulation techniques. Multi Disciplinary Design Optimization, Knowledge Based Engineering, Value Based Engineering, Collaborative Engineering, derivate modeling principles and tools are developments based on computational optimization methods which apply an abductive form of reasoning (Van Tooren 2003 , Raymer 2012, Torenbeek 2013).

Cause versus correlation

The extent to which evidence can be convincing is based on one hand on the factual information and the physical evidence that can be collected and on the other hand on the likelihood of inferences that may exist between variables. Two principle search strategies exist for understanding and explaining causes, correlations and contextual conditions. During the generation and testing of hypotheses, cause and effect relations can be either tested in controllable laboratory conditions or can be extracted from documented field experiences and raw data registration systems. The toolbox for such testing varies between the two extremes from field observations towards mathematical modeling. Selecting each of these tools depends on criteria such as required similarity with reality, simplicity of the investigative model, control over the investigation conditions, repetitiveness of the testing and involvement of the number of disciplines in the testing (Rand 1998). Either statistical data or causal models may substantiate the likelihood of the conclusions, depending on the availability of data, their inherent restrictions, scarcity or uniqueness. A most relevant question is raised on how convincing the evidence derived from such analysis is if learning is required and actions have to be taken. Such convincing potential depends on the perspectives and logic reasoning process characteristics of the change agents and whether they apply a scenario based or a frequency based perspective (Hendricks 1991). This raises the issue on whether conclusions should be based on cause, correlation or preferentially both (MacIntosh 2010).

In addition to these practical critical reflections from within the investigation community, the use of statistical data and information has to face challenges with respect to the validity of unpredictable and rare events in complex and dynamic systems. The existence of outliers with huge impact (such as stock exchange crashes, air crashes, earthquakes and tsunami's) cannot be dealt with by predictions and extrapolations of existing patterns and trends. Although it appears to be acceptable in common practices to discharge outliers from data bases in optimizing strategies, such data processing should be done carefully in order not to eliminate unique, safety critical performance indicators from the data sets. Deficiencies in data processing such as outliers, redundancy, noise and the application of incorrect algorithms may reduce the reliability of predictions on future behavior and performance of complex and dynamic systems. Black Swans and Unknown Unknowns emerge if such extrapolations of past performance exclude highly improbable but physically feasible events (Taleb 2007). In Non Plus Ultra Safe systems, such exclusions are not acceptable for ethical, commercial, business continuity and public confidence reasons. A reality check remains indispensable.

Because highly improbable events are very hard to predict, Taleb suggests not to predict such events, but to build in redundancy and robustness in order to make systems resilient to the negative and to focus on the positive for learning purposes.

With the prospects of modern data mining and process mining capacity due to the emerging of big data systems and quantum computing capacity, questions are raised about their validity and usefulness, with their potential to replace 'old fashioned' and obsolete tin-kicking activities in the field.

Is aviation no longer in the business of combing through ashes and wreckage to find answers? SMS gives us the intelligence we need before the problem reaches the headlines. SMS uses hard data to point us in the direction we need to go and do not wait for something bad to happen (as quoted in Guzzetti 2014).

Solid arguments are be made that this is not the case (MacIntosh 2010, Guzzetti 2014). Existing data bases are not complete, suffer from a loss of details and contextual information and above all, are biased towards conventional definitions of human error as the dominant causal factors. The more complex the situation, the less conventional metaphors and models such as the Heinrich pyramid, Swiss Cheese model, Domino Theory are applicable. Incidents are no trivial precursors of low-frequency, complex and major events. In low frequency cases, it proves to be difficult to place statistically reliable significance on findings due to their small sampling size. In the Boeing 777 Heathrow engine rollback case features identified from the data mining could be incorporated in the laboratory testing and, similarly, laboratory testing results could be applied to the data analysis. Data mining proves to be largely complementary to laboratory testing, requiring a balance between the two approaches (MacIntosh 2010, Guzzetti 2014). As stated by MacIntosh and Guzzetti, accident investigation brings together diverse groups of experts, in a focused and structured environment. A synergy of human experience and motivation should be matched with data base analysis. Perhaps the most challenging agony is the credibility of the findings in order to convince decision makers to take the actions needed to prevent the next accident (MacIntosh 2011, Guzzetti 2014).

6.3 scientific perspectives in investigations

The usefulness of forensics

Putting investigations in a systems perspective triggers several questions about the shift in perspective that accompanied such a transition. Almost implicitly, during both investigations two shifts occurred: on one hand a shift from the event itself towards the system under scrutiny, on the other hand a shift from identification of explanatory variables towards change and control variables. These shifts trigger questions about underlying methodological issues as well.

Socio-technical systems are in general, also open, dynamic and complex, but not hierarchically designed as integral systems. The incremental growth and development of such systems does not follow the rationale of a Program of Requirements of a technological design. There is neither a single problem owner nor a commissioner for designing a socio-technical system. At the highest level an inquiry into the malfunction of such a system is the responsibility of parliament by conducting an Parliamentary Inquiry, as the ultimate political responsible authority and problem owner in case of a ''governance accident''. Socio-technical systems are evolving bottom-up, submitted to incremental and evolutionary change in performance. Many of the sociological models are not able to cope with this complexity and dynamics, due to their descriptive and linear nature or aggregation level, lacking the identification of change variables and control mechanisms. With respect to managing safety in such systems, the work of Rasmussen and Svedung on systems hierarchy and control has provided new perspectives, but lacks a link towards technological engineering design (Rasmussen 1997, Rasmussen & Svedung 2000). In accident investigations, such a link is pivotal (ATSB 2006). Sociological models also still lack a dynamic modeling of the hierarchically ordered supply and demand of service provisions, allocation of resources and responsibilities, business optimization and management control mechanisms.

Designers need an intellectual counterpart; an investigator, capable of reconstructing the actual and factual sequence of events, the operating conditions and context, the actual functioning of the designs in practice. Such a reconstruction ability should not only reproduce the physical reality, but also should encompass the knowledge, assumptions, decisions and safety critical issues which have been taken into account and assessed with respect to their acceptability. Such reconstruction ability should also incorporate the ability to recompose the socio-technical context and operating environment (Stoop & Dekker 2007). From such an investigator perspective, three kinds of systems designers should be supplied with a counterpart, each qualified with diagnostic and analytical skills from a technological/engineering design, organizational/managerial or governance/control perspective in order to cover the architecture of the overall socio-technical system. This is expressed in the DCP diagram (See fig 10).

Accident investigation cannot evade the challenge of dealing with 'cause' by switching to modelling accidents or computer supported probabilistic simulation of scenarios. Failure models and logical models which are applied in accident investigation are different from physical and mathematical models in Quantitative Risk Analysis and Bayesian Belief Networks. Accident investigations have to provide evidence, based on scientific methods, principles and knowledge (Stoop 1997, Sklet 2004). The Australian and Canadian safety boards have done groundbreaking work in this area (CTSB 2000, ATSB 2006). By their mandate accident investigation agencies do not have the option but to render advisory opinions to assist the resolution of disputes affecting life or property. They play a role of public safety assessor in a multi-actor environment in open decision making processes (Kahan 1998).

This recomposition of events in their systemic context is the domain of forensic sciences. Beyond their present judicial applications, forensic sciences could support investigations:

- into the ability of companies to deal with conflicting goals and sacrificial decision making issues,
- into the working conditions of operators which may degrade their cognitive or physical performance and
- into the effectiveness of safety management systems in ensuring corporate safety oversight.

In general, forensic sciences comprise of the science, methodology, professional practices and principles involved in diagnosing common types of accidents and failures.

Determination of causes of technical failure and gaining oversight over social systems performance requires:

- *familiarity with a broad range of disciplines*
- *ability to pursue several lines of investigation simultaneously.*

As such, it is an essential skill of accident investigators, engineers, researchers, inspectors and managers who are dealing with safety investigations.

Applying time as a diagnostic dimension

Investigating events deals with three lines of exploration:

- re-construction of the physical behavior of the system, according to the physical laws that were in force during the event regarding energy, forces, power, kinematics, dynamics
- re-enactment of decisions that have been made during the event by each of the operators and actors, based on their functions, responsibilities, tasks, skills, competences, perception, cognition and awareness aspects.
- re-composition of the event in its systemic context and operating environment, dealing with the structure, culture, content, context and history of the systems.

Investigating events deal with three time lines:

- the timeline of the event itself, looking upstream and downstream from the moment of impact as t=0, enabling establishment of the point of no return where the triggering event was set in motion and the event became inevitable
- the timeline of the systems life cycle history, starting from the initial design until the adaptations, modifications and operating conditions that were circumstantial during the event
- the timely transparency over an event should facilitate a quick recovery of a system to a stable state and safe operations. Prevention of similar, emerging properties should enable a system to return to its previous state by building in resilience, while on the longer term it might be necessary to change inherent properties by redesign, conceptual change and technological innovation.

Analyzing time as a systems dimension enables assessment of a discrepancy between the time necessary and the time available to recover at the bifurcation point where the system became safety critical.

6.4 the system

Four dimensions of a system

A problem can be approached from different viewpoints by making a number of ‘cross sections’ through a problem. If these cross sections are properly chosen, each cross section shows different ‘dimensions’ of the problem. Such a structured search is referred to as the ‘dimensions’ technique (Stoop 1990). The objective of the ‘dimensions’ technique is to establish a description of the problem in the context of a socio-technical system which makes a reference to the life-cycle, dynamics and structure, culture, context and content the relevant system dimensions.

The ‘dimensions’ are respectively:

- historical: this dimension provides insight into the development of the problem and the long-term development of its technical, organizational and social factors and hence, of the controllability of the problem in the context of the long term development of the system. This dimension covers the Context of a system
- life-cycle: from design, development, manufacturing, through us towards demolition. This dimension gives insight into the feed-forward and feed-back coupling of knowledge and expertise between system phases and knowledge about the criteria relevant for improvement and change. This dimension covers the Structure of a system

- process: this dimension gives insight into the ‘normal’ use of the system and describes the content of the processes that occur in ‘normal’ functioning. It gives insight into the tasks, activities, procedures, tools, equipment, operating environment, inputs and outputs of the system. This dimension covers the Content of a system
- culture: this dimension characterizes the system as occupational, transport. leisure or domestic. It gives insight into the (social) objectives, of the system, the role, positioning and functioning of stakeholders, their views, norms, values and codes of conduct. This dimension covers the Culture of a system.

The ‘dimensions’ technique collects data from normal as well as disturbed functioning, addressing all available performance indicators from intended and actual use and develops from broadly descriptive towards detailed explanatory. Data collection can be conducted by literature study, interviews, document analysis, on-site investigation and other forensic techniques.

The four dimensions are explored in parallel and should result in credible, plausible and verifiable description of the problem under scrutiny in its systemic environment.

A system life cycle approach: the DCP diagram

In order to integrate safety in design and operations, a new notion of vectoring safety through the systems landscape should be defined. Such a notion consists of three principal elements, being Design, Control and Practice (DCP). They can be interrelated along three dimensions, being a systems approach, a life cycle approach and a design approach. Together they constitute an integrated systems architecture prototype: the DCP diagram as depicted in fig 12.

A systems dimension defines three levels: the micro level of the user/operator, the meso level of organization and operational control and the macro level of institutional conditions.

The life cycle dimension defines a series of subsequent phases, being design, development, construction, operation and modification. At this dimension, the coordination of decision making among actors across the phases is crucial.

The design dimension identifies three principal phases in design, being goal –expressed by a program of requirements, concepts and principles-, function –expressed by design alternatives- and form, expressed by detailed design complying with standards and norms. At this dimension, the potential of technical innovation for new safety solutions is crucial.

The operational dimension. Eventually, only in practice safety is visible and actual consequences of accidents occur. At each of the other levels and phases however, separated in time or space, safety critical decisions have been made by different actors. The diagram demonstrates who, how, at which moment can contribute to safety and risk assessment

To manage consequences of new technology and innovation in transport systems engineering design, three principal lines are available:

- the Practice-Control line. Along this line, an upgrading in interventions takes place. The focus shifts from the performance of individual operators towards the meso level of organization and management in allocating resources, skills, operating procedures and responsibilities. At a macro governance level, rules, regulations and legislation, inspection, certification and governance oversight are addressed as safety enhancement opportunities.

- the Design-Control line. Along this line, decision making and safety assessment methods and standards should be elaborated, to facilitate coordination among stakeholders and actors, participating in major project developments. Several initiatives have already been taken such as safety impact assessment techniques, harmonization of standards by drafting EU Guidelines and Directives on specific topics such as tunnel safety, land use planning or external safety.

- the Design-Practice line. Engineering design methods for integration of safety in technological innovation are in their earliest phases of development. Historically, an impressive variety of design techniques is available. However, these instruments focus on specific industrial sectors and detailing levels of engineering design of components and are not always generically applicable across modes, disciplines or sectors.

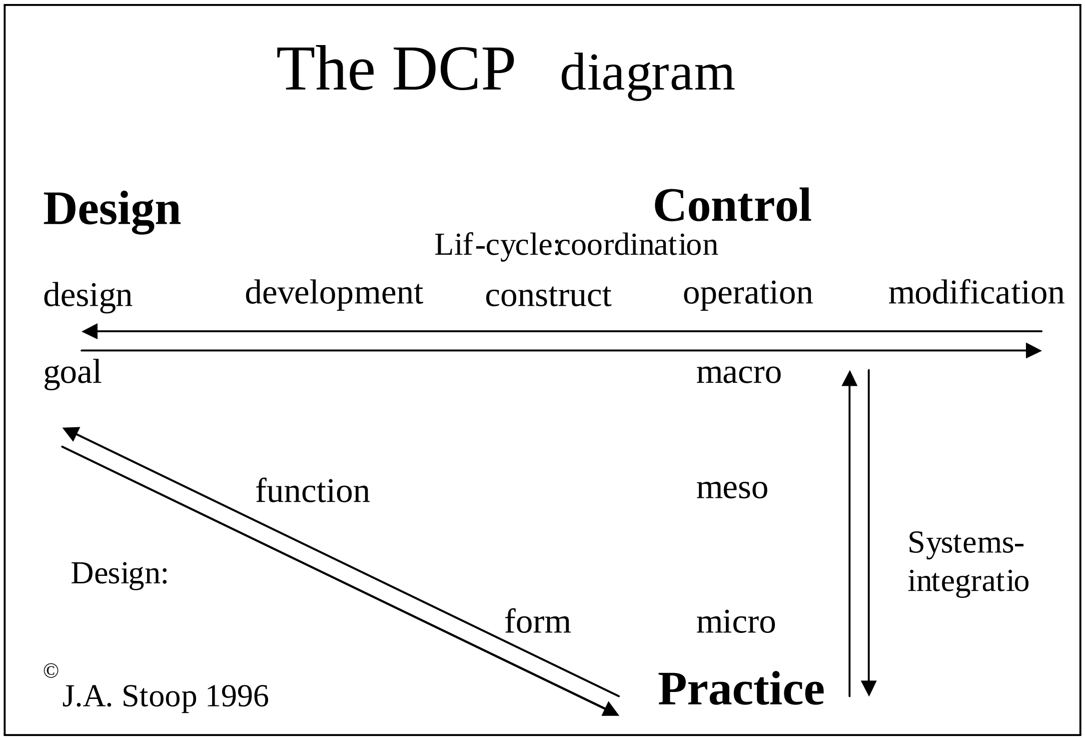


Fig 12: the DCP diagram

In order to design a coherent system and to maintain oversight over the system functioning, a system safety integrator role should be defined. During the design of complex transport systems, a dedicated responsibility should be allocated to assure continuous monitoring of the safety aspects along both lines during its design.

System performance levels

Business models and earning systems as incentives for efficiency versus thoroughness trade-offs are very powerful drivers for cost-efficient operations. In modern business concepts, calls for lean production, faster, cheaper and better performance are frequently heard.

With the introduction of New Economy principles in the transportation sector, three simultaneous developments have changed the drivers for cost-effective decision making. Changes in economic and logistic infrastructures, safety philosophy and selection mechanisms for preferential solutions have shifted from safety performance criteria towards exploitation, availability and cost-efficiency criteria. Cost-benefit considerations and environmental constraints in operations have become dominant. Instead of covering technical deficiencies by an array of technical provisions, a 'willingness to pay' and cost-effectiveness of solutions have become prevalent. Other arguments than safety have to been taken into account in decision making.

Differences in expertise are considered hindrances or even unjustifiable instruments to control the outcomes of a consensus process. Such an environment of 'participative policy making', assumes equality between parties and change the role of experts. Public private partnerships are favored as an answer to hierarchical ordered governmental projects on major infrastructural projects in tunneling, railways and aviation. Safety becomes a 'social construct' instead of an outcome of objective assessment based on professional experiences, quantifiable performance parameters and expert opinion. Such a 'new approach' in safety thinking shifts the focus towards prevention, flexibility, cost-benefit considerations , quantification of key performance indicators and institutional arrangements. This 'new' approach is a response to the inadequacy to provide substantive progress in conventional safety in the context of a 'new economy' context. Implicit assumptions are that the market should be best prepared to bear the risks and supply the knowledge, while a process approach should drive out substantive approaches. Private parties should not be disturbed by approval of their technical solutions, but should have their hands free to inform government about their selection of preferential solutions. Performance of a systems is reduced to measurable and quantifiable performance indicators. Safety is not such a parameter. Such a regime may reduce or improve the overall safety performance level of a system.

In comparing similar concepts, two options emerge:

- a low systems safety level, characterized as a earning system. In such a system, liability issues, blame and performance are pivotal. Willingness to pay and ALARA techniques prevail, while rule compliance and inspections are important control mechanisms. Safety is controlled at the organizational and company level.
- a high systems safety level, characterized as a learning system. Such a level is guaranteed by quality performance, transparency, communication and cooperation. Sharing responsibilities and information is essential for common learning and indirect cost are recognized. Responsibilities and roles are guaranteed by institutional arrangements.

In selecting either of such options as preferential, specific criteria should be available. Identifying systemic values in a multi-agent based environment has become a topic.

6.5 State space vectors: a value approach

Value Operations Methodology

Value Driven Design (VDD) is a methodology which promotes the use of a more complete value function as the objective function to be solved through optimisation, rather than using a more limited formulation typically related to some performance metric or through managing the process of meeting requirements. However, this principle can be extended to consider not only the value of today's basic economic drivers but also to incorporate the ultimate value for the customer and even society. Such ultimate values depend on who is implementing the Value Operations Methodology (VOM) that focuses on the ultimate value realised in through-life operation. Consequently, it is extremely well aligned to the problem of how to incorporate safety analysis into engineering and policy making decisions.

This has been incorporated into the fundamental VOM hypothesis as follows (Stoop and Van der Burg 2012):

> *'the true value of an engineering solution is subjective, temporal and of an inherently transient nature, and therefore engineering value analysis and optimisation is more meaningful if formulated as the evaluator's preference for one state over another as a function of the quantitative difference in a number of key value levers related to the operational realisation of the intrinsic value of the product, process or service being considered.'*

Consequently, this principle is further expressed in equation (1):

$$\Delta V\left(x_i,\ldots,x_n\right)=\sum_{i=1}^{N}\alpha_i\sum_{j=1}^{M}\omega_j\frac{\left(v(x_{ij})\right)_{end}}{\left(v(x_{ij})\right)_{start}}+\varepsilon_{ij} \qquad (1)$$

Where a change in value ΔV is caused by a change in a set of associated value levers x_i, when moving from some start-state to some new end-state. Each value lever of the set $i=1\ldots N$ has an associated scaling factor α_i and error ε_i and is in turn defined by a subset of lower level value parameters, x_{ij} for $j=1\ldots M$ and associated scaling factor ω_j , that describe the causal nature of each of each driver.

Consequently we can conclude that certain events have a contextual, cultural, content, structural and temporal dimension, whether at a component, sub-system, systems or system of systems level. Most importantly, rather than just stating safety factors, we now have a concept of real safety related events having an impact magnitude and a directional bias relative to the four dimensions of the model. The model suggests multi-vectorial design solution spaces which have meaning relative to the four dimensions of safety in terms of the contribution or impact within each dimension and the overall resulting orientation or direction of the safety issue being considered. Consequently, safety is significantly elevated from the very basic consideration of factor, to a new level where it is being quantified as a multi-dimensional quantity with a resulting orientation that defines the choice of the designer or operator relative to their values regarding safety. With reference to the Value Operations Methodology this leads us to the position where safety can be integrated into the general design approach of the air transport system according to an equation relating KPI to some delta value of the form in the equation (2):

$$\Delta V = \alpha_C(C_1/C_0)+ \alpha_U(U_1/U_0) + \alpha_M(M_1/M_0) + \alpha_E(E_1/E_0) + \alpha_P(P_1/P_0) + \alpha_S(S_1/S_0) + \varepsilon \qquad (2)$$

where Cost efficiency is represented by *C* (revenue/cost), Utilization by *U*, Maintainability by *M*, Environmental Quality by *E*, Passenger Satisfaction by *P*, Safety by *S* and finally including an error *ε*, consideration. Consequently, safety as a function of: safety=fn(context, culture, content, structure), can be characterised with the individual drivers associated with each dimension so that safety in its vectorial and most realistic form can be integrated into the overall integrated system of systems design solution spaces.

Simulation and prototyping

In making the transition from a linear safety intervention towards a dynamic safety intervention, the concept of critical load is applied. Accident scenarios can be considered critical loads on a system: once the critical load is applied, the system will fail if the loads is increased, exceeding the load capacity under the given operational conditions and system configurations.

In complex interventions, the focus is on safety critical events in a systems context rather than on isolated factors and generic aspects, as is the case with linear interventions. The re-composition of events takes place by identifying and synthesizing explanatory variables as ‘building blocks’ into scenarios in their specific operating environment and constraints. Such synthesizing is primarily evidence based. The redesign of the systems is conducted along the lines of engineering principles by generating design alternatives in the enlarged design space into the form of a limited set of prototypes. These prototypes contain a relocation and addition of functions, changing the morphology and configuration and incorporate additional actors and aspects. The testing of these prototypes is conducted by running scenario tests, definition of limit state loads and simulation of complex and dynamic systems in virtual reality. Analyzing system responses, before they are put into practice, are based on First Time Right and Zero Defect strategies. The responses of systems can be determined experimentally by a gradual enlargement of the disruptions which are inflicted upon the system, until oscillation and instability occur. Responses of systems may become visible by a gradual or sudden transition to another system state by passing a bifurcation point. After such a transition, the safety of the systems can be assessed according to the acceptability of the new safety integrity level, also in a technological sense.

Technology in itself contains many forms, incorporating invisible knowledge, notions, principles and decisions from previous life cycle phases. The physical appearance of a product and process does not disclose inherent properties, principles or interactions to end-users in their operational environment. Design decisions are frequently made under conditions of high uncertainty. Safety margins and design standards, identification of failure mechanisms, probability assessment, consequence analysis and identification of a design envelope should reduce the uncertainty again to an accepted level. Designers deal with optimizing performance and are not in a position to gain oversight into all uncertainties and unforeseen behavior of their designs. Such behavior however can be designed into their processes such as with the Japanese design philosophy of Limit State or

Critical State Design methodologies. Designers need an intellectual counterpart in assessing the safety and operational performance of their designs. Such a role is historically provided by accident investigators, operators and safety managers. To fulfill their role, their expertise should become available in the design process. This also implies that there is no objective 'truth' of a exclusively correct designed performance, because each of the actors will contribute their expertise, experience and competences in controlling their performance and providing trust in a safe, reliable, efficient and beneficial use of their resources under the operational conditions and constraints. In operating under such a variance of control options, a discrepancy may emerge between designed performance and actual, applied operational practices, revealing disclosed interrelations and operating restrictions as 'emergent' properties.

Consequently, such a collaborative engineering design methodology may provide a perspective for improving the safety performance of complex systems at a socio-technical level.

The potential for systems engineering design in providing safer solutions requires to:

- Identify inherent properties before they manifest themselves as emergent properties
- Deal with complexity and dynamics by focusing on functions rather than on factors
- First focus on design principles and system properties before optimizing performance
- Identify the working processes, operating envelope and operating conditions
- Identify the specific system state and operating configurations
- Introduce systems dynamics by synthesizing interrelations into accident scenarios
- Apply a proof of concept by testing solutions in a dynamic simulation environment

Therefore, it is necessary to:

- develop event scenarios separated from systems models
- develop prototypes of safer solutions
- create dedicated virtual systems models, representing their specific characteristics
- facilitate testing and validation in these virtual models, parallel to the real system.

Instead of identifying causes in order to establish the involvement of factors, actors, their motives and interrelations during the event, the operational performance of the *system as such* becomes relevant in the potential change towards a safer performance and the ability to learn from undesirable disruptions. Historically, safety oriented interventions have been focusing on elimination or mitigation of factors, actors or aspects, breaking up the sequence of events in order to prevent its recurrence. Instead of the event and the causal relation to the mishap, identifying systemic deficiencies and knowledge deficiencies become the critical issue in system change and knowledge development.

Synchronizing Eigen Values and Added Values

Design solution spaces can be discriminated as either linear and static –within the existing design envelope with a focus on performance in the operational phase- or complex and dynamic –beyond the design envelope, focusing on changing system properties and configuration-. The scope of such design interventions varies from an improved control over the event in case of a linear intervention to identification of multiple stable or unstable systems states and transition phases across these states.

Complex systems modeling takes the form of representation by *system state vectors,* expressed by five primary systems dimensions –culture, structure, contents, context and time-, each with their own characteristic key performance indicators and metric values.

Similar to such a system state vector, an *event vector* should be identified expressed by its own characteristics such as hazards, actors, factors, aspects, causal relations, operating variance, interactions and operating conditions. Navigating such an event vector through the systems operating envelope indicates proximity to operating limits and a drift into failure. The challenge in optimizing safe solutions is the synchronization of these two vectors by *transforming the event vector problem space into systems vector solution spaces* within the boundaries of the available engineering design solution space. Such synchronization requires transparency over the various system state transitions, as well as a consequence assessment of the residual risk and side effects that remains after a transition. Inherent to multi-dimensional systems, a proximity to a certain optimum may be identified. It is very likely that more than a single optimum exists for the multitude of dimensions and functions that are identified in the performance requirements.

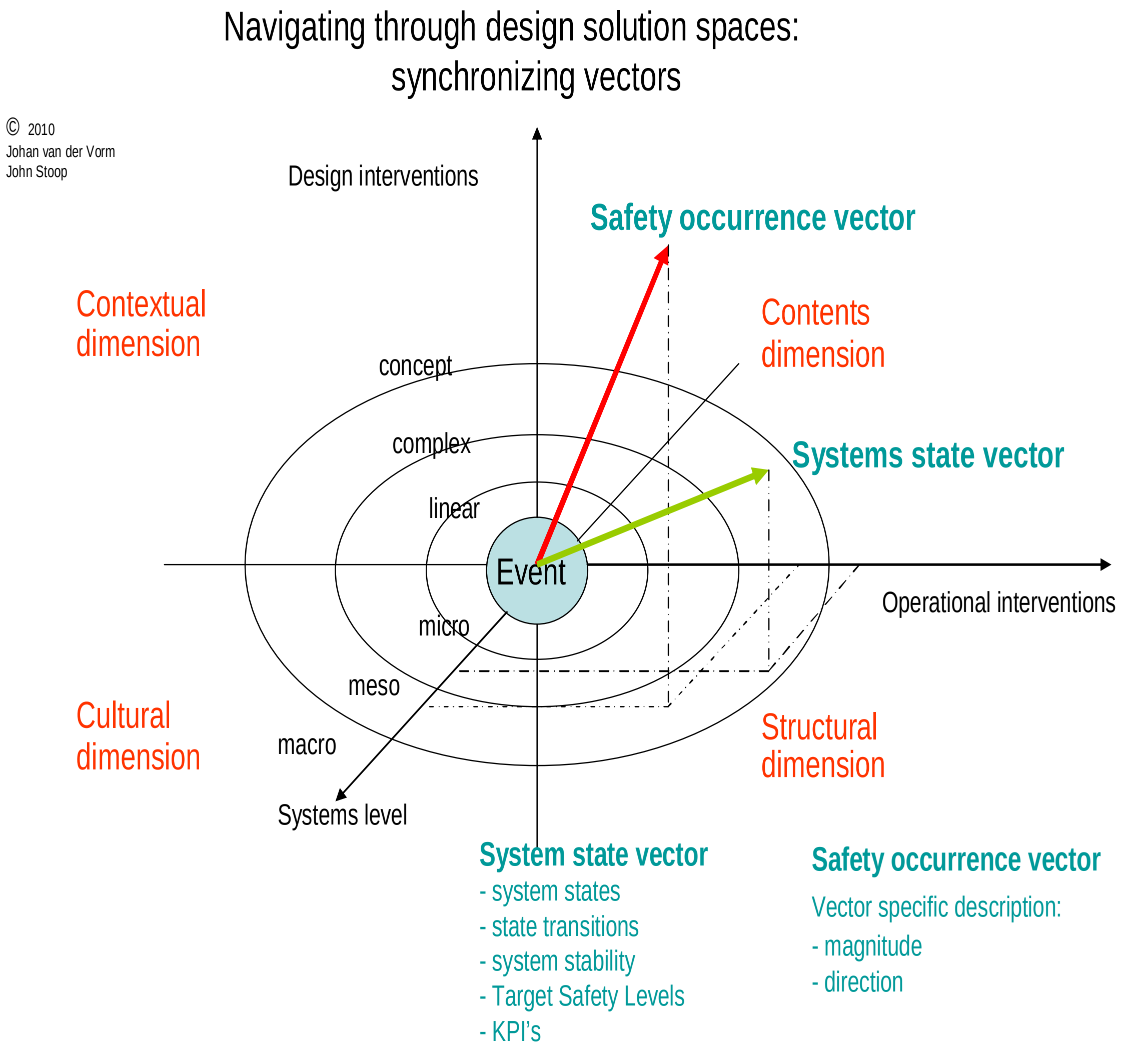


Figure 13: Multivectorial safety design solution spaces

In order to facilitate such synchronization, the Eigen Values of the event vector and system vector should be established to avoid oscillation and resonance (fig 13). Analyzing the potential systems responses is supported by testing the solutions in a virtual design environment by simulation and serious gaming techniques before the changes are implemented in the real world. By exposing the redesigned systems to the original ultimate load –the event scenario- the support for safety enhancement in terms of commitment for change, acceptance of the residual risk and feasibility for engineering design improvements are tested and validated. To this purpose, new scientific notions (such as Value Engineering, Knowledge Based Engineering and Resilience engineering) focusing on system properties such as resilience, reliability, redundancy, recovery, reliance, reconfiguration, rescue and emergency handling and can be assessed for their applicability during the (re-)design.

Vector transformation towards Cartesian coordinates

The use of multivectorial solution spaces is the common language for communication in the mathematical and engineering design community. Value operations, multi parameter optimization, dynamic modelling and simulation have become practical applicable through a wide variety of software applications. Computational Fluid Dynamics, Finite Element Methods, optimization algorithms are widely used in practice for research and education purposes. Although through such vectorial representations a state/space modeled advanced engineering of technically complex and sophisticated designs becomes feasible, their application requires qualified skills and experience in both modeling and software domains.

Therefore, this representation has to be transformed into notions that are accessible for practitioners, operators and investigators in a variety of industrial sectors such as energy production, process industry and transport.

To this purpose, the three dimensions of the ESReDA Cube –aspects of operations, depth of learning and stakeholders effected- are considered the X-Y-Z transformation of the vector representations of values, variables and attributes of the r(t), Ω(t) and θ(t) coordinates in the Polar system into the Cartesian coordinate system. .

The dimension of time –representing the dynamics of operational processes and systems development- is considered as the sequential/temporal solution algorithm which is required to enable the coordinate transition from one systems state and space to another. It is expressed by manipulating the ESReDA Cube by rotating surfaces and colors towards a new desirable configuration.

Such a rotation of the Cube indicates which changes in what direction and in which order should be achieved to accomplish the desired change in state and space modeling .

The ability to manipulate this transformation is submitted to internal control laws and rules of engagement of the system under scrutiny.

In this respect, the analogy with the Rubik Cube becomes clear by the fact that manipulating the Rubik Cube is also submitted to mathematical logic, sequencing of necessary steps in the solution process and solution algorithms. The ways to achieve a satisfactory final result are almost infinite. In this respect, an analogy with infinite possible design solutions serves the use of the ESReDA Cube as a metaphor.

Using a Cartesian instead of a Polar coordinate system is not a matter of being ‘old fashioned’ or applying an obsolete ‘Newtonian’ perspective. Each of the systems has its own origin and is appropriate for application in specific domains such as ranking of multiple variables, navigation and describing functions in a multi variable space and complies with the language of its users. Transformation between Cartesian and Polar coordinates follows strict rules of mathematics, where preferences for one or another system is defined by the usefulness of their applications and complexity of the calculations required to describe the issues.

In general, Cartesian coordinates serve purposes of ranking, scaling and graphical representations of functional relations between variables, while Polar coordinates serve determination of distances, positioning, navigation and change. Defining curves and movements of bodies in a three dimensional space can be relatively simple and intuitively modeled using Polar coordinates. Cartesian coordinates are applicable as abstractions of problem representations in a n-dimensional space superimposed on a problem application. To model such problems, coordinates can be expressed in units of distances, assigning a specific spatial origin and a definition of the axis for directional cues in the n-dimensional space. Agreement between experts involved in the problem space on the definition of coordinates is crucial for novel applications and the creation of a novel coordinate system. In non-engineering and non-technological applications, units do not necessarily have identical units over the axis or numerical values, but may be composed of qualitative units. Although higher dimensional spaces are hard to visualize, the mathematics involved in their calculations can be extended relatively easily and represented in a Polar coordinate system with vectorial representations.

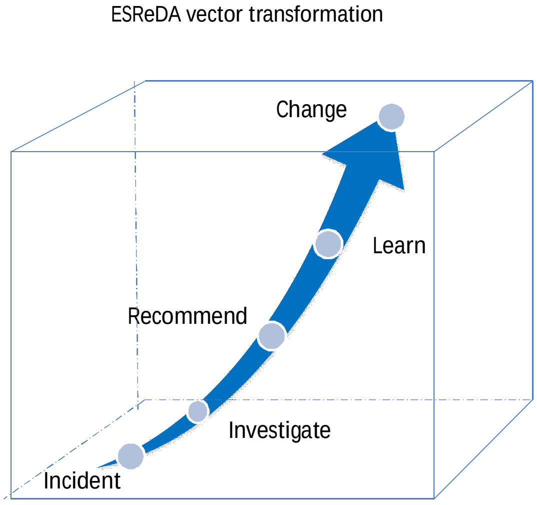


Fig 14: vector transformation

In this perspective, the three X-Y-Z axis for the dimensions of Aspects of Operations (X), Depth of Learning (Y) and Stakeholders Effected (Z) can be expressed respectively in qualitative units such as Structure, Culture, Content and Context along the X-axis, Optimize, Adapt and Innovate along the Y-axis and Micro, Meso and Macro along the Z-axis (See fig 14).

Transition rules between the Cartesian representation of the ESReDA Cube and the Polar coordinates of the value vector optimization are gives by algebraic equations, linking the two systems (fig 15).

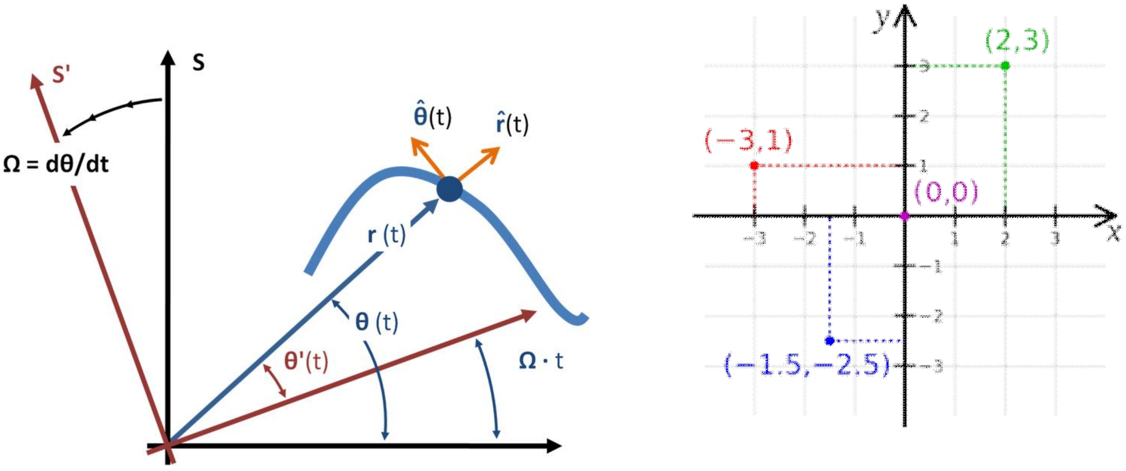


Fig 15: Transformation from Polar to Cartesian coordinates

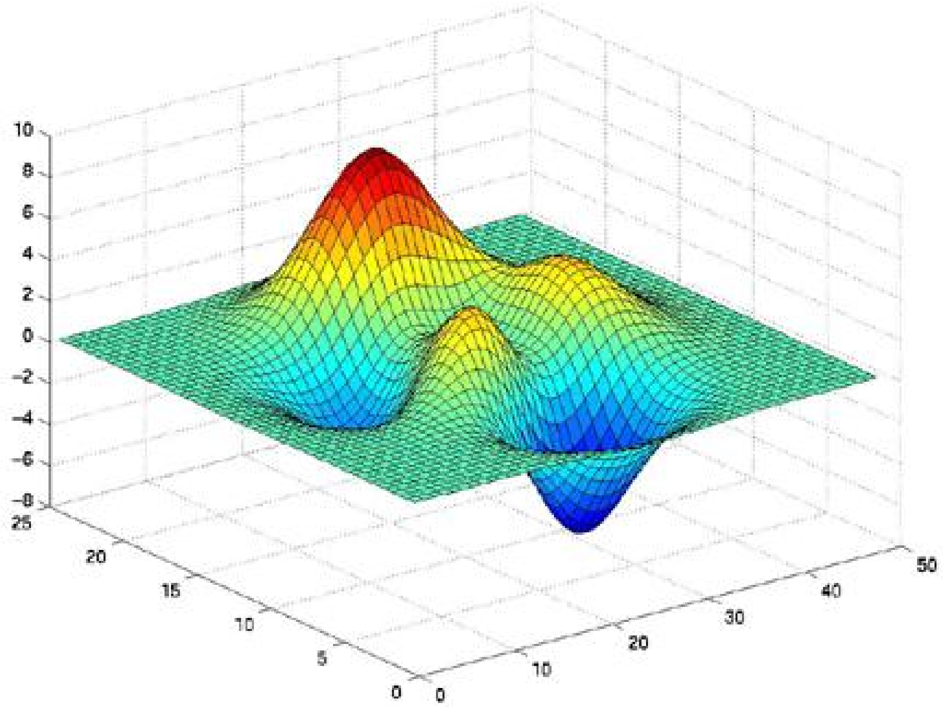


Fig 16: Multi dimensional optimization surfaces

In particular for navigating the safety vectors through a value landscape, polar coordinates and vector connotation are appropriate in answering questions such as; where are we, where do we want to go, how to get from problem spaces to solution spaces. Managing the required change by informed decision making can be supported by the shift from a factor notion towards a vector notion (Stoop and Van den Burg 2012). Eventually, optima can be represented by the use of multi-dimensional optimization surfaces such as available through Multi-Disciplinary Optimization software applications (fig 16).

6.6 system state and mental mode transitions

System state transitions in complex systems can be described by the Cynefin model of Snowden (2007). In 1999 Snowden introduced a framework for understanding that different situations and different system states required different responses. In his model, Snowdon discriminated various states of systems in relative disorder:

- simple: standard operating practices prevail, best practices are legitimated, while cause and effect relations are repeatable, perceivable and predictable
- complicated: analytical and reductionist approaches prevail, adhering to preplanned scenarios, while cause and effect are separated over time
- complex: patterns and trends are to be identified, complex and adaptive actions are required, while cause and effect are only coherent in retrospect
- chaotic: crisis intervention is required, focusing on stabilizing actual system behaviour, while no cause and effect relations are perceivable and outcomes are undefined.

These states are to be recognized by their characteristics as depicted in fig 17.

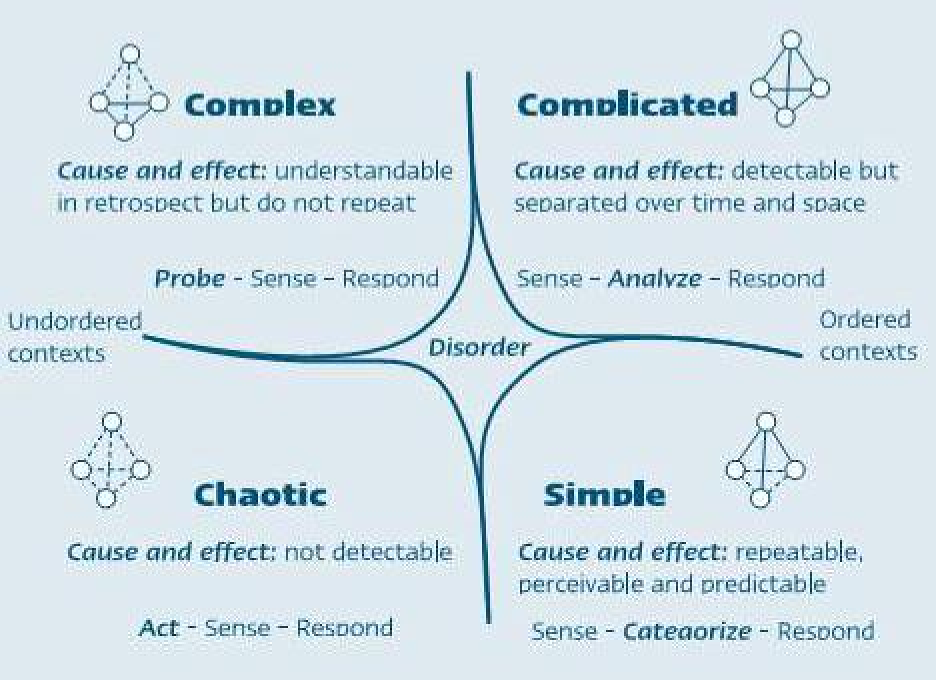

Complex
Cause and effect: understandable in retrospect but do not repeat
Probe - Sense - Respond
Complicated
Cause and effect: detectable but separated over time and space
Sense - Analyze - Respond
Undordered contexts
Disorder
Ordered contexts
Chaotic
Cause and effect: not detectable
Act - Sense - Respond
Simple
Cause and effect: repeatable, perceivable and predictable
Sense - Categorize - Respond

Fig 17: The Cynefin model

Solutions how to deal with these system states may be detrimental if they are applied in an incorrect context. Their internal dynamics may produce unforeseen and undesirable consequences, either for the best or for the worst. In avoiding an undefined performance with ‘emergent’ responses, a correct approach should be applied, compliant with the actual, corresponding state of the system. In such an intervention in the chaotic state, it is important to recover from this ‘chaotic’ state and regain control over the system performance, returning to a ‘normal’ state of complicated, but manageable conditions.

During the analysis of accidents, the ability of a crew or operator to successfully return from the chaotic state to a complicated state is investigated. Such a recovery however cannot be achieved by offering an operator resources of a ‘normal’ nature, because the transition from chaos to normal has to transit either the ‘complex’ or ‘simple’ state. The normal, complicated state is not directly accessible for operators in a system which is in the chaotic state. Two potential recovery strategies are available: either reducing complexity by diagnosis, communication and cooperation or regaining control beyond the level of reflexes and intuitive responses is required.

Recovery through the simple state requires a correct mind-set of the operator to avoid an intuitive, reflex based ‘fight, flight or freeze’ type of response. Applying Standard Operating Procedures of a simple nature alone does not work with unstable and chaotically perceived states in case of unanticipated and unusual events. Such an approach requires intrasubjective adaptation by an autogenic feedback or neurofeedback training nature in order to achieve individual reflection on the subject’s mental mode (Cowings and Toscano 1993, Den Hertog 1994, Gorter and Jaeger 2014).

Recovery through the complex system state requires providing transparency and oversight over the situation by creating intersubjective adaptation by sharing information, communication and cooperation, mobilizing all available knowledge about the actual behaviour of the system. A return to safe operations is possible through creating a timely and shared mental models across all actors (Rochlin 1999, Mohrmann 2013).

In practice, it is not possible for an operator to arbitrarily or objectively select one of these two strategies to transit the simple or complex system state. A predetermined transition management strategy is required by a proactive and precautionary support of the operator by a timely provision of a crew with skills and resources (fig 18). Such a transition management strategy eliminates the notion of ‘human error’. The concept of ‘human error’ is replaced by a specific system state and specific mental mode approach, taking into account the operating envelope and operating conditions of the system.

Human error is replaced by a concept of:
discrepancy between the operator and the system if the operators mental mode and the system state are not synchronized.

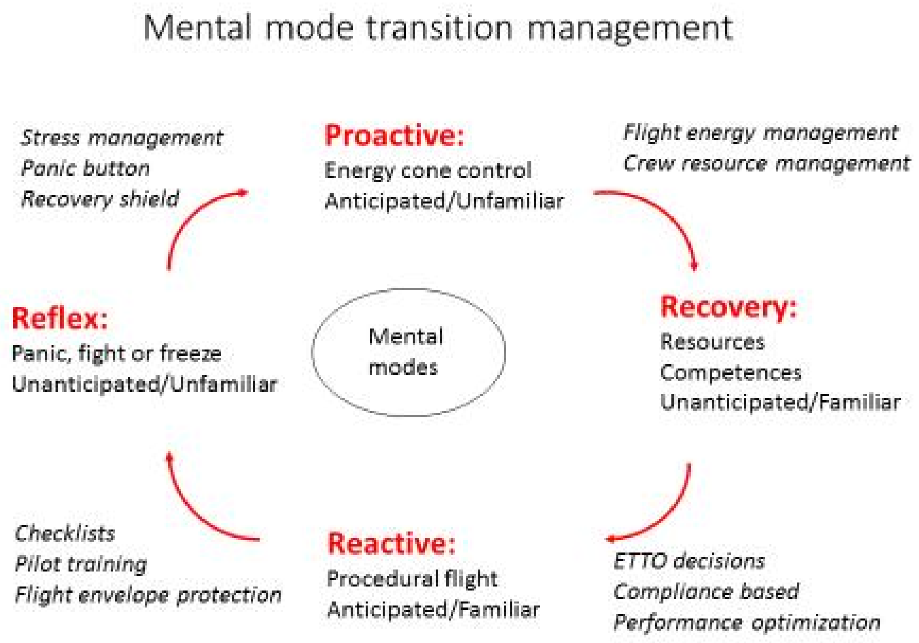


Fig 18: Regaining mental oversight and transparency

Cognitive resistance to change

During a mission, humans may encounter various mental modes that cannot be changed spontaneously and in a random order. The selection of a mental mode is consciously or unconsciously influenced by automation actions, crew expectations and either a high or low work load. Changing a mental mode may be difficult due to a natural inclination to resist change and remain in the equilibrium basin of the system. A transition to another equilibrium basin in the system depends on strong external stimuli and conscious reflection of the need to change.

The identification of influences of unconscious and intuitive processes, group dynamics and environmental interferences enabled the scientific research community in the field of psychology to identify three explicit basic modes of operandi:

- rational decision making
- emotional decision making
- social decision making.

In developing an affective model of cognitive resistance to change, De Boer identified two modes of operandi in the interference between emotional and social decision making: unconscious resistance to change and conscious stubbornness (De Boer 2012). These phenomena represent a switch between two cognitive systems of thinking -a system 1 of rapid intuitive interferences and a system 2 of slower deliberate interferences- resenting a change in the mental model and type of reasoning of an operator (Kahneman 2011). Such a resistance may create a cognitive lockup in supervisory control tasks and eventually can create disruptions in the performance, causing accidents. Such a resistance to change is related to an individual's beliefs, persistence, change blindness, cognitive mismatch, fixation and mental lockup when dealing with contradicting signals

(De Boer 2012). Resistance to change may block sharing of mental models in a team, hindering a common understanding of a situation. It may have either a positive or negative effect on performance, depending on the situation. As such, this phenomenon bears no normative judgement on the correctness of an outcome or qualification by labelling the decision and subsequent action as 'human error' if the consequences do not comply with the normative expected outcome. Cognitive resistance as a behavioural component, its indecisive positive or negative consequence and situation dependency has additional positive properties such as extraversion and agreeableness.

Resistance to change may have advantages as well: vigilance, danger avoidance, stimuli detection, achievement and rapid reflection may induce less automatic, intuitive behaviour and enhance analytic competences. Such properties may be characterized as attributes of Good Airmanship or Good Seamanship. The interaction between emotion and reflection may improve the overall performance and the sensitivity to types and intensities of contradicting stimuli. Applying cognitive resistance and conscious stubbornness can be instrumental as a deliberate strategy to cope with unsubstantiated desires to change and disruptions of thought processes (De Boer 2012).

In analogy with the notion of functional resonance and system oscillation of Hollnagel, a notion of *'cognitive dissonance'* could be introduced to indicate the discrepancy between the actual mental mode of an operator and the required mental mode in order to comply with the system state in which specific operating tasks have to be fulfilled (Stoop and Van Kleef 2014).

7. Summary

In investigating accidents and incidents several notions, theories and principles on composing events and modeling systems have to be taken into account. In order to gain transparency over investigation processes, their results and conclusions, an understanding of often implicit backgrounds, rationales and paradigms has to be taken into account. Such understanding provides structure to the investigation processes. Otherwise, an effective intervention in the characteristics of the events and the systemic properties of the operating environment is not possible and learning abilities are hampered.

7.1 theoretical notions and principles

To investigate accidents and incidents the investigator needs to understand that :

- *systems are decomposed* along lines of:
  = four dimensions: structure, culture, content and context
  = three systems levels of hierarchical organization and interrelations: micro, meso and macro
  = three orders of change capacity: optimization of standing operating procedures and practices, adaptations and reconfiguration of formal procedures and practices and innovation by redesign and conceptual change
  = a Design, Control and Practice landscape representation of the systems history, dynamics and complexity
- *events are recomposed* along lines of spatial and temporal relations between occurrences under specific operational conditions in a specific systemic context represented in flow diagrams. The temporal dimension represents the dynamic sequencing of events in a systemic context into a scenario
- *events are described along lines of a narrative scenario*, containing a communicable inventory of facts and findings that may contribute to the sequence of event, dealing with factors, actors, aspects and assessment of safety criticality of these contributing entities, decisions and relations as 'building blocks' for further input in the investigation process
- *communication, transparency and oversight* is achieved by visualization of the event process in a time related flow chart, characterized by its critical path and correctable action potential for each of the typical events that can be derived from a collection of frequent and similar events that occur during normal operations. Communication with the outside world is represented by applying a metaphor and a structured solution space description: the ESReDA cube
- *several shifts occur*: in applying a method of investigation for diagnosing complex systems, a shift occurs from performance to properties and from causal factors to systemic and knowledge deficiencies. New areas of concern emerge from investigation findings: after describing the event, based on observable physical facts and findings, documented information naturalistic decision making processes and clarification of mental models, such an accident scenario is diagnosed through adequate modeling in its systems and

operational context. This perspective facilitates a shift from descriptive variables towards explanatory variables and eventually into change variables

- *transitions in logic thinking take place*: through the various phases of the investigation process, various forms of logic thinking are applied. Selection of each of these forms depends on the phase of the problem solving cycle, dealing with problem recognition and definition, fact finding, analysis, generation of solution spaces, selection of preferential options and implementation.

7.2 putting theory into practice

In making the transition from theory to practice, several tools are required, each specifically dealing with their own goals regarding communication, analysis and change process management.

1. common language for cross disciplinary and multi-actor cooperation requires common understanding and application of shared notions in dealing with complexity: in the ESReDA approach the notion of Eigen Values as vectors are introduced. Such a notion however also requires the definition of three new dimensions –operations, learning and stakeholders- and their scaling along the axis as defined in the ESReDA Cube. Such publicly accessible dimensions should be transformable to vector optimization methodologies
2. communication metaphors serve the purpose of providing appealing analogies with familiar notions for the lay public and non-experts in order to communicate with experts using scientific methods; the Rubik Cube analogy provides an interesting metaphor for simulating and manipulating simultaneously multiple dimensions in decision making
3. tools are required to enable the transitions between the various phases of the investigation process from fact-finding, to analysis, towards change. Such transitions necessitate the distinction of 2 types of cubes for descriptive variables and for change variables, which subsequently facilitate
   = structuring the problem: allocating descriptive variables to their position in the cube
   = solving the problem: positioning change variables and change drivers in the cube
4. Translating notions about state/space vectors and synchronizing event and systems vectors require a transition from Cartesian coordinates to Polar coordinates to incorporate safety value vectors into the optimization of the system, taking into account all relevant values of the system under scrutiny. However, such transition algorithms have not yet been developed.
5. Combining single event cubes with classes of events: provide a repository of solutions for similar events by adding up solution space cubes for each of the events into a class of problem spaces and solution spaces. By classifying problem spaces into categories, systemic deficiencies can be identified and solution spaces can be developed.

7.3 investigations: a firewall against failure

Essentially, in complex and dynamic systems, two principal fire walls can be identified between the various transitions across system life cycle phases:

1. Certification represents a firewall between the engineering design phase and the operational phase. This firewall assesses whether the intended performance, as formulated in the program of requirements during the assignment, actually can be

exercised under specified and foreseen operational conditions and constraints, for a defined series of missions and system states. Certification identifies compliance with the agreed operating envelope and allocates responsibilities and efficient performance to the integral system.

2. Investigation represents a firewall between operations and system design, manufacturing, maintenance, repair, overhaul and modification. This firewall assesses whether a continuation of safe and efficient performance can be maintained, taking into account the assumptions and limitations in knowledge and operational experiences as they manifest themselves in the actual operating environment, culture and context in which the system has to perform. Investigation identifies the systemic and knowledge deficiencies that create unacceptable degradations of system performance and system states.

Applying these two firewalls implies a series of interventions that are possible in the actual operating mode. Such interventions will depend on the nature, severity and extent of the deficiency, available resources and efficiency and effectiveness considerations. Urgent safety issues will require immediate actions, such as grounding of aircrafts, closing of airspace or airport facilities in case of imminent danger as has been shown in the Boeing 787 case on the Lithium ion battery fire risks. To this purpose, the precautionary principle is widely applied, such as in aviation where a pilot may be forced to repeat a landing procedure by making a go around after a missed approach or unstabilized approach. Such safety measures are characterized as optimization of current procedures. A second level of intervention is characterized as adaptation of procedures such as by introducing more stringent criteria for go around procedures with cross wind restrictions, touch down area overshoots or glide slope deviations. A third level of intervention is to innovate the landing phase process by introduction of new technology, new pilot training concepts or continuous descent approach strategies.

Investigations close the feedback loop in the product, process and system life cycle (fig 19).

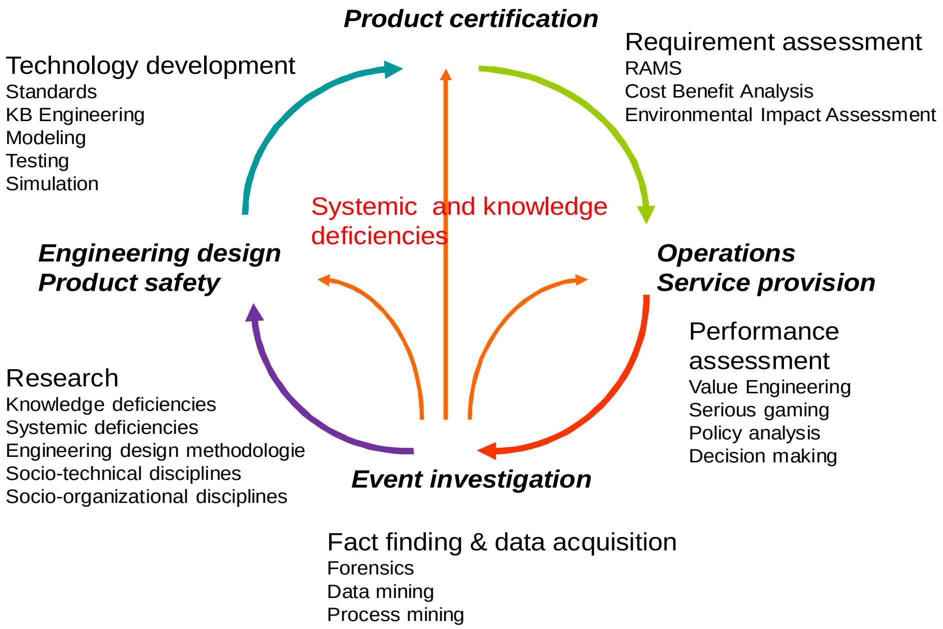


Fig 19: The two principal fire walls

7.4 a change of mindset

Technological systems have become more complex and less tractable, posing new challenges by the introduction of ICT and global network configurations. Those new challenges have induced new safety notions. Technological systems evolved into socio-technical systems, becoming less tractable and adaptive. This encouraged the development of resilience engineering, because traditional reliability and analytical methods could no longer cope with the problems introduced by those systems. New areas of scientific research are explored, such as the intuitive, emotional and social dimensions of human interactions of operators with their environment. Methodological issues have been raised in establishing a safety investigation methodology, based on a systems engineering design perspective. Moving from factor to vector, from performance to properties and transitioning from descriptive to explanatory and change variables is part of checking a complex and dynamic reality and achieving a safe operating environment. Applying new metaphors opens up new perspectives for communication on understanding of complex events.

A transition towards new notions on human performance, systems engineering design and safety investigation methodology and introducing a new metaphor for communication also requires coping with resistance to change. Cognitive resistance does not restrict itself to operators, it also covers experts in other disciplines than pilots, such as researchers, designers and investigators.

Such a transition should be based on highly reliable, credible, tractable, plausible and encompassing data: facts and findings are the start of the investigation process.

Safety investigations are evidence and case based, providing raw data for further exploration, interpretation and deliberations. a timely transparency in the factual functioning of socio-technical systems.

For investigators, applying forensic reasoning and the ability to think along multiple lines of reasoning, including abduction and construction provides a starting point in the investigation process:

forensic sciences comprise of the science, methodology, professional practices and engineering principles involved in diagnosing common types of accidents and failures.

The determination of the causes of failures require:

- familiarity with a broad range of disciplines and
- the ability to pursue several lines of investigation simultaneously.

The objective of the investigation is:

- to render advisory opinions to assist the resolution of disputes
- affecting life or property.

In such a perspective, investigations may play a role of problem integrator.

Cognitive resistance is a common human property. Thinking along lines of safety investigations and system change may not only overcome one's own resistance to change. It may provide a mind-set that enables us to face the unanticipated with an open and impartial mind and to make unbiased observations. It also may create a safer world.

Hora est.

References

Almeida I. and Johnson C. (2011) Extending the borders of Accident investigation: Applying Novel Analysis Techniques to the Loss of the Brazilian Space Programme's Launch vehicle VLS-1 V03. pp1-25

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

Anderson J. (1999) Aircraft performance and design. University of Maryland. WCB McGraw-Hill

Arslanian J.P. (2012) Key note Paul-Louis Arslanian as recipient of the Jerome Lederer Award. ISASI Salt Lake City, ISASI Forum, Oct-Dec 2011.

ATSB (2006) Manual accident investigation, Australian Transport Safety Bureau

ATSB. (2010) In-flight engine failure - Qantas, Airbus A380, VH-OQA, overhead Batam Island, Indonesia, 4 November 2010 Australian transport Safety Bureau (2010)

BEA. (2012) Final report on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France Flight AF447 Rio de Janeiro – Paris. Bureau d'Enquete et Analyses pour la securite de l'aviation civile. Paris, July 2012

Benner L. (2010) Transforming Experience Data into performance Improvement. Open Access Document, http://creativecommons.org/licenses/by-nc-nd/3.0

Benner L. (2010) Accident data for the Sematic Web. Safety Science 2010 doi:10.1016/j.ssci.2009.12.013 Article in press

Benner L. (1975) Accident theory and accident investigation. Proceedings of the Society of Air safety Investigators Annual Seminar, Ottawa, Canada, 7-9 October 1975.

Benner L. (1980) Accident investigations – A case for New Perceptions and Methodologies. Society of Automotive Engineers INC SP-461, Febr 1980

Benner L. (1985) Rating Accident Models and Investigation Methodologies. Journal of safety research, Vol. 16, pp. 105-126, 1985

Benner L. (1996) Accident Investigations: a case for new perceptions and methodologies. National Transportation Safety Board. Washington, USA. The Investigation Process Research Resource Site.

Benner L. (2009) Five Accident perceptions: Their implications for Accident Investigators. Journal of System Safety, pp 17-23September-October 2009.

Benner L. (2013)

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

Bird, J., and Di Paolo, E. A., (2008) Gordon Pask and his Maverick Machines. Pp 185-211. In P. Husbands, M. Wheeler, O. Holland (Eds), The Mechanization of Mind in History, Cambridge, MA: MIT Press.

Bor R. and Hubbard T. (2006) Aviation Mental Health. Psychological implications for air transport. Ashgate

Carper K. (1989) Forensic engineering. CRC Press, First Edition 1989

Cowings P. and Toscano W. (1993) Autogenic-Feedback Training (AFT) as a Preventive method for Space Motion Sickness: Background and Experimental Design. NASA Technical memorandum 108780, August 1993

Cross N., (1989) Engineering Design Methods. John Wiley & Sons.

CTSB (2000) ISIM methodology. Transport Safety Board of Canada, 2000

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

Den Hertog R. (1994) Human Factor Issues from other Aircraft contributing to incidents & accidents. 47th Annual International Air Safety Seminar Flight Safety Foundation. Lisbon, Portugal, Oct 31-Nov 3 1994

Den Hertog R. (1999) Safety starts at the Manufacturer. Lecture presented to the Netherlands Association of Aeronautical Engineers (NVvL) on May 27, 1999

Dekker S. and Hollnagel E. (2004) Human factors and folk models. Cogn. Tech. Work (2004) 6: 79-86

Dekker S. (2005) Ten question about Human Error. A new View of Human Factors and System safety. Taylor and Francis e-Library 2008

Dekker S. (2006) The Field Guide to Understanding Human Error. Ashgate.

Duffey R. and Saull J. (2008) Managing Risk. The human element. Wiley 2008

Dym C. and Little P. (2004) Engineering design. A project based introduction. Second Edition. Wiley International Edition.

Edwards E. (1972) Man and Machine: systems for safety. Loughborough University of Technology. United Kingdom

Eisner H. and Stoop J.A. (1992) Incorporating safety in the Channel Tunnel design. Safety Science, vol 15, no.2, July 1992.

ESReDA (2005) Roed-Larsen S., Stoop J.A. and Funnemark E. 2005. Shaping public safety investigations of accidents in Europe. An ESReDA Working Group Report. DNV, Oslo, February 2005

ESReDA (2009) Guidelines for Safety Investigations of Accidents. ESReDA Working group on Accident Investigation. Oslo, June 2009

ETSC (2001) Transport accident and incident investigations in the European Union. European Transport Safety Council. Brussels 2001

ETSC (2005) Europe and its road safety vision – how far to zero? The 7th European Transport Safety Lecture. European Transport Safety Council, Brussels, 2005

European Union (2011) Flight Path 2050. Europe's vision for aviation. Report of the High Level Group on Aviation Research. European Union 2001.

Eurotunnel (1994) The Channel Tunnel, a safety case. Channel Tunnel Publications. Langton green, Tunbridge Wells, Kent, UK

Evers J., Bovy P., De Kroes J., Sommerhalder R. en Thissen W. (1994)
Transport, Infrastructuur en Logistiek: een proeve van een integrerend onderzoeksprogramma. Onderzoeksschool voor Transport, Infrastructuur en Logistiek. Technische Universiteit Delft

FAA (2014). ADS-B benefits are limited due to a lack of advanced capabilities and delays in user equipage. Office of Inspector General, Audit Report AV-2014-105, September 11, 2014. Federal Aviation Administration, USA

Fahlbruch B. and Wilpert B. (2002) System Safety Challenges And Pitfalls Of Intervention. Pergamon

Faleiro L. and Lambregts A. (1999) Analysis and tuning of a 'Total Energy Control System' control law using Eigenstructure assignment. Aerospace Science and technology, 1999, no 3, 127-140

FSI (2013) Flight safety information, No 6, April 2013. www.fsinfo.org

Gorter D. and C. Jaeger (2014) Project Samurai Pilot. Cognitive Enhancement of Airline Pilots with Biofeedback Training. Delft University of Technology, Faculty of Aerospace Engineering, and University of Leiden, Faculty of Psychology. August 2014

Guzzetti J. (2014) Safety Data: The Agony and the Ecstasy of Their Use. ISASI Forum, Jan-March 2014, pp 6-10

Hale A.R. (2005) In: Review of the scientific research program of the Safety Science Group, Faculty of Technology, Policy and management, Delft University of Technology.

Hale A.R. (2006) Method in your madness: System in your safety. Valedictory Lecture Delft University of Technology, Delft, September 2006.

Harris D. (2011) Human performance on the flight deck. Ashgate 2011

Hendrickx L. (1991) How versus how often. The role of scenario information and frequency information in risk judgment and risky decision making. Doctoral Thesis Rijksuniversiteit Groningen, the Netherlands.

Hersman D. (2012) Ensuring safety in aviation's second century. ISASI Forum, Jan-March 2012

Hersman D. (2012) Improving Aviation Safety: Reactive, Predictive, Preventive.
Aero Club of Washington, DC, Washington, DC - September 18, 2012

Hollnagel E. and Woods D. (2006) Resilience engineering: Concepts and Precepts. Aldershot: Ashgate Publishing Ltd 2006

Hollnagel E., Pieri J. & Rigaud F. (2008) Proceedings of the Third Resilience Engineering Symposium. October 28-30, 2008 Antibes-Juan-les- Pins. MINES, Paris, Collection Sciences Economiques. France

Hollnagel E. (2009) The ETTO Principle: Efficiency-Thoroughness Trade-off. Why things that go right sometimes go wrong. Ashgate

Hollnagel E. (2012) FRAM: the functional resonance analysis method. Modelling complex socio-technical systems. Ashgate 2012

Hoppe E.A. (2011) Ethical issues in Aviation. Ashgate

Hudson P. (2009) Causal Model for Air Transport safety. Final report. Directorate General of Civil Aviation and Maritime Affairs. Ministry of Transport, Public Works and Water Management, the Netherlands.

Hudson P. (2010) Safety Science: it's not rocket science, it's much harder. Inaugural lecture Delft University of Technology, 24 September 2010.

IDAIP (2001) Main Points Memorandum on independent accident investigation. Independent Disaster and Accident Investigation Project. Ministry of the Interior and Kingdom Relations. The Netherlands

Johnson C. (2003) Failure in Safety-Critical Systems: A Handbook of Accident and Incident Reporting, University of Glasgow Press, Glasgow, Scotland, October 2003.

Johnson C. (2012) The future of sub-orbital and orbital space accident investigation. Dept of Computing Science, University of Glasgow, Scotland

Kahan J. (1998) Safety Board Methodology. In: Second World Congress on Safety of Transportation. Editors Hengst S., Smit K. and Stoop J. Delft University Press 1999.

Kahneman D. (2011) Thinking, F ast and Slow. Penguin Books.

Katsakiori P., Sakellaropoulos G. & Manatakis E. (2008). Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models. Paper accepted for Safety Science, November 2008

Kletz T. (1991) An engineer's view of human error. Second Edition. Institution of Chemical Engineers. Warwickshire, UK

Klir G. (1987) The role of methodological paradigms in systems design Dept. of System 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 on fuzzy systems. Vol 2, no1 Febr 1994.

Koningsveld H. (1989) De samenhang is ons ontglipt. Fragmentatie en integratie in de landbouwwetenschap. Wetenschap en Samenleving, jaargang 41, no 3.

Launius R.D., Krige J. and Craig J. (2013)
Space Shuttle Legacy: How We Did It and What We Learned (Library of Flight) AIAA, September 30, 2013

Lees F. (1960) Loss prevention in the process industry. Vol 1, Oxford Butterworth Heinemann

Leveson N. (2002) System Safety Engineering: Back to the Future. Massachusetts Institute of Technology, department of Aeronautics and Astronautics. June 2002

Leveson N. (2003) White paper on Approaches to Safety Engineering. April 23, 2003. From: Safeware, 1995, Addison-Wesley

Leveson N. (2004) A new accident model for engineering safer systems. Safety Science 42 (2004) 237-270

Lintsen H. 1980. Ingenieurs in Nederland in de 19e eeuw: Een streven naar erkenning en macht (Dutch engineers in the 19th century: The quest for recognition and power). PhD Thesis

MacIntosh R. (2010) The Accident "Cause" Statement. Is It Beyond Its Time? ISASI Forum, April-June 2010, pp 5-9.

MacIntosh R. (2011) Major Investigations, New (and revised) Thinking Ahead. http://www.isasi.org/Documents/library/technical-papers/2011.

MacIntosh R. (2012) Major investigations, 'Nextgen' thinking ahead. ISASI Forum Jan-March 2012

McCay R.E. (2007) Using the Universal Model in Accident Investigation. www.iprr.org/papers/mccay.htm, April 2007

McIntyre G. (2000) Patterns in safety thinking. Ashgate 2000

Matthews E. (1978) Max Weber, Selections in translations. Cambridge University Press.

Martins E. and Soares M. (2012) Automation under suspicion – case flight AF-447 Air France. Work 41 (2012) 222-224 DOI: 10.3233/WOR-2012-0160-22 IOS Press (2012)

Meech J.F. (1992) Addressing operator errors in supervisory systems. British Aerospace Sowerby research centre, UK International Conference on Information-Decision-Action Systems in complex organizations, 6-8 April 1992.

Mohrmann F. (2013) Investigating flight crew recovery capabilities from system failures in highly automated fourth generation aircraft. Delft University of Technology, Faculty of Aerospace Engineering, December 2013.

Morales Napoles O. (2010) Bayesian Belief Nets and Vines in aviation safety and other applications. Doctoral Thesis Delft University of Technology, Febr 2010.

Morel G., Amalberti R. and Chauvin C. (2008) Articulating the differences between safety and resilience: the decision-making process of professional sea-fishing skippers. Human factors, Feb 2008 Vol 50 issue 1, 2008

McIntyre G. (2000) Patterns in Safety Thinking. Ashgate

Nelson P. (2008) A STAMP analysis of the LEX Comair 5191 accident. MSc thesis Lund University, Lund Sweden, June 2008

NLR-ATSI (2009) CATS: Causal model for Air Transport Safety. NLR-ATSI, March 2009. Ministry of Transport and Infrastructure, the Netherlands

Obert E. (2009) Aerodynamic Design of Transport Aircraft. Delft University of Technology Faculty of Aerospace Engineering IOS Press 2009

Petroski H. (1992) To engineer is human. The Role of failure in Successful Design. Vintage Books, New York

Pritchett A. (2009) Aviation automation: general perspectives and specific guidance for the design of modes and alerts. Reviews of Human factors and Ergonomics, 5, 82-113.

RAND (1993) Hillestad et.al. Airport Growth and Safety. A Study of the External Risks of Schiphol Airport and Possible Safety Enhancement Measures. EAC-RAND, Supported by the Netherlands Ministry of Transport, Public Works and Water Management. Santa Monica, USA

RAND (1998) Wegverkeer binnen de Raad voor de Transportveiligheid. J.A. Stoop, L. van Dorp, J.P. Kahan, J.L. de Kroes. RAND Europe. RE-98.025. RAND Europe Leiden, the Netherlands

Rasmussen J. (1997) Risk management in a dynamic society: a modeling problem. Safety Science Vol. 27, No2/3, pp183-213, 1997

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

Raymer D. (1999) Aircraft Design: A Conceptual Approach AIAA Educational Series Third Edition

Raymer D. (2012) Aircraft Design. Fifth Edition textbook and RDS ^{win} Student software. ISBN 978-1-60086-9321-1. AIAA

Rimson I. and Benner L. (1996) Mishaps investigations: Tools for Evaluating the Quality of System Safety Program Performance. In: Proceedings 14th International System Safety Conference, august 12-17, Albuquerque, New Mexico. pp 1C2-1 – 1C2-9

Robinson G. (1982) Accidents and socio-technical systems; principles for design. Accident Analysis and prevention, Vol 14, No 2, pp 121-130.

Rochlin G. (1999) Safe operation as a social construct. Ergonomics, 1999, Vol 42, No 11, 1549-1560

Roozenburg, N. and Eekels, J. (1995) Product Design, Fundamentals and Methods, Wiley, Chichester, UK.

Rosenthal U. (1999) Challenges of Crisis Management in Europe. In: International Conference on The Future of European Crisis Management. Crisis Research Center, Leiden University and The Swedish Agency for Civil Emergency Planning. November 1999

Roskam J. (2007) Lessons Learned in Aircraft Design. The Devil is in the Details. DAR Corporation, Lawrence, Kansas, USA

Schneider M. and Rosa J. (2011) CARESSI Commercial Airliner Emergency Safety System. ICINCO 2011 – 8th International Conference on Informatics in Control, Automation and Robotics. pp 223-226

Sheridan T. (2008) Risk, Human Error, and System resilience: Fundamental Ideas. Human factors, Vol 50, No 3. June 2008 pp 418-426

Sklet S. (2004) Comparison of some selected methods for accident investigation. Journal of hazardous Materials 111 (2004) 29-37

Slovic P. (1994) Trust, emotion, sex, politics and science: surveying the risk-assessment battlefield Risk Analysis, Vol 19, no 4, 1999 pp 689-701

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 pp1-12

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

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

Stoop J.A. (1997) Airport growth and safety; improvement of the external and internal risks of airports. Aviation Safety; pp 615-629. Ed. H. Soekkha, VSP 1997

Stoop J.A. (2002) Accident investigations: trends, paradoxes and opportunities. International Journal of Emergency Management . Vol 1, No 2, 2002, pp 170-182

Stoop J.A. and Beukenkamp W. (2003) Monitoring safety in design and construct; the HSL-South case study. ITA World Tunneling Congress 2002, (Re)Claiming the underground space. 12-17 april 2003, Amsterdam the Netherlands.

Stoop J.A. (2003) Critical size events: a new tool for crisis management resource allocation? Safety Science, Vol 41, no 3, pp 465-480.

Stoop J.A. (2004) Independent accident investigation: a modern safety tool. Journal of hazardous Materials 111 (2004) 39-44

Stoop J.A. and Dekker S.A. (2010) Accident modelling: from symptom to system. In: D. de Waard, A. Axelson, M. Berglund, B. Peters and C. Weikert (Eds). Human factors: a system view of human, technology and organization (pp 185-98). Maastricht, the Netherlands, Shaker Publishing.

Stoop J.A., (2011) Timeliness, an investigators challenge. Investigation – A shared process, ISASI 2011, 12-15 Sept, Salt Lake City Utah, USA

Stoop J.A. and Dekker S. (2012) Are safety investigations pro-active? Special Issue Safety science 50 (2012) 1422-1430. Future challenges of accident investigation , some insights from the 33rd ESReDA Seminar. Safety science 50 (2012) 1422-1430.

Stoop J.A. (2012) Time as a safety critical system integrator for stall recovery in aviation. Human Factors and Ergonomics Society Europe Chapter, Human Factors: a view from an integrative perspective. October 10-12, 2012, Toulouse, France

Stoop J.A. and Van der Burg R. (2012) From factor to vector, a systems engineering design perspective on safety. ESREL Helsinki

Stoop J.A. and de Kroes J.L. (2012) Stall shield devices, an innovative approach to stall prevention? Third Air Transport and Operations Symposium, 18-20 June 2012, Delft University of Technology, The Netherlands

Stoop J.A. (2013) Towards a Failsafe Flight Envelope Protection: the Recovery Shield. Advances in Risk and Reliability Technology Symposium. University of Nottingham, UK. 21-23 May 2013

Stoop J.A. and Van Kleef E.A. (2014) Reliable or Resilient: Recovery from the Unanticipated. Paper submitted to the Special Issue of International Journal of Performability Engineering on Transport System Safety, Risk and Asset Management 2014

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

Stroeve H., Blom H. and Bakker G. (2013) Contrasting safety assessments of a runway incursion scenario: Event sequence analysis versus multi-agent dynamic risk modeling. Reliability Engineering and System Safety 109 (2013) 133-149

Taleb N. (2007) The Black Swan: The Impact of the Highly Improbable. Random House, new York ISBN 978-1400063512, 2007

TCI (2004). Onderzoek naar infrastructuurprojecten. Nr 10 Grote infrastructuurprojecten: inzichten en aandachtspunten (Achtergrondstudies). Vergaderjaar 2004-2005 Tweede Kamer der Staten Generaal. Kamerstuk 29283 nr. 10. Gepubliceerd 9 december 2004 (Interim Committee for the Infrastructure, Parliamentary Hearing Committee, Dutch Parliament, background studies)

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

Torenbeek E., (2013) Advanced Aircraft Design. Conceptual Design, Analysis and Optimization of Subsonic Civil Airplanes. Wiley, Aerospace Series

Troadec J.P. (2013) The final word: Air France flight 447. ISASI Forum, Jan-March 2013.

Van Kleef E. and Stoop J. (2014). Reliable, Resilient: Towards a Dialectic Synthesis? 46th ESReDA Seminar on Reliability Assessment and Life Cycle Analysis of Structures and Infrastructures, May 29th - 30th, 2014, Politecnico di Torino, Turino, Italy

Van Meer P.M. (2010) Development of a systemic safety investigation methodology for KLM E&M. Masters' thesis Delft University of Technology. Faculty of Aerospace Engineering, 26th April 2010.

Van Vollenhoven P. (2002) Independent Accident Investigation: Every Citizen's Right, Society's Duty. Dutch Transportation Safety Board. The Hague, the Netherlands

Van Vollenhoven P. (2006) RisicoVol. Inaugural lecture University of Twente, 2006

Yin R. (1994) Case Study Research. Design and Methods. Applied Social Research Methods Series Vol 5, SAGE Publications

Young M., Braithwaite G., Shorrock S. and Faulkner J. (2005) The (R)Evolution of human factors in transport safety investigations, ISASI Forum. July-September 2005.