REPORT

Interdisciplinary Risk Assessment of Integrated Operations addressing Human and Organisational Factors (RIO)

www.sintef.no/rio

Essays on socio-technical vulnerabilities and strategies of control in Integrated Operations

Albrechtsen, E. (editor); Andersen, S.; Besnard, D.; Grøtan, T.O.; Hollnagel, E.; Hovden, J.; Mostue; B.Aa.; Størseth, F.; Vatn, J.

February 2010

MINES ParisTech SINTEF NTNU 100 skapende år

SINTEF REPORT

SINTEF Technology and Society
Safety Research

Address: NO-7465 Trondheim, NORWAY
Location: S P Andersens veg 5 NO-7031 Trondheim
Telephone: +47 73 59 03 00
Fax: +47 73 59 28 96

Enterprise No.: NO 948 007 029 MVA

TITLE

Essays on socio-technical vulnerabilities and strategies of control in Integrated Operations

AUTHOR(S)

Albrechtsen, E.[1](editor); Andersen, S. [2]; Besnard, D. [3]; Grøtan, T.O. [1]; Hollnagel, E. [3].; Hovden, J. [2]; Mostue; B.Aa. [1]; Størseth, F. [1]; Vatn, J. [4]

[1] SINTEF Technology and Society, Safety Research;
[2] Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology
[3] Industrial Safety Chair, MINES ParisTech, Sophia Antipolis, France;
[4] Department of Production and Quality Engineering, Norwegian University of Science and Technology

CLIENT(S)

The Research Council of Norway
The Norwegian Petroleum Safety Authority
Center for Integrated Operations in the Petroleum Industry

REPORT NO.	CLASSIFICATION	CLIENTS REF.	
SINTEF A14732	Unrestricted	Tor Petter Johnsen (The Research Council of Norway); Grete Løland (PSA); Jon Kvalem (IO center)	
CLASS. THIS PAGE	ISBN	PROJECT NO.	NO. OF PAGES/APPENDICES
Unrestricted	978-82-14-04866-7	504185.20 Interdisciplinary Risk Assessment in Integrated Operations (RIO)	74
ELECTRONIC FILE CODE		PROJECT MANAGER (NAME, SIGN.)	CHECKED BY (NAME, SIGN.)
U:\prosjekt\6050-Sikkerhet\Prosjektoversikt\Rapporter\Åpne\2010\SIN		Eirik Albrechtsen E. Albrechtsen	Ranveig K. Tinmannsvik R.K. Tinmannsvik
FILE CODE	DATE	APPROVED BY (NAME, POSITION, SIGN.)	
U/Prosjekt/6050	2010-01-28	Lars Bodsberg, Research Director Lars Bodsberg	

ABSTRACT

The report consists of 11 essays written by authors representing different disciplines and schools in risk research. The collection of essays elaborates various threats and vulnerabilities in Integrated Operations (IO) and/or discusses strategies for assessing and handling risk in IO. As a result, the report shows a variety of interpretations of IO-related threats and vulnerabilities as well as a range of approaches to assessing risks related to IO concepts.

The report is written within the RIO project (Interdisciplinary Risk Assessment in Integrated Operations) project, sponsored by the Norwegian Petroleum Safety Authority, the IO Center at NTNU and the PETROMAKS program of the Norwegian Research Council.

KEYWORDS	ENGLISH	NORWEGIAN
GROUP 1	Risk	Risiko
GROUP 2	Risk assessment	Risikovurdering
SELECTED BY AUTHOR	Integrated Operations	Integrerte operasjoner
	Interdisciplinary	Tverrfaglig

Table of contents

Preface: Views on socio-technical vulnerabilities and strategies of control in Integrated Operations

Eirik Albrechtsen

SINTEF Technology and Society, dep. of Safety Research, Trondheim, Norway
UUUUeirik.albrechtsen@sintef.no

Denis Besnard

Industrial Safety Chair, MINES ParisTech, Sophia Antipolis, France
denis.besnard@mines-paristech.fr

Background and purpose

There is an ongoing trend at the Norwegian continental shelf to create new ways of working based on digital infrastructure and information technology. The industry claims that the development creates efficient reservoir exploitation, optimisation of exploration and operation processes, managed development of fields and installations, as well as improved HSE performance (OLF, 2003). There is no straightforward description of the concept Integrated Operations (IO), as there are different IO solutions among companies and among installations. We can however identify some generic properties of IO:

- use of information technology and digital infrastructure to enable new work practises;
- increased capture of offshore performance data;
- use of real-time data to monitor and manage operations across geographical and organisational borders;
- use of collaborative technology to link different actors in a more efficient and closer way;
- access to expert knowledge.

Altogether these properties result in tighter integration of technology, data, competency, activities and organisations, thereby emphasising interdisciplinary approaches. The general belief is that these IO concepts will provide “better, faster and safer decisions”. However, new opportunities and challenges for risk assessment and management arise in this IO environment. The purpose of this report is to show some of the challenges and opportunities that need to be considered when assessing and managing risk in IO.

The report consists of 11 essays, which are written by 9 authors. The authors represent different disciplines and schools in safety research. As a result the set of essays shows a range of different perspectives on IO and risk management strategies. All essays were presented at a seminar with participation from the petroleum industry and public authorities on November 17th, 2009, in Trondheim (Norway).

The report is written within the RIO project (Interdisciplinary Risk Assessment in Integrated Operations)

About the RIO project

The RIO project (Interdisciplinary Risk Assessment in Integrated Operations addressing Human and Organisational Factors), sponsored by the PETROMAKS program of the Norwegian Research Council; the Norwegian Petroleum Safety Authority; and the Center for Integrated Operations in the Petroleum Industry at NTNU. The project period is 2008-11.

The overall project goal is to develop new knowledge (theories, models) and frameworks for reasoning, as a basis and platform for risk assessment in relation to petroleum production in an integrated operations environment, and to provide guidance on the practical use of these results to relevant practitioners in the petroleum industry.

The main project activities are:

- Establish an IO baseline by identifying a list of change elements characterising IO with respect to human and organizational aspects, and the scientific foundations for dealing with the risk aspect of the identified changes.
- Analyze which effects the IO-related changes have on risk by using different perspectives on organizational accidents and resilient organizations.
- Position different methods for risk and safety assessments according to system characteristics, and evaluate the applicability of methods related to IO-related changes.
- Develop a guideline document for risk assessment of IO concepts to be used by practitioners in the petroleum industry.
- Establish a risk assessment course for IO

More project information is found at the project website: www.sintef.no/rio

Contents of the report

Jan Hovden opens this volume by making the point that there are various positions regarding risk, various perspectives, and various epistemological approaches; a point that is also illustrated with this collection of essays. Hovden's decomposition highlights the simple fact that there is not a single, unified way to look at risks. Furthermore, Hovden presents arguments for various and interdisciplinary approaches to risk management in an IO environment. Indeed, there is a danger to study risks in silos, the various levels of an organisation being understood via different sets of knowledge. This creates a challenge since risk-related problems can be selected according to a favourite discipline or research paradigm. A change is needed. Risks must be tackled via a collaboration of disciplines, and systems must be decomposed into functions as opposed to e.g. organisational layers.

Also following the argument that a change is needed, Erik Hollnagel, states that it is necessary to look for approaches that can be used for intractable systems, i.e. systems that are incompletely described or underspecified. It is not enough to consider risks that arise from recognisable failures or malfunctions, alone or in combination. It is also necessary to consider risks that are due to unforeseen combinations of the variability of normal performance. According to Hollnagel, there has been an evolution of how risks are approached. A technological approach (using probabilistic risk assessment) has progressively been complemented by a human factors approach (and human reliability assessment). Closer to our time, safety management promoted the idea that the organisation also had to be taken into account. However, the large picture is that socio-technical systems have historically been understood in terms of how we would like them to operate as opposed to how they actually work. One reason for this is that socio-technical systems are intractable, thereby calling for a new approach. Hollnagel proposes resilience engineering as an answer to this challenge, by a functional approach looking to create robust yet flexible processes, to monitor and revise risk models and use resources in a proactive way for unexpected developments. He also calls for methods that focus on safe functioning, as opposed to the many methods focusing on unsafe

functioning, since safety is not only about preventing unsafe functioning but also about sustaining safe functioning.

In his essay on risk assessment and system characteristics, Jørn Vatn provides an operational definition of risk: an uncertainty regarding the occurrence and severity of undesired events. When defining risks, Vatn also deals with the notion of domains, of which he lists the following: the real world, the scientific cause and effects, uncertainty, and values and preferences. According to his view, as long as the risk assessor stays in the uncertainty domain, the risks associated to any kind of system characteristics can be assessed. Also, by using opportunity-based maintenance as a case study, he puts forward the argument that increased complexity and intractability represent uncertainties. Eventually, these can be structured by using his proposed risk analytical framework.

Tor Olav Grøtan makes the point that IO can cause disruptions and unanticipated effects. This requires an assessment of the IO contribution to risk. With this perspective, the essay attempts to differentiate complexity-induced risks from others. But contrary to Vatn's essay on risk assessment and system characteristics, Grøtan poses that emergence cannot be treated as uncertainty. Also, his position is that some of the complexity induced by IO cannot captured by more analytical effort. An alternative view is the Cynefin model that Grøtan presents as a possible way of understanding complex systems. Depending on their specific traits and configuration, these systems can be seen or classified as: chaos, complex, knowable or known. Cynefin embeds some general descriptions that depart from a purely analytical view of risk assessment: in systems that cannot be anticipated, not all causal relationships can be known. This has implications for risk management: how do we handle emergence and the ever-changing nature of systems?

In his second essay (on perspectives and risk assessment in IO), Jørn Vatn uses various perspectives on organisational accident and resilient organisation for a quantitative risk assessment framework. He addresses this issue by employing a set of control questions for every perspective used. The responses to the control questions can be transferred into what Vatn denotes safety critical functions (SCF). They are functions whose failure or operational deficiency will increase the risk level. With his essay, Vatn makes the demonstration that social science-based knowledge on organisational accidents and resilient organisations can be incorporated into quantitative risk analysis.

Siri Andersen also considers risk assessment and perspectives by adopting a high-level view on IO. She attempts to compare HRO (High Reliability Organisations) and Resilience Engineering in terms of sources of evaluation criteria for risk assessment methods. This reflection is applied to system elements that are left unchanged by the introduction of IO, and to the new elements as well. From a preliminary comparison of this ongoing piece of research, Andersen proposes that no risk assessment method will fit all the dimensions of HRO and Resilience Engineering. Her conclusion is that the problem is that of knowing what are the key dimensions of HRO and Resilience Engineering, and work on that basis towards identifying evaluation criteria for risk assessment methods.

Bodil Mostue and Eirik Albrechtsen question themselves on the possible side-effects of IO and attempt to identify a range of possible decision-related shortcomings. Namely, they look into the constraints imposed by IO on decision makers, especially during emergencies. Some situation-specific constraints are already known. However, other dimensions are at play. For instance, in a given company, not everyone is equally close to hazards. Also, not all actors involved in crisis management have the same level of knowledge about the overall system

reaction or preparedness. Finally, the possibility of information overload and interactions between decisions provide the final touch to a complex risk picture. These are only some examples that Mostue and Albrechtsen discuss. The key message is that as far as decisions are concerned, the introduction of IO is an organisational change that will solve some problems but will also create a series of new challenges.

In a psychological perspective, Fred Størseth also touches on one of the possible unwanted side-effects of IO. The latter being a change in the organisation, it might require workers to adapt. However, IO can have a truly detrimental effect if it increases demands on people when at the same time leaving them in a conditions of limited control. The overarching argument is that despite what IO is desired to be by some, its implementation will have effects on people, and these effects must be identified. Størseth puts forward the argument that resilience engineering approaches place too much demand on individuals by not taking into account organisational layers and segments. How can adaptation to change be engineered into a system without impacting on the workers, Størseth asks.

The individual dimension is also the spin adopted by Denis Besnard. He looks into the human-machine interaction issues that might arise when introducing IO. Namely, a drilling assistance system called eDrilling is being deployed in the Norwegian petroleum industry. The essay investigates some general automation issues related to eDrilling (e.g. system authority, change of control cues, mode confusion). Besnard discusses some automation-related accidents that happened in other industries to make the point that some human-machine interaction issues could be used as risk assessment criteria in IO, especially where automation is to be introduced.

In his second essay, Grøtan examines how collaboration technology reduces or reinforces complexity in IO. Based on his arguments from his first essay on complexity as an embedded “wildness-in-wait” and as un-order, he argues that information and communication technology is a significant source of complexity. The latter can create escalation effects not only in the real world but also in the realm of information, interpretation, knowledge and sensemaking.

In his single-authored essay, Albrechtsen is investigating the positive counterpart of introducing IO into systems. Namely, IO can have a number of desirable effects that can be captured by safety views ranging from the barriers perspective to the high reliability organisations perspective. In other words, IO could be beneficial to organisations in many ways, and at various organisational levels. But beyond operations, IO can also have an effect on risk assessment. For instance, better availability of real-time information or better access to expert knowledge is expected to improve the performance of the assessment exercise. Such features could be those of a safety operations support centre, in the form of pre-assessment, information and methods databases, interdisciplinary capabilities and so on. Albrechtsen concludes his essay by bridging his point with resilience engineering and the positive contribution to safety that is made by normal performance. IO could be instrumental in enhancing the latter.

Towards interdisciplinary views of risk related to integrated operations

In his essay, Hovden describes an interdisciplinary field as an area of study that crosses traditional boundaries between academic disciplines or schools of thought, as new needs for problem solving and organising of research communities emerge. The set of essays presented in this report was written by authors representing different academic disciplines, schools and

perspectives. It shows the variety of an interdisciplinary approach to risk assessment of IO concepts. Concretely, the essays illustrate a range of interpretations of vulnerabilities and opportunities in IO, as well as various approaches to assess and manage related risks. Only by communicating, recognising and respecting various understandings and approaches can interdisciplinary perspectives converge towards a better understanding of the consequences of the introduction of such a change as IO.

The interdisciplinarity of this report is the *collection* of essays, not each essay. This means that discussions sometimes revolve around the same system properties, but exhibit diverging positions. For example, both Vatn (in his first paper) and Grøtan (in both his papers) address how to assess risk for systems with regard to complexity. They consider the same system characteristic. However they propose different approaches in expressing risk. One will find a similar divergence between Hollnagel's approach to intractable systems and how Vatn (in his first paper) proposes to assess risk in intractable systems. Another example of a contrasting view on IO and risk management is the expected positive safety effect of the introduction of IO (Albrechtsen), as opposed to the downsides of IO identified by Grøtan (second paper), Mostue and Albrechtsen, as well as Størseth. Furthermore, we find a contrast between Hollnagel's arguments for a resilience engineering approach and Størseth's questions on whether resilience engineering considers all levels in an organisation adequately. Also, Vatn's arguments that different system characteristics can be handled by a quantitative risk approach are challenged by Hovden's arguments in favour of relying on different disciplines in risk management.

The importance of different perspectives on risk governance of Integrated Operations - a need for interdisciplinary risk research?

Jan Hovden

Department of Industrial Economics and Technology Management
Norwegian University of Science and Technology, Trondheim, Norway
jan.hovden@iot.ntnu.no

Abstract. The essay discusses the scientific basis for the RIO project (Interdisciplinary Risk Assessment of Integrated Operations addressing Human and Organisational Factors – RIO) by looking at some of the presumptions and positions in the project description in terms of epistemological approaches to risk and the distinction between discipline-based academic risk research and applied, problem oriented risk research.. Arguments, reasons and elaborations for major hazard prevention in an IO context based on an interdisciplinary framework for risk governance is presented.

Introduction and background[1]

Why asking this question in the title of the paper? Different perspectives and interdisciplinary approaches is a presumption for the RIO project (Interdisciplinary Risk Assessment of Integrated Operations addressing Human and Organisational Factors – RIO). The aims are therefore to give some arguments, reasons and elaborations for what is already decided. The scope given by the RIO project proposal is major hazard prevention in an IO context based on an interdisciplinary framework for risk governance (Renn, 2008). A core challenge for studying safety performance in a context of Integrated Operations is distributed, collaborative decision-making in control of hazardous processes and the adaptive response of decision-makers to internal and external stressors and variability.

Compared to the stable conditions of the past, the present dynamic oil and gas industry brings with it some dramatic changes of the conditions of industrial risk governance:

- The very fast pace of development of technology, especially information technology, leads to a high degree of integration and coupling of systems and effects of a single decision can have dramatic effects that propagate rapidly. It is thus becoming increasingly difficult to explain accident causation by analysing local factors within a work system. Safety and risk management increasingly become system problems.
- Furthermore, companies today live in a very aggressive and competitive environment. The German sociologist Ulrich Beck (1992) summarizes these challenges for risk management by the phrases: “produced uncertainties and organized irresponsibility” and “strategic uncertainty and structural vulnerability” as key words for risk research.

The socio-technical system involved in risk management is normally de¬composed according to organizational levels and specific hazard phenomena, which are the subjects for studies within different disciplines, see the “vertical” risk management model of J. Rasmussen (1997). This raises the problem of the constraints of specialized academic risk research.

[1] The paper is to a large extent based on a report for SRV (Hovden & Rasmussen, 1998) and a report for the Research Council of Norway (Hovden, 1999), and inspired by discussions and confrontations at meetings within the RIO group of researchers.

The RIO Problem Context

The overall objective of RIO is to develop new knowledge (theories, models) and frameworks for reasoning, as a *basis and platform for risk assessment* in relation to petroleum production in an integrated operations (IO) environment. The interdisciplinary RIO framework is described in terms of:

- Incorporation of the impact of human cooperation (human factors) and organisational factors (MTO) on risk in IO, with a special emphasis on their contribution to *systemic*[2] risks.
- Risk governance processes that addresses various degrees of risk *novelty* – whether being emerging, re-emerging, increasing in importance, current (topical) and/or institutional (already subject to management attention and decisions).
- Interdisciplinary planning and engineering of new projects or large modification projects, aiming to benefit from IO. Coordination of risk governance between cooperating parties in an IO environment.

According to the project plan the RIO emphasis is on *generation of knowledge for appraisal, assessment and characterisation of risk*, as a preparation for subsequent risk management.

Approaches to risk research

To avoid confusions in the discussion of approaches to risk research it may useful to make a distinction between:

- Paradigms – theories of knowledge, e.g. regarding the "risk" concept, - see own section below.
- Academic disciplines[3] - traditions of organising subjects at universities. In 1231 the University of Paris consisted of four faculties: Theology, Medicine, Canon Law and Arts. Today the main grouping is humanities, social sciences, natural sciences, formal sciences, professions and applied sciences. There is no consensus on how academic disciplines should be classified, and most new disciplines are merges of subjects from old disciplines driven by developments in society and technology.
- Schools - within the disciplines and even for the same subject we will find competing schools or perspectives on understanding and analysing a research object.
- Perspectives - combining schools and theories from different fields in analysing a research problem.
- An interdisciplinary field or multidisciplinary field is a field of study that crosses traditional boundaries between academic disciplines or schools of thought, as new needs for problem solving and organizing of research communities have emerged. Cross-disciplinarity is the act of crossing disciplinary boundaries to explain one subject in the terms of another, foreign subject or method.

The RIO approach implies interdiciplinarity in terms of combining mathematical statistics in risk analysis modelling with safety engineering and different social science perspectives on organisational accident risks.

---

[2] Risks that cannot be understood as (or reduced to) solely a result of a technical component or human malfunction/failure, see also Hollnagel (2008)

[3] An academic discipline, or field of study, is a branch of knowledge which is taught and researched at the college or university level. Disciplines are defined, and recognized by the academic journals in which research is published, and the learned societies and academic departments or faculties to which their practitioners belong. Fields of study usually have several sub-disciplines or branches, and the distinguishing lines between these are often both arbitrary and ambiguous (from Wikipedia)

Perspectives or frames

Organizational theory has a rich history from Plato on leadership, Aristotle on communication, Machiavelli on power, Weber on rational organisations, Taylor on scientific management, Mayo on human relations, Simon & March on organisational behaviour, control and stress modelling, Minzberg's organigraph, institutional theory, contingency theory, cultural explanations of organisations and change, organisational learning, to Weick's mindful organisations, etc. In their famous textbook Bolman & Deal (1984) recapitulated almost all the different schools into just four main *frames, perspectives or approaches* for analysing organisations, namely structural, human resources, political, and symbolic. It was a great success both for the understanding of organisations and as a basis for organisational diagnosis in practical organisational development projects.

I guess this frames also inspired Rosness et al (2004) in their report "Organisational Accidents and Resilient Organisations. Five Perspectives." These perspectives have generated new research and have also been successful for educational purposes. The demarcation of the perspectives and the choice of perspectives can, of course, be questioned and need further elaboration. The five perspectives seem to be a mix between a grouping of accident models and organizational models. Therefore, many different and competing accident models highlight different aspects of a risk problem. Kjellén (2000) classify them as Causal-sequence models, Process models, Energy models, Logical tree models, Human information-processing models and Safety management models. Hollnagel (2004) defines three groups: Sequential, Epidemiological, and Systemic accident models. It may be interesting to make a table of organizational perspectives on one axis and main accident models on the other and discussing the contents of the cells of the table. An alternative grouping of perspectives could be:

- Safety engineering, including risk analyses, MTO modeling, safety management systems
- Resilience engineering including human factors and information processing
- Organization analyses and diagnosis (like OD/"pentagon"-model, etc.) including safety culture, leadership, etc.
- Power and politics, including actors and decision-making, risk governance

In a way safety research is much wider than risk research, and that will of course influence the choice of perspectives.

Epistemological approaches to risk - theories of knowledge

The meaning of the "risk" concept represents a huge cleavage between academic disciplines which leads to problems for multi-disciplinary communication and collaboration in risk research. The pre-modern use of the Latin term riscum was attributed to maritime ventures and natural disastrous events, an objective danger, an act of God, a force majeure, excluding the idea of human or societal responsibility and influence. The Enlightenment in the seventeenth century and industrialisation demanded more objective knowledge, rational thinking, and a search for predictability and control (Giddens, 1991). The development of a science of probability theory and statistics gave a basis for modern risk analysis. The modernist concept of risk represented a new way of viewing the world and its chaotic manifestations (Lupton, 1999). Risks could be predicted, prevented and controlled by human action and technology, and thereby replacing earlier concepts of fate or fortune (Giddens, 1990). The probability estimates of an event and its consequences rely upon knowledge regarding modelling and data to fit into a risk analysis.

A basic critique of shortcomings of the risk analysis methodology and reductionist risk evaluations in dealing with modern complex, non-linear socio-technical systems was given by Perrow (1984). On societal risk issues we often get science/risk experts against lay people, and a lack of confidence and trust in experts and institutions. In general, the new uncertainties created by the fast developments in science and technology combined with the features of global systemic risks produce an attribution of risks and a feeling of helplessness with some similarities to the Middle Ages.The distrust of risk assessments of science and technological disciplines, opened for social science based risk research revealing that people perceive and assess risks in a different and richer way than the two dimensions of probabilities and consequences calculated by the risk analysts (Klinke & Renn, 2001).

There is a difference between the risk itself and public perception of it. The main dichotomy between natural-scientific objectivism and cultural relativism can be detailed and paraphrased as follows (partly based on Lupton, 1999):

- *Rationalist* – risks as real world phenomena to be measured, and estimated by statistics, prioritised by normative decision theory (economics) and controlled by "scientific management".
- *Realist* – risks are objective hazards or threats that exist and can be estimated independently of social and cultural processes, but may be distorted or biased through social and cultural frameworks of interpretation.
- *Middle positions between realist and constructivist* – risk is an objective hazard or threat that is inevitably mediated through social and cultural processes and can never be known in isolation from these processes.
- *Constructionist* – nothing is a risk in itself – what we understand to be a "risk" is a product of historically, socially and politically contingent "ways of seeing".

To simplify: in risk research most sociologists are in the middle position, e.g. Renn (2008), psychologists in the realist position, and anthropologists in the constructionist position. The needs for cross-disciplinary approaches in risk research revealed by Rasmussen (1997) and his model of convergence of scientific paradigms are in line with a middle position regarding epistemology[4] .

The risk analyst has traditionally been a "rationalist" defining risk as an expected value of probability and consequence. In recent years we can observe a move in parts of the risk analysis society towards a middle position. This is best illustrated by the coming book of T. Aven (2010) on "Misconceptions of Risk", plus Aven (2009) and Aven & Renn (2009) where a risk is defined as f(A,C,U,P,K) where A is events/accidents, C is consequences, U is uncertainty about A and C and underlying phenomena, P is P(A)=P(A|K) (Bayesian), K is the knowledge base the probability is based on. According to Aven are all other definitions of risk are misconceptions. This seems very academic and arrogant; even though I agree with his definition, it totally disregards the knowledge and reasons behind the different approaches to risk assessment in different disciplines, - engineering, medicine, economics, etc. These definitions are not always consistent in a strict scientific sense, but can be useful and practical in real risk management of specific hazards and contexts. In well defined, known technical systems bow-tie modelling and calculations of expected values may be valid as input for decision-making. For hazards representing systemic and emerging risks the wider definition of Aven and also the "black swan" metaphor are relevant. The risk picture of RIO covers the

[4] Alternative epistemological positions and theoretical frameworks are also reviewed in Krimsky & Golding (eds), 1992; Adams, 1995; Turner & Pidgeon, 1997, Lupton, 1999.

whole scale from violation of simple technical barriers to the “black swans” in complex ICT systems.

Discipline-based Academic Risk Research

The constraints of a “closed” academic research environment should be considered. Academic research is organized in more or less classic disciplines with well-defined boundaries, each served by particular scientific journals giving merits to the researcher and money incentives to the university. In academic research it is necessary to focus on problems which can be formulated with reference to established concepts, methods and paradigms within a discipline, problems that are manageable and will give experience adding up to a recognizable academic competence and status.

For safety research, the consequence is that each of the organizational levels involved in risk management is studied by a different discipline (engineering, psychology and human-machine interaction, management and industrial engineering, organizational sociology, economics, and law) and because researchers will have little or no competence with respect to the functionality of particular work systems, safety research will have a 'horizontal' orientation across work systems and hazard sources. The integration is on the premises of the disciplines as it is not driven by a united problem-definition and understanding of the primary hazardous processes. – Nevertheless, applied problem oriented risk research needs the results and achievements in such basic academic risk research as an input, inspiration and corrective.

Applied Problem Oriented Risk Research: interdisciplinary, multidisciplinary or cross-disciplinary

In contrast, problem driven safety research cannot select the problems to consider according to a favorite research paradigm. In addition, actual safety problems cut across disciplines. Furthermore, complex, cross-disciplinary research, by nature, requires an extended time horizon. It takes considerable time to be familiar with the functioning of complex work systems and with the paradigms of other disciplines, and often time consuming field studies are required. Quite naturally, such studies are less tempting for researchers who have to present a significant volume of publications each year to ensure tenure.

Results of recent studies of major accidents point to the kind of paradigm required to capture the nature of accident causation and to guide a proactive design of safer systems. Commercial success in a competitive environment implies exploitation of the benefit from operating at the fringes of the usual, accepted practice. Exploring the boundaries of normal and functionally acceptable boundaries of established practice during critical situations necessarily imply the risk of crossing the limits of safe practices (Rasmussen, 1997).

We have to understand the mechanisms generating the actual behavior of decision makers at all levels, to identify the information needs of decision makers both with respect to the actual state of affairs and to values and objectives, and we have to identify aspects that are sensitive to improvement and, therefore, the targets of guidelines for industrial risk management. Considering the dynamic nature of a modern society with companies and institutions constantly striving to adapt to a dynamic environment, a shift is needed in the basic modeling philosophy from the decomposition of the system into elements studied separately toward a system oriented model describing the role of actors at all levels in the control of the hazard sources of a work place. To have a system modeling language capturing performance at all levels, functional abstraction rather than structural decomposition is required. This is a “cross-

disciplinary" issue, involving technical as well as several social sciences. *In writing this essay I had an intention to differentiate between inter-,multi- and cross-disciplinary, but I am not able to be consistent, and I also started wondering about how important or useful these distinctions are in context of the RIO project.*

Some Important Research Questions

The research in the RIO project should focus on a study of risk assessment for the distributed, collaborative decision making in control of hazardous work processes and the adaptive response of decision makers in collaboration centers. This will involve research questions such as the following:

- Mapping the decision-making network and information flow within the total socio-technical system involved in operational and risk management, that is, the structure of the distributed decision making network and the information flow found in the different categories of hazard. How is, for instance, the interaction among the engineering design community, the operating community and the regulatory community with respect to risk management.
- The dynamic interaction between the levels and layers; what is the effect of the different time constants for adaptation to change at the various functional levels; and the influence of this difference on the interaction among decision making at the various levels;
- Analysis of accident and threat scenarios to identify critical decision situations that are sensitive to environmental stressors and inadequate information, i.e. some sort of vulnerability analysis.
- Analysis of the characteristics of hazard sources within the various work domains. Accidents and intentional violations are caused by loss of control of hazardous processes. A taxonomy of hazard sources and related modes of risk management strategies should be developed. The control structure involved in risk management for different categories of hazard sources should be described.
- Analysis of the transfer of information from design to the operating community. Present accidents have indicated that this transfer does not keep up with the present technological pace of change. The problem goes also the opposite way, as experience transfer from operation to design is lacking and consequently creates problems for operations and work places not adjusted to the operators and leading to unforeseen hazards and complicated redesign.

Conclusions

I recommend an open-minded, and mindful research approach as the RIO project has to deal with a great variety risk assessments of IO from the well known old hazards of the industry to emerging systemic risks and "black swans" introduced by the interface between new ICT and organisational changes for work practices. Therefore, applying different and multidisciplinary perspectives on risk management seems appropriate.

From Failures to Emergence: The Changing Nature of Risk

Erik Hollnagel

Industrial Safety Chair, MINES ParisTech, Sophia Antipolis, France
&
Department of Industrial Economics and Technology Management
Norwegian University of Science and Technology, Trondheim, Norway
erik.hollnagel@crc.ensmp.fr

Abstract. The focus of safety, and therefore also the assumptions about the sources of risks, must change to match the nature of industrial environments. Technology and engineering were from the beginning the major concerns, and models and methods were developed to a address that. Developments from the 1970s and onwards demonstrated the need to consider also human and organisational factors. Today it is, however, no longer possible to describe socio-technical systems precisely enough to apply the time-honoured methods. A change of approach is required where safety is seen as the ability to function effectively in expected and unexpected situations, rather than as the absence of adverse outcomes.

Introduction

The rationale for safety has always been to prevent that something – a process or an activity – failed or malfunctioned with unintended and unwanted outcomes as the result. The manifestations of failures or malfunctions usually lead to the loss of property, material, or life, and also bring normal functioning to a halt for a shorter or longer period of time. In the aftermath of such failures or malfunctions it is necessary both to recover or restore normal functioning and to find an explanation of why things went wrong, so that steps can be taken to prevent it from happening again.

The realisation that things can go wrong is as old as civilisation itself. Early evidence can be found in the Code of Hammurabi, written in 1.760 BC. It was nevertheless not until the late 19th Century that industrial risk and safety became a common concern, as described by Hale & Hovden (1998). These authors described three distinct ages in the scientific study of safety, which they named the age of technology, the age of human factors, and the age of safety management.

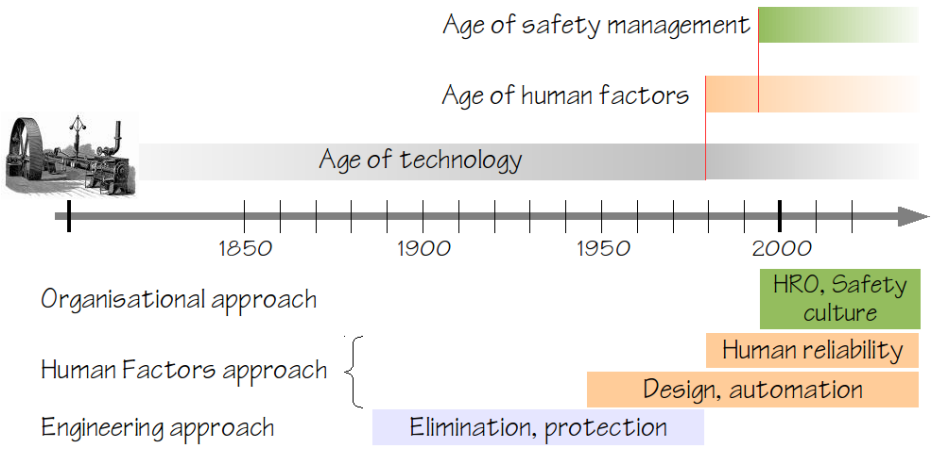


Figure 1. Distinct ages in the scientific study of risk

Risk in the Age of Technology

In the age of technology, the main concern was to find ways to guard machinery, to stop explosions, and to prevent structures from collapsing. This 'age' began with the industrial revolution around the middle of the 18th century and lasted until the end of the 1970s. During the first two centuries the development of technology was relatively slow and was proportional rather than exponential. There was, on the whole, time to learn by experience and formal methods for risk assessment were uncommon.

In the 1940s, during and around the period of the Second World War, the development of new sciences and technologies (such as information theory, cybernetics, digital computers, and the transistor) radically changed the setting. From the 1950s and onwards, the technologies for control, computation, and communication began to develop in an exponential fashion. This meant that the size, and therefore also the complexity, of technological systems expanded very rapidly. Industrial and military systems alike soon became so complex that they challenged established practices and ways of working and new methods were needed to address risk and safety issues. Reliability engineering, as a combination of probability theory with reliability theory, became a separate discipline by the early 1950s. Fault tree analysis was developed in 1961 to evaluate the Minuteman Launch Control System for the possibility of an unauthorized missile launch. Other methods such as FMEA and HAZOP were developed not just to analyse possible causes of hazards (and later on, causes of accidents), but also to identify hazards and risks. Probabilistic Risk Assessment (PRA) was successfully applied to the field of nuclear power generation with the WASH-1400 'Reactor Safety Study' (Atomic Energy Commission, 1975). The WASH-1400 study considered the course of events which might arise during a serious accident at a large modern Light Water Reactor, using a fault tree/event tree approach, and established PRA as the standard approach in the safety-assessment of modern nuclear power plants.

Risk in the Age of Human Factors

The second age started rather abruptly with the accident at the Three Mile Island (TMI) nuclear power plant on March 28 1979. This accident made it clear that something was missing in the PRAs, namely the human factor. Although Human Factors Engineering had been practised since the middle of the 1940s, the focus had been on the efficiency (productivity) side of system design with little or no interest for safety issues. That changed completely after 1979.

Since PRA had become the industry standard for how to deal with the questions of safety and reliability of technical systems, it also became the natural starting point for addressing the human factor. This quickly led to the development of human reliability assessment (HRA), which at first was an extension of existing methods to consider 'human errors' in analogy with technical failures and malfunctions, but which soon developed more specialised approaches (cf., Hollnagel, 1998; Kirwan, 1994.) These developments effectively extended reliability engineering to cover technological and human factors alike. The use of HRA quickly became established as the standard analysis for NPP safety, although there have never been any fully standardised methods (e.g., Dougherty, 1990) – or even a reasonable agreement among the results produced by different methods (Poucet, 1989).

The third age came about for two reasons. One was that several accidents demonstrated that the established approaches, including PRA-HRA, had their limits. The other was a growing doubt that safety could be ensured by normative approaches, e.g., by forcing procedure compliance. Although the third age began less suddenly than the second age, accidents such as Challenger and Chernobyl, both happening in 1986, made it clear that the organisation had to be considered in addition to the human factor (Reason, 1997). In consequence of that, safety management systems became a target of research and development efforts.

The extension of reliability engineering to cover also organisational issues was, however, even more complicated than the extension to include human factors. It was initially hoped that the impact of, for instance, organisational factors on nuclear power plant safety could be determined by accounting for the dependence that these factors introduced among probabilistic safety assessment parameters (Davoudian, Wu & Apostolakis, 1994). It was, however, soon realised that other ways of thinking were required. In some sense it is hardly surprising that the proportional development of safety assessment methods was unable to match the exponential nature of system development.

At present, the practices of risk assessment and safety management find themselves trying to cope with the complexity of socio-technical systems. On the one hand it is still widely hoped that the established approaches somehow can be extended to include organisational factors and organisational safety issues. In other words, that organisational 'accidents' and organisational 'failures' can be seen as analogous to technical failures. On the other hand there is a growing recognition that methods should address the organisation as a whole and not just treat it as a 'context.' If accidents sometimes can be caused by organisational factors, it follows that any changes to these factors must be the subject of a risk assessment. To put it bluntly, neither human factors nor organisational factors can be adequately addressed by methods that rely on the principles on which technical safety methods are based. There is therefore a need to revise the traditional assumptions and instead take a fresh look at what risk and safety mean in relation to organisations.

One reason for that is that safety is invisible whereas the lack of safety is visible. Safety has traditionally focused on what goes wrong and rarely noticed what goes right. We are preoccupied with failures because it is assumed that we know how a system should work in minute details. In reality, however, we may know how we would *like* to system to work, but we do not always know how it *actually* works. This difference between the ideal and the real is due to the fact that many large industrial systems, and all socio-technical systems, are underspecified or intractable.

Tractable and intractable systems

In order to manage the safety of a system, it is necessary to know what goes on 'inside' it, i.e., to have a sufficiently clear description or specification of the system and its functions. The same requirements must be met in order for a system to be analysed and in order for its risks to be assessed. That this must be so is obvious if we consider the opposite: if we do not have a clear description or specification of a system, and/or if we do not know what goes on 'inside' it, then it is clearly impossible effectively to manage it and therefore also to make a risk assessment. We can capture these qualities by making a distinction between tractable and intractable systems, cf., Table 1 below.

The established approaches to risk assessment require that systems are tractable and that it is possible to describe them in detail. In order for this to be the case, systems must be relatively simple and reasonably stable. Neither of these conditions are fulfilled by socio-technical systems which generally are intractable or underspecified. This means that the established methods are not suitable and that it is necessary to develop approaches that can be used for intractable systems.

Table 1: Tractable and intractable systems

	Tractable system	Intractable system
Number of details	Description are simple with few details	Description are elaborate with many details
Comprehensibility	Principles of functioning are known	Principles of functioning are partly unknown
Stability	System does not change while being described	System changes before description is completed
Relation to other systems	Independence	Interdependence

Resilience Engineering represents such an approach. Traditional approaches to risk and safety depend on detailed descriptions of how systems are composed and how their processes work in order to count 'errors' and calculate failure probabilities. Resilience Engineering instead starts from a description of characteristic functions, and looks for ways to enhance an organisation's ability to create processes that are robust yet flexible, to monitor and revise risk models, and to use resources proactively in the face of unexpected developments or ongoing production and economic pressures. Because socio-technical systems are incompletely described, hence underspecified, individuals and organisations must always adjust their performance to the current conditions in order to accomplish their objectives. Since resources and time are finite, it is inevitable that such adjustments are approximate. In resilience engineering, accidents and failures are therefore not seen as representing a breakdown or malfunctioning of normal system functions, but rather as representing the converse of the adaptations necessary to cope with the real world complexity.

Early Signs of Intractability: The Sneak Circuit

Although doubts about the traditional approaches to safety only have become widespread from the 1990s and onwards, the intractability problem had been known from the early 1960s – although by another name. On November 21, 1960, a Redstone rocket with a Mercury capsule was launched and began lift-off. However, after a 'flight' of a few inches, lasting a mere 2 seconds, the engine cut off and the vehicle settled on the launch pad. The escape tower rockets fired to separate the Mercury capsule from the rocket, which deployed the re-entry parachutes and landed 1,200 ft. away. The whole area was cleared for 28 hours both because the reason for the engine shutdown was unknown and to allow the Redstone batteries to drain down and liquid oxygen to evaporate.

Since this clearly was a both unexpected and unwanted event, every effort was made to find the effective cause. In the end it was realised that the rocket's engine cut-off because the tail plug that connected the rocket to the launch assembly was prematurely pulled out before the control cables. This created a 'sneak circuit' or 'sneak path' that caused the engine to shut-off. The existence of such sneak circuits made people realise that unwanted outcomes could occur

even when nothing went wrong. Indeed, sneak circuits were formally defined as "... conditions which are present but not always active, and (which) do not depend on component failure" (Hill & Bose, 1975). (It is interesting to note that the concern for sneak paths has reappeared to confront the so-called cyber threats, defined as 'access by adversaries who want to obtain, corrupt, damage, destroy, or prohibit access to very valuable information.')

Despite the open acknowledgement that unwanted outcomes not always were a consequence of failures, the mainstream of safety methods remained faithful to the assumption that things work until they fail, and further that they worked as intended. This assumption may be quite reasonable in the case of technological systems, which have been designed to work in a certain way, with as little variability (and with as much precision and reliability) as possible. But the same assumption does not hold for socio-technical systems.

Accidents and risks

Since accidents traditionally have been explained in terms of something that has gone wrong or failed, risks have naturally been seen in the same light. There is a simple and inevitable symmetry between accidents and risk: risks are accidents that may happen in the future, and accidents are risks that became reality in the past. Risk assessments has, for historical reasons, relied on representations of how failures – or malfunctions – can occur and on how they can combine, using representations such as fault trees and event trees. Considerable efforts have been invested in finding ways to determine and/or calculate the probability that something may fail or malfunction, and describing how multiple failures may combine.

Most of the methods for risk assessment and accident investigation that are used today have their origin in the 1960s. This was the period where the technical or engineering analysis methods were developed in response to the growing complexity of technological systems. As technical systems have grown in size, classical risk assessments have become harder to perform, both because the number of components and events to be considered has grown disproportionally and because the ways in which failures can combine and interact has become more difficult to describe. In other words, because systems have become intractable.

In the 1970s it was gradually realised that safety could not be considered in relation to technical systems alone but had to include the people working with the systems as well, not just as potential victims but more importantly as factors that could affect safety in either a positive and negative direction. This has resulted in the understanding that all technical systems fundamentally speaking are socio-technical systems. The TMI accident in 1979 made it obvious that human factors played a significant role in system safety, and that it therefore was necessary for risk assessment and accident investigation methods both to go beyond the technological system as such and to consider 'accidents without failures' such as sneak paths. The concern for the human factor was later extended to cover organisations and organisational factors as well, with the prominence of 'safety culture' as a good example. Since the mid-1990s considerable efforts have been made to answer the perceived need among theorists and practitioners of a re-orientation in thinking about safety, and to develop methods and approaches that are both more efficient in use and have better grounded concepts and constructs.

From causes to emergence

For systems that are tractable it makes good practical sense to associate risks and accidents with causes, to look for causes, to estimate or calculate their probability, and to calculate how

the consequences of failures can propagate. But for systems that are intractable, risks are emergent rather than resultant, and typically represent combinations of socio-technical performance variability rather than the consequences of failures and malfunctions. It is therefore necessary to deal with emergent risks in a different way. In technological systems (especially electronics and software), a few failures may be due to sneak paths but on the whole cause-effect thinking will suffice. In socio-technical systems, and in general in systems that are more complex or intractable, adverse outcomes can happen in the absence of failures or malfunctions and cause-effect thinking is therefore inappropriate. Indeed, for socio-technical systems, such 'accidents without failures' are likely the norm rather than the exception. For these it is necessary to think in a different way and to have access to different methods. The most important risks are not those that arise from single events but rather those that emerge from the underspecification that is a part of everything that happens.

Safe and unsafe functioning

It is reasonable to assume that any method is just about adequate for the problems that were typical at the time it was developed. Methods are developed to solve known problems, but not unknown problems. New methods must therefore be developed when the existing methods encounter problems for which they are insufficient or inadequate. Today this happens mainly because socio-technical systems continue to develop and to become more complex and more tightly coupled. The inevitable consequence of this argument is that even new methods after a while become underpowered because the nature of the problems change, although they may have been perfectly adequate for the problems they were developed for in the first place.

Today we have many methods that focus on unsafe functioning but few, if any, that focus on safe functioning. We spend considerable efforts on how to prevent unsafe functioning, but almost none on how to bring about and sustain safe functioning. We naively assume that the absence of unsafe functioning – or rather the absence of the conspicuous outcomes of unsafe functioning – means that we are safe. Yet the aim of safety should not only be to reduce the number of adverse events (the 'visible'), but also to improve the ability to succeed under varying conditions, to deal with the 'invisible.' This requires a more proactive approach, and that methods are developed to support that.

How to assess risk related to system characteristics relevant to IO

Jørn Vatn

Department of Production and Quality Engineering
Norwegian University of Science and Technology, Trondheim, Norway
jorn.vatn@ntnu.no

Abstract. In this paper risk is defined as uncertainty regarding occurrence and severity of undesired events. In order to improve understanding of the risk concept four different domains are defined, i.e., the real world, the cause and effect, the uncertainty and the value & preference domain . To express risk quantitatively uncertainty statements are expressed in terms of probability statements. It is emphasized that risk always is conditioned on various framing conditions. Here we emphasize the consequences dimensions established by dialogue and risk communication processes, the understanding and knowledge basis of the assessor, and the result of third party verification. With such a basis for risk assessment we especially discuss system characteristics like complexity and intractability which are words often associated with IO. An example is presented in order to discuss elements of the risk analysis process.

Introduction

Risk analyses are today recognized as an important element of safety management. However, there are claims that some "risks" never can be revealed by risk analysis, i.e., there are some escaping risks we need to approach by other methods than risk analyses methods. One aspect of IO is more use of opportunity based maintenance where many safety critical decisions are made more or less on the fly, hence system complexity and intractability is believed to increase making risk assessment almost impossible. In this essay I discuss this example and the implications on risk assessment.

For the principal discussion I distinguish between four different *domains* which are important when defining risk, and applying the result of a risk analysis.

1. ***The real world domain***. In this domain things are happening. Important attributes of the real world domain are events such as accidents or component failures, and states such as level of competence or the actual position of a valve. The real world attributes are generally denoted observables since they in principle are observable.
2. ***The scientific cause and effect domain***. This is the set of laws, models, methods, theories etc we use to establish relationships between observables in the real world. We differentiate between deterministic relationships such as the force of gravity, and probabilistic relationships in terms of tendencies or correlations, e.g., correlation between the number of fatalities and the average speed level on a road. We use the label *scientific* to emphasize that these laws and research results are general, and have been exposed to critique by the scientific society. We do not consider the cause and effect domain to be "true". The point is that there exist results and arguments that have been exposed to public critique, and the user of this knowledge will always be responsible for bringing it into his analysis in an adequate manner.

3. ***The uncertainty domain.*** This domain comprises all the uncertainty regarding the value of the observables in the real world domain, and the relevance and correctness of the scientific cause and effect domain. In a risk analysis context the uncertainty reflects that the *assessor* do not have certain knowledge i.e., uncertainty is not an inherent property of the real world. Probabilities are used to express uncertainties quantitatively.
4. ***The values and preferences domain.*** In order to make decisions regarding risk it is necessary to state how desirable or undesirable the various outcomes in the real world domain are. Such statements belong to the values and preferences domain.

I will here use the word risk entirely to express issues in the uncertainty domain. Although the phenomena in the real world domain and perspectives in the cause and effect domain are quite different, I argue that *the principal issue of risk will be the same independent of which system, or system characteristics that are analyzed*. This means that I will not *move* the risk definition to some of the other three domains.

Risk definition

Here, the term risk is defined with respect to use in risk analysis and risk assessment. This means that the term will be defined as it is proposed used in the risk analysis discipline, i.e., among risk analysis scientist, researchers and consultants within risk analysis. Further it is meaningful to distinguish between the conceptual content of risk, and an operational definition of risk.

Conceptually, risk can be defined as the *uncertainty regarding the occurrence and the severity of events*. There are three elements of this definition, (i) the uncertainty whether an event will occur or not, (ii) the event, and (iii) the uncertainty in the severity given that the event occurs. From a principal point of view there are three categories of uncertainty regarding the event aspect. First is the uncertainty regarding whether we are able to identify or foresee the relevant events. Next is the uncertainty regarding whether we are able to identify and structure the causes leading to the undesired event. Finally there is an aspect of “degree of uncertainty” which we express by probabilities. The borderline between the two first types of uncertainties depends on the level of analysis. If the “events” we are focusing on is very “late” in the course of events leading to an accident it is usually not a major challenge to identify the events, i.e., by focusing on traditional events like hydro carbon leakages, blow-outs, fires, explosions, loss of structural integrity etc it is hard to imagine “new accident types”. On the other hand if we let our “event” level in the risk definition correspond to the very first deviating events in an accident scenario it is rather obvious that the identification problem becomes much more challenging. Thus phrases as “some risks are escaping the risk analysis” need to be seen in relation to how risk is defined. I will later discuss the challenges related to identification and structuring in the “left” part of the risk scenario, i.e., related to the early phase.

In the Norwegian offshore oil and gas industry the concept Defined Situations of Hazard and Accident (DSHA. In Norwegian: DFU (definert fare eller ulykkessituasjon) has been introduced. A DSHA is a defined hazard or accidental situation. Although the introduction of IO brings along several concerns related to safety, we have not seen arguments that calls for new DFUs (Skjerve et al., 2008) This means that if the “event level” in the risk analysis definition is in the “right” part of the accident scenarios the identification problems would not be of any concern. We remain with the challenge to identify and structure those causes and patterns that are leading to the events in our risk definition.

The conceptual definition of risk considers risk as uncertainty regarding the occurrence and severity of events. This definition will be the basis for risk analysis, but we need a definition where we are able to express risk in quantitative terms. Probabilities are used to express uncertainty in quantitative terms. We use p as a likelihood measure of the occurrence of an event, say e. Here p is either a probability of occurrence in a time period, or a probability distribution function for the number of occurrences per time unit. Returning to the conceptual definition, p is thus representing the uncertainty regarding the occurrence of the event e of interest. We let **S** represent the severity of the event. Note that **S** is a multidimensional random quantity. This means that we treat several dimensions like personnel safety, environmental impacts, material damages, economical aspects etc. A probability distribution function is used to reflect the uncertainty in the actual value of the severity, e.g., the number of fatalities given the event e. Since there are more than one event to treat, we use i as an index to run through all relevant events to include in the risk definition, and an operational definition of risk is the set of all relevant triplets:

$$R = \{<e_i, p_i, \mathbf{S}_i>\} \quad (1)$$

The definition in equation (1) is very similar to the proposed definition of risk by Kaplan (1991). As pointed out by Aven (2009), Kaplan applies the probability of frequency interpretation with respect to the probability measure, rather than an epistemic[5] interpretation as suggested by Aven. Often the risk picture covered by equation (1) is compiled into more simple expressions, e.g., the expected number of fatalities in a time period, the f-N curve etc.

The risk expressed by equation (1) is a very naked expression where assumptions, what to focus on etc. is vanishing. The risk is always *conditioned* on some background information, theories, understanding etc. We let $\mathcal{U}$ represent the relevant information, the theories, the understanding, the assumptions etc. which are the basis for the risk assessor when risk is assessed.

It is often argued that risk is a one dimensional quantity which cannot capture the variety of aspects that are relevant in a decision situation. It is therefore important to avoid an oversimplification when a risk metric is established. For example focusing only on expected number of fatalities is often considered insufficient. On the other hand, introducing too many dimensions will make the risk analysis very complex and costly, and the result may also be hard to utilize in a decision context. Generally it is recommended to have various types of dialog processes among the stakeholders to focus on a reasonable set of dimensions to include in the analysis. We let $\mathcal{D}$ represent the result of such dialog processes and risk communication. Note that $\mathcal{D}$ is related to the *values and preferences* domain with respect to which dimensions to include in the risk metric.

A risk analysis of high quality shall have a list of assumptions. Next the analysis shall also identify those assumptions that are most critical to the result of the risk analysis. This will enable the decision maker to evaluate whether he should ask for other positions regarding $\mathcal{U}$. Although a risk analysis lists a set of assumptions, and documents the main elements of the calculations, it is usually impossible for the decision maker and other stakeholders to verify the "correctness" of the work, i.e. what is really $R \mid \mathcal{U}$? Third party verification is the instrumental mechanism to verify the correctness of the calculation given $\mathcal{U}$. Third party verification is also an instrumental mechanism to ensure that the relevant positions regarding

[5] The word epistemic is often used in connection to uncertainty when the uncertainty is related to lack of knowledge.

$\mathcal{U}$ are documented. We let $\mathcal{V}$ represent the result of the verification processes. The final principal definition of risk is then given by equation (2):

$$R = \{<e_i,p_i,\mathbf{S}_i>\} \mid \mathcal{C}, \mathcal{U}, \mathcal{V} \quad (2)$$

This operational definition of risk for use in risk assessment has the following strengths:

- The operational definition of risk expressed by equation (2) is consistent with the conceptual understanding of risk, i.e., uncertainty in occurrence and severity of events.
- The operational definition of risk is consistent with the way risk can be assessed in a quantitative risk analysis.
- The fact that risk is conditioned on $\mathcal{C}$, $\mathcal{U}$, and $\mathcal{V}$ will not guarantee, but to a great extent ensure that the positions of the relevant stakeholders are taken into account, that various disciplinary disagreements have been elaborated, and that the analysis is of high quality.

Risk assessment in an IO context

Since risk according to our definition essentially express uncertainty, the focus in risk assessment in an IO context should also focus on uncertainties. There are three elements of this uncertainty. The first category of uncertainty deals with the completeness problem, i.e., to which extent we are able to identify undesired events. A wide range of studies have been conducted to face any new safety problems with the introduction of IO. As far as we know (Skjerve et al., 2008)[6], none of these have revealed new accident types corresponding to the "event" definition used in the risk analyses. The second category of uncertainty is related to the causal analysis where the aim is to structure available knowledge regarding causes that may lead to the undesired event. A major challenge here is that we are not able to prescribe in detail all possible patterns leading to an accident. Thus we need to reflect in our uncertainty statements. The third category deals with the uncertain severity *given* the occurrence of an undesired event. As far as IO is concerned the structuring of uncertainty with respect to severity primarily deals with crisis handling in distributed teams. Important key words are shared situational awareness, but in this presentation I will not focus on that aspect of uncertainty.

It has been argued in more general terms that certain types of risks escape the risk analysis framework. Also when IO is considered phrases like "intractable systems", and "complexity" have been used as arguments for other approaches than risk analysis. A fundamental question is therefore whether the risk analysis framework has reached its limitation, or whether we may work within the risk analysis framework we have defined above taking the new challenges into account. This question can not be answered without working with the practical cases. In the following these challenges with the basis of one explicit concern that have been put forward in relation to IO will be discussed. The introduction of IO gives, among other benefits, new possibilities with respect to more or less on-line planning. For example in maintenance there is now a shift from fixed interval to opportunity based maintenance. Opportunities for maintenance execution "outside" the schedule may be due to shut-down on another installation, other failures on the actual installation etc. What will then be the safety impact of the resulting real time decisions regarding important safety issues? At a first glance we instinctively respond to such a change by stating that new risks may emerge. But what is an emerging risk? Since our understanding of risk is related to uncertainty, the relevant

[6] In this report new challenges are identified, but it is not pointed to any new event types to include in the risk analysis.

question would be whether there are new uncertainties and which types of uncertainties this could be.

Even though the implementation of IO will affect working practice, the technical system to control the energies will be the same. This means that we expect more or less the same technical solutions. Lack of protection may be classified into two types of failures, i.e., physical failures (degradation), and functional failures where the technical system fails to control the energies due to other factors than physical degradation. Current risk analysis approaches deals with both physical and functional failures. The introduction of IO is primarily assumed to affect *functional* failures with respect to controlling the energies. The challenge would be to map our knowledge regarding the introduction of IO with respect to the impact of potential functional failures. There are two concerns that are put forward in this context. The first one is the increased complexity anticipated with the work practices that are introduced by IO. This complexity or intractability may manifest in terms of problems we are not able to anticipate, and hence result in functional failures, and loss of control of the energies. However, the impact of these problems will be the same event types that we already are treating in the risk analyses, e.g., gas leakages. The second area of concern is the emergency handling after a crisis eventually occurs. In this presentation the focus is on the first concern, i.e., the new causes that may result in loss of control of the energies.

Since there are no indications of new accident types by introduction of IO, we would rather investigate whether there are new failure causes. In the last decade gas leakages have been put on the agenda in the offshore sector, so we will narrow the scope of this discussion to focus on gas leakages of pressurized hydrocarbon systems. Direct causes may be grouped into physical equipment failures such as a flange that is not tight, or a functional failure such as a valve left in wrong position. Investigation of historical gas leakages gives several examples of both categories. Our main concern with respect to the new work practice of IO would be functional failures. However, the challenge of functional failures is not unique for IO. A huge research activity is currently running to address these challenges, e.g., the OMT[7] project. To be more explicit on the gas leakage scenario we consider gas leakages which occur in relation to maintenance on process vessels. In order to execute such maintenance it is necessary to change the state of a range of process valves before the maintenance starts, and set these valves back to the original state when maintenance is completed. A valve in a wrong position may be very critical from a safety point of view. A dedicated blending and isolation plan is worked out to ensure that all valves are in the correct position during and after maintenance. A number of persons are involved with preparing this plan, which again relies on updated and correct documentation of the plants instrumentation (P&ID). We recognize that there are a large number of involved “agents” here, e.g., engineering staff responsible for initial and updated instrument plans, manufactures of the equipment, persons installing the systems, the documentation office etc. In principal there exist knowable relationships between the agents. Some of these relations have delayed effects, e.g., updated documentation needs to be approved after the changes have been implemented. It is therefore no surprise that e.g., a wrong version of the instrument documentation now and then is used as basis for the blending and isolation plan. Upfront it seems almost impossible to foresee all such undesired patterns which might become critical from a safety point of view. However, in retrospect it is often very easy to “explain” why it went wrong, e.g., a logical error in the version control system.

[7] See Risk_OMT Risk Modelling - Integration of Organisational, Human and Technical factors, www.preventor.no/omt

Introduction of IO is anticipated to enable opportunity based maintenance in large scale compared to today's practice. This will push the system even harder since the blending and isolation plan has to be prepared more or less "on the fly", or at least downloaded from a "prepared list" without any verification wrt changes in assumptions. Again, it is underlined that such a description relates to the real world domain, the list is in the real world, any errors in the list are in the real world etc. Such a system description would ideally be a consideration of all possible situations which would have resulted in an erroneous blending and isolation plan. But we recognize that the system is too complex to be described in detail, the system is only *retrospective coherent* in the words of Kurtz and Snowden (2003).

Now as we have some understanding of our example in the real world domain, we would return to the question of risk and probabilities in the *uncertainty domain*. At a first glance it seems reasonable to argue that the system complexity prevents us from assessing the risk. But we have to be very precise regarding the interpretation of risk. Risk is essentially uncertainty, and for our discussion we restrict the discussion to the uncertainty in whether the blending and isolation plan is correct or not. It is rather obvious that we deal with uncertainty. We are not able to state with certainty whether the isolation and blending plan is correct or not, basically due to the complexity of the problem. But this complexity in the real world domain does not necessarily translate to a complex probability assignment in the uncertainty domain, where probability is used as a measure of uncertainty, as seen through the eyes of the assessor. The literature provides a range of methods that seems relevant in order to assign the probability figure of interest.

We will elaborate on the approach being developed in the OMT project. Here a set of so-called risk influencing factors (RIFs) are introduced to modify a baseline probability of interest, e.g., the probability of an erroneous isolation and blending plan. To develop the model we need at least to consider (i) the importance of the RIF (weighting), (ii) the status of the RIF (scoring), (iii) interaction with other RIFs (typically negative synergy effects), and (iv) the impact on "common cause failures" between safety functions (or barriers). One relevant RIF here would be the quality of the plant documentation. Another RIF would be time pressure in the planning. The time pressure RIF would be very important as opportunity based maintenance is anticipated to be more common. It is reasonable to claim that time pressure will increase the probability of failure in the isolation and blending plan. However this depends on many other factors like competence, accessibility to documentation etc. A typical argumentation would go as follows: *By a detailed investigation of the correctness of the isolation and blending plans used offshore we have found an historical average failure probability of 5%. Further since company X seems to have a first class document office, the failure probability is set to 4%. The adjustment is based on the investigation of seven erroneous isolation and blending plans. On the other hand company X will utilize IO benefits such as more use of opportunity based maintenance. Although the isolation and blending plan according to procedures shall be prepared in advanced for maintenance activities to be executed on opportunities, the possibility to verify e.g., changes in system configuration is believed to be reduced. Hence failure probability is increased from 4% to 8%.* Although we would like to see a more comprehensive list of arguments, e.g., inclusion of more than one RIF, more elaboration on how the experience data have been interpreted and used, the above statements basically summarizes the knowledge and understanding available for the risk assessor. Note that the claimed increase in complexity as a result of introducing IO did not reveal new failure causes at the level we find it reasonable to treat in this analysis, i.e., error in the isolation and blending plan. Complexity has been revealed as an implicit potential cause

of errors, hence we find it more likely that “something” unexpected will happen compared to existing practice. The result from this limited study need to be integrated in the total risk picture, and we refer to the OMT project for further details..

Conclusions

The essay has discussed aspects of assessing risk in an IO context. Risk is defined as uncertainty regarding occurrence and severity of undesired events. By investigating anticipated effects of IO we do not anticipate new event types, but we foresee challenges with new failure causes. Complexity due to real time decision with safety implications is anticipated. There are reasons to believe in increased level of “intractability” compared to traditional operation. However, our example demonstrates that this increased complexity and intractability represent uncertainties which we may structure within our risk analysis framework. We are able to structure the problems in terms of failure causes. At some level we are no longer able to structure the failure causes by means of explicit cause and effect relations, the system will be too complex to describe explicit in detail. However, system complexity and intractability do not imply impossibility or meaningless risk assessment. By focusing on, and structuring uncertainty an approach is outlined by indirect arguments via e.g., risk influence factors, where a risk picture is obtained that is valuable for decision making. As for all type of risk analysis it is of vital important to stress the assumptions and understanding behind the numbers. This also applies for risk assessment in an IO context. The example in this essay has not revealed new principal aspects related to IO which calls for methods “outside” the risk analysis framework set up as far as risk assessment is concerned. One single example does not prove that other characteristics of IO meaningfully may be treated by such a risk analysis framework. To increase our insight we need to study the various system characteristics one by one.

How to address the complexity issue in IO Risk Governance

Tor Olav Grøtan

SINTEF Technology and Society, dep. of Safety Research, Trondheim, Norway
tor.o.grotan@sintef.no

Abstract. IO Risk Governance implies an activity of making sense of a (complex) systerm that is itself sensemaking and adaptive. This paper investigates how the notion of complexity can constitute a difference (of risk assessment) that makes a difference (for risk management):. Renn's (2008) typology of knowledge is criticized and elaborated by use of the Cynefin framework (Kurtz and Snowden 2003). This provides the grounds for a finer granularity of Risk management options, as well as a contribution to a more differentiated understanding of Resilience (Engineering) as a risk management strategy.

Introduction

This essay will use the notion of risk in a broad sense that reflects the view that our conception of risk is influenced of our *conception of* the system's characteristics, e.g. according to Renn (2008) and Figure 1 below, in which pre-assessment is foundational to risk governance acitivities. For a reader that prefers to reserve the notion of risk (assessment) for an expression of uncertainty that is more detached from the system characteristics, e.g. according to J. Vatn's essay (pp.17-23), the notion of "risk" can be replaced by the notion of "hazard".

The essay attempts to address the systemic risks (hazards) that may emerge from IO in terms of failure, disturbance or collapse in densely coordinated action across organizational boundaries. A premise for this approach is taking the liberty to reinterpret IO as "*Integrating* Operations", that is, a virtually never ending quest for tighter coordination of action and improved effectiveness and efficiency in concerted activities at various levels, e.g. work processes, work flows and value chains. In contrast, the "official" term of Integrat*ed* Operations connotes the existence of some final degree of integration being the goal of IO.

IO on the NCS is not a singular case of industrial or organizational transformation. Integration is the "holy grail" of ICT-driven organizational transformations at a global scale, and there is clear empirical evidence that such transformations experience a lot of unexpected trouble and unintended effects (Hanseth and Ciborra, 2008) that can be attributed to complexity as an overarching theme. There is thus an urgent need to address the systemic IO contributions to risk, in the sense of emerging patterns of (coordination) failure, and as unexpected propagation of effects and consequences. It is assumed that a meaningful discussion can be conducted based on such a general conception of systemic risk, with a backdrop of steadily increasing scope of control, detached from the limitations of time and place.

According to Renn (2008), systemic risk denotes the "embeddedness" of risks in a larger context of social, financial and economic consequences, and increased interdependencies both across risks and between their various backgrounds. Paying attention to systemic risk thus implies relatively less focus on intended behavior of isolated targets and defined deviances from norms, and correspondingly more awareness of possible unintended effects and propagations.

The RIO framework (Figure 1) derives from the IRGC Risk Governance framework (Renn 2008). As highlighted in terms of its multiple arrows, in the risk assessment phase various types of risk may be analyzed and characterized on different analytical grounds, with the important implication that various options for risk management must be considered. Some risks may even escape analytical grasp, but must be managed nevertheless. This essay aim to shed light on how systemic risks due to complexity can be differentiated from other risks in terms of risk assessment, and subsequently how these differences make a difference in terms of options for risk management. *It does not claim that the IO risk at large is solely a matter of complexity, but that complexity is a difference that makes a difference for risk management.*

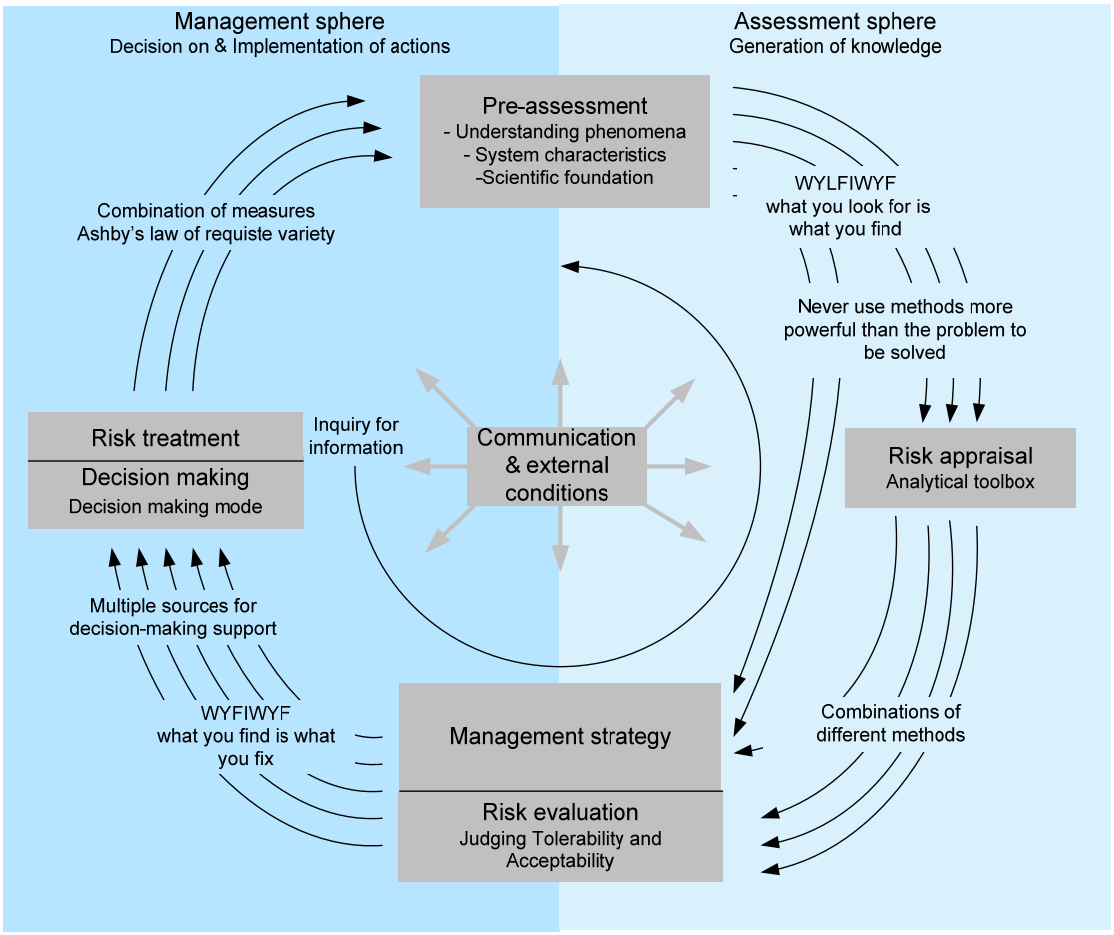


Figure 1: RIO Risk Governance Framework (Albrechtsen, Mostue & Hovden, 2009)

Renn (2008) offer a classification scheme for risk assessment based on the notions of complexity, uncertainty and ambiguity in order to create a corresponding set of options for risk management. Renn's scheme is based on a reduction strategy of which the first move is an attempt to reduce complexity into uncertainty. Complexity is hence regarded as something that can be reduced by means of analytical effort. However, there are other conceptions of complexity of which complexity defies or resists such analytical reduction. Whether the residual complexity should be denoted complexity or something else, say uncertainty or ambiguity, is a question that is much more than a play with words. It touches upon some very fundamental assumptions about how far we can understand complex systems – and *emergent* phenomena in particular, and thus how we can deal with them in terms of risk assessment and

management. Such assumptions and distinctions may imply an additional kind of difference in terms of risk management options, and their legitimization. The main purpose of this essay is thus to employ an alternative conceptualization of complexity to the task of investigating whether (risk assessment) differences make any difference for risk management. Renn's approach will hence be re-interpreted according to an alternative framework, along with some other contributions to the issue of complexity.

Complexity as wildness-in-wait and un-order

Chesterton (1909) made the following observation decades before ICT became a prominent technology for organizational transformation, and a century before "IO" emerged on the scene:

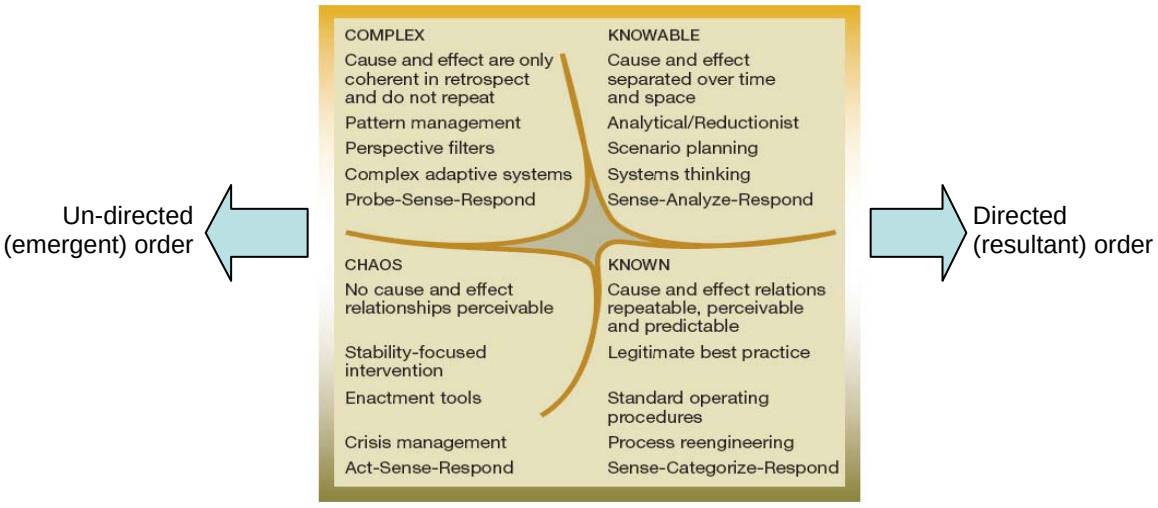
"*The real trouble with this world of ours is not that it is an unreasonable world, nor even that it is a reasonable one. The commonest kind of trouble is that it is nearly reasonable, but not quite. Life is not an illogicality; yet it is a trap for logicians. It looks just a little more mathematical and regular than it is; its exactitude is obvious, but its inexactitude is hidden; its* ***wildness lies in wait***" (this author's emphasis)

Considered as an expression of complexity, such a *wildness-in-wait* cannot be reduced by analytical effort, nor can it be clinically separated from the logical order it is embedded within. If we take a closer look at contemporary systems and contexts that are similar to IO, we will find it hard not to recognize that human beings cognize, interpret, make sense of, act on forecasts, look for meaning and purpose; thus they act with an agency that is not necessarily corresponding with the "official" or original purpose of a system. The normality of the "irrational" organization was described by Etzioni (1986), the limits of (bounded) rationality was described by Simon (1947), and the "Garbage Can" decision model convincingly presented by Cohen et al. (1972). Add the psychological insights into e.g. group dynamics, and the purpose of the system presents itself as complex. It should thus not be difficult to acknowledge the inherent *contextual complexity* of IO constituted by the fact that humans are not limited to having one identity, not limited to acting in accordance with predetermined rules, and not limited to acting on local patterns (Kurtz and Snowden, 2003).

In a risk governance perspective, addressing an uncertain future, the challenge is to make sense of a complex system (that is in itself sense-making of nature), based not only on available knowledge and information, but also based on a *reflexive* attitude to the *underlying assumptions engaged* in interpreting that body of knowledge and information. The *Cynefin* framework (Figure 2 below) by Kurtz and Snowden (2003) can be used to capture the mixed ontological and epistemological dimensions of the sensemaking situation. According to which, complexity, surprise and "wildness" may be a feature of the world as such, but may also stem from our habitual, but unrealistic expectations of stable, repeatable and "rational" patterns.

Cynefin distinguishes between directed (resultant) order residing on the right side (knowable or known), while undirected (emergent) order resides on the left side (complex or chaotic). The Cynefin domains reflect different options for system understanding and comprehension; being *Known* in terms of predefined categories, *Knowable* in terms of analyzable cause-effects relation-ships; *Complex* in terms of being (only) retrospectively coherent; *Chaotic* in terms of no cause-effect relationships perceivable. The four domains also implies different styles for maintaining control; responding categorically; responding according to system analysis; responding according to probing (cautiously interfering with the system) and making sense of the system's reactions; and acting (intervening) in order to enact and enforce stability by coercion an imposition.

Cynefin offers holistic premises at different levels. First, by distinguishing between the four domains that carries unique premises of unity. Second, by claiming that there may exist a higher-level order denoted *un-order* which unifies the "vertical" divide between directed and un-directed order. Un-order is thus a term conveying a paradox, connoting two things that are different, but in another sense the same. Un-order challenges the distinction that any "order" which is not directed or designed, is invalid or unimportant. Un-order thus paraphrases Chestertons description of the wildness-in-wait, embedded in apparent order, by incorporating recent chaos and complexity theory. This implies that we cannot reveal wildness (un-directed order) by rejecting (directed) order at large. However, while order can be instituted and verified by managing discrete events or scenarios of such events, potential un-order may only be spotted as patterns at a systemic level.



Cynefin (Kurtz and Snowden, 2003)

Figure 2. Un-order according to Cynefin (Kurtz and Snowden, 2003) : the wildness within order

Differences that make a difference

How can Cynefin be used for the purpose of assessing differences that make a difference? At an overall level, we can see that the assumption of directed order corresponds nicely with classical risk analysis, as risk in a Known system may be expressed in probabilistic terms on the basis of experience with similar systems, and that risk in a Knowable system may be based on defined scenarios comprised of known parts to which probability and consequence may be assessed. With the assumption of directed order, the notion of "wildness" must be reduced to the case of improper categorization or analysis. With the assumption of emergent order however, the notion of "wildness" addresses the case of a systemic breakdown, the presence of chaos (no cause-effect relationships perceivable), or a (new) order *emerging* without explicit intention, e.g. as a result of human adaptation. For risk assessment, the assumption of emergent order will also have implications for defining possible outcomes, and not at least their probabilities.

Let us now try to re-interpret the approach of Renn in terms of Cynefin. Renn (2008:74) state that the act of appraising risks and concerns is best described by the terms *complex*, *uncertain* and *ambiguous*. These three challenges are not related to the intrinsic characteristics of hazards or risks themselves, but to the *state and quality of knowledge* available about them. According to Renn (ibid:75), *complexity* refers to the difficulty of identifying causal links

between a multitude of potential causal agents and specific observable effects. Hence, complexity requires sophisticated modeling, which often defies common-sense or intuitive reasoning. Yet, if resolved, it produces a high degree of confidence in the results. There is nothing about this description that connects with the underlying assumptions of Complexity (un-directed order) in the Cynefin sense. Rather, the underlying assumptions match those of a Knowable system, in which only directed order is assumed. The focus on results defying common-sense or intuitive reasoning is actually something that connotes directly to the System Dynamics way of thinking. Some proponents of System Dynamics explicitly links this field to Cynefin, and points at the Knowable domain as the most suitable (Maani and Cavana, 2007:21).

Uncertainty is according to Renn (ibid) different from complexity, but most often results from an incomplete or inadequate reduction of complexity in modeling cause-effect chains. Renn emphasizes that human knowledge is always incomplete and selective, and, thus, contingent upon certain assumptions, assertions and predictions. He asserts that the means for distinguishing between the key components of uncertainty may be e.g. target variability, systematic and random error in modeling, indeterminacy or genuine stochastic effects, differences in system boundaries, or sheer ignorance or lack of knowledge. None of these can be linked to the (Cynefin) premise of emergent order, hence Renn's uncertainty concept must be interpreted merely as difficulty (or impossibility) of penetrating the Knowable system. Finally, *ambiguity* is strictly understood by Renn as giving rise to several meaningful and legitimate interpretations of accepted risk assessments results, and can be of the interpretive or normative kind.

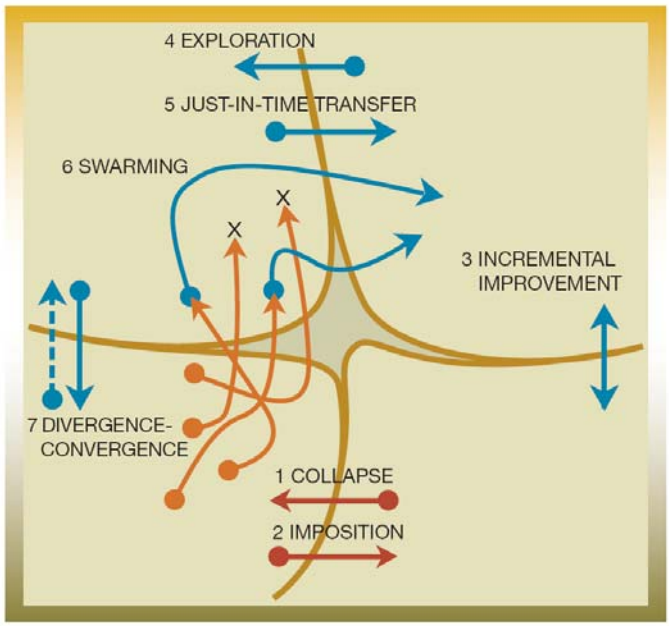


Figure 3: Cynefin Dynamics (Kurtz and Snowden, 2003)

It should be noted that both Renn's approach and Cynefin are based on the assumption that there is always a cause behind an observed effect. Cynefin opens up for the possibility that the causality is emergent, that playing the game changes the rules, so that the system is only retrospectively coherent (that is, attributable to novel cause-effects relationships), while Renn hesitates to move farther ahead when meeting the signs of uncertainty and ambiguity at the

end of the road for a priori causal modeling. It might therefore be argued that Cynefin is more sensitive or receptive to the possibility of systemic risks due to emergence.

An equally important divide in relation to risk governance is that assuming directed order suggests one set of options for risk management, while assuming the possibility of emergent order suggests a different set. It is interesting to note that Renn (2008:182-3) recommends a *robustness* focus on risk management when dealing with complexity-induced risk problems. That is, aiming for a risk-absorbing system based on improved coping capacity and "high-reliability" organization. In dealing with uncertainty-induced problems, Renn recommends a *resilience*-focused risk absorbing approach, based on improved capability to cope with surprises as well as preparedness for adaptation. In Cynefin terms, robustness and resilience can hardly be associated with the Knowable domain (where we claim that Renn's notion of complexity resides), as they connote more directly with the Complex domain, or at least visits to this domain (as in "exploration" and "just-in-time transfer" in Figure 3.

Resilience (Engineering) is also a distinct discipline in contemporary safety science (Hollnagel et al, 2006). It can be argued that Resilience (Engineering) can be interpreted in at least two ways as a risk management strategy. First, as a way of actively preparing to deal with exceptional situations (connoting to exploration and just-in-time transfer, Figure 3). Second, as a way of ensuring proper functioning under various conditions, ultimately corresponding to the mastery of divergence/convergence at the "edge of chaos", or "swarming" as a way of rescue (Figure 3). In the second case, resilience is justified by the aspiration of dealing with Perrows (1984) Normal Accident Theory, which can be interpreted as the more fatalistic claim that the combination of high interactive complexity and tight coupling enforces a move from Knowable to Complex, with the "normal" eventuality of falling into Chaos (divergence) with no ticket home.

Cynefin also encompass a diversity of risk management strategies within a reflexive modernization perspective. Paraphrasing Giddens (1990), safety science development is conduced by an overall process of knowledge spiraling in and out of the universe of industrial protection, continually reconstructing both itself and that area as an integral part of that process. There is no stable (social) world to know, but that knowledge of that world contributes to its unstable and mutable character. Safety science at large have had a progression starting with asking the question what (which component, technical or human) can *fail*, proceeding with the question of what can *go wrong* despite absence of failure, and proceeding with what may *go right* despite surprise and unexpected variation in system behavior (Resilience Engineering). In Cynefin terms, we thus see a pattern of a "modern" process of discovering, when scientific results "impregnates" practice, that fundamentally new questions may have to be asked in order to find feasible answers. Resultant order thus carries the embryo of (new) undirected order, however not only at the institutional level. Nathanael and Marmaras (2008) argues that we should use a similar perspective in addressing organizational resilience as a work practice, focusing on the (in Cynefin terms, un-ordered) dialectic between practice and prescription. This approach also resembles Hepsø's (2006) focus on *articulation work*.

Hence, we see that complexity in the Cynefin sense provides a richer set of interpretations of complexity, and a broader range of risk management options than its alternative (Renn, 2008). Is also hints very strong to the implication that risk management is neither on the directed nor the un-directed side of order. As they are intertwined in the higher-level concept of un-order, they might be mutually dependant in order to work. That is, in the face of complexity, neither

the traditional risk management “hard-liners”, nor resilience, can provide success solely on its own terms.

Using Cynefin as a sensemaking device requires the acknowledgement of risk governance based on a social process of knowing. It will be useless of risk governance is short-circuited into an exclusive relation between a prerogative risk analyst and a sovereign decision maker. And, it signifies some very critical managerial and organizational challenges of actually implementing resilience as a risk management strategy.

How can different perspectives on organizational accidents and resilient organizations be utilized in quantitative risk assessments?

Jørn Vatn

Department of Production and Quality Engineering
Norwegian University of Science and Technolgy, Trondheim, Norway
jorn.vatn@ntnu.no

Abstract. The paper reviews various perspectives on organizational accidents and resilient organization in order to support risk assessment. Previous research has identified certain so-called control questions to each of the reviewed perspectives. In the paper we discuss explicitly how these control questions supports the background knowledge for assessing risk. Risk is here to be understood as the uncertainty regarding occurrence and severity of undesired events.

Introduction

The aim of this essay is to link various perspectives on organizational accident and resilient organization to the risk assessment framework. First I define risk as uncertainty regarding the occurrence and severity of undesired events. Quantitative risk assessment is a structured way to identify undesired events, and quantify the uncertainty regarding occurrence and severity of these. I adopt the following operational definition of risk[8]:

$$R = \{<e_i,p_i,\mathbf{S}_i>\} \mid \mathcal{K}, \mathcal{U}, \mathcal{M} \quad (1)$$

Where e_i are the identified undesired events, p_i are likelihood measures quantifying the uncertainty regarding the occurrence of e_i, and $\mathbf{S}_i$ describe the uncertainty regarding the severity given that e_i occurs. I use the notation $\mid \mathcal{K}, \mathcal{U}, \mathcal{M}$ to emphasize that a quantitative risk measure always is conditioned on various aspects. Here $\mathcal{K}$ represent the result of various dialogue processes necessary to agree, or at least discuss various metrics to include in $\mathbf{S}_i$, $\mathcal{U}$ represents background assumptions, understanding, information etc, and $\mathcal{M}$ is the result of third party verification, or as a minimum "checked by" if third party verification has not been undertaken. It is important to emphasize that risk essential is uncertainty regarding occurrence and severity of undesired events, and to make this definition operational I use equation (1).

In this presentation I will discuss six perspectives. The five first perspectives are presented by Rosness et al. (2004) and the sixth perspective is presented by Hollnagel et al. (2006). Grøtan et al. (2008) present these perspectives in order to shed light on IO, where a set of control question for each perspective is introduced. I will utilize these perspectives and control questions in a more formal risk assessment approach. When control question are listed in the following the basis is always Grøtan et al. (2008)

The energy and barrier perspective

The baseline for this perspective is that energies or hazards are controlled by various barriers. To some extent the safety is increased by adding more barriers to the system, i.e., *defence in depth*. Although the term 'barrier' is used to describe this perspective, I will introduce a

[8] See Vatn's essay: *How to assess risk related to system characteristics relevant to IO (pp. 19-25)* for further discussion on the conceptual definition of risk, and how the conceptual definition is taken into an operational definition

broader concept and use the term ‘safety critical function’(SCF) rather than barrier. An SCF is a function for which a failure or deficiency in the function will increase the risk level. I use the word “critical” to underline that there is a significant increase in risk. The SCFs are structured in three levels, i.e., primary, secondary and tertiary SCFs. The primary SCFs are those functions we need to control the immediate hazards such as energies, toxics etc. The secondary SCFs are those functions necessary to ensure that performance of the primary SCFs is maintained over time. The tertiary SCFs are management functions necessary to reveal new hazards, drifting issues etc. As an example the emergency shutdown system provides a primary SCF. Maintenance in terms of e.g., functional tests and replacement of wear parts represent a secondary SCF. Maintenance management in terms of follow-up of maintenance backlog, root cause analysis of problems revealed during maintenance etc. represent tertiary SCFs.

The energy and barrier perspective only focuses on primary SCFs. Thus, this perspective fails short to reveal a realistic risk picture due to ignorance of secondary and tertiary SCFs. Very many quantitative risk analyses conducted in the design phase of an installation focus on the primary SFSs due to lack of knowledge regarding secondary and tertiary SCFs. In relation to the definition of risk given by equation (1) this would correspond to $\mathcal{U}$ expressed in terms like ’given that reliability parameters will be the same as the average in the industry’ or similar.

If we allow the energy and barrier perspective to consider SFS on all three levels this perspective will capture most aspects which I will discuss in the perspective the follows next. I refer to such an approach by *the extended energy barrier perspective*.

The control question for the energy and barrier concept is “*Are barrier functions strengthened or weakened*?” The control question is formulated with IO aspects in mind, i.e., there is an anticipation of change in one or the other direction. In my discussion I will for this control question and the other questions not focus on changes, but rather the “baseline”. Here this means that the important aspect is the function of the barrier. What is then meant by the term ‘barrier function’ as used here, or in my framing a safety critical function. The control question introduces the adjectives *strengthened* and *weakened*, and we may ask whether it is possible to “measure” a barrier function. In reliability theory we measure a function by the failure rate, where failure refers to lack of function, and rate is a probability measure. It is important to distinguish between physical failures and functional failures. A physical failure refers to physical degradation, e.g., fatigue resulting in a breakage. Functional failures refer to errors in the functional description or implementation of a function. For example if a valve is left in wrong position during maintenance there is no *physical* failure, but the *valve function* is not fulfilled. I will not discuss quantification within this perspective any further, since this has been the main focus within risk analysis for decades.

The information processing perspective

This perspective is very well captured in the following control question: “How is an organization’s ability to utilize safety-relevant information, observations and ideas, regardless of position and status of the person or group who possess this, influenced by the IO development?”

The aspects captured by the control question obviously affect the safety level. From a quantitative risk analysis point of view the challenge is to relate these aspects to probabilities. The natural approach would be analyze the various accident scenarios one by one to find links

to information processing aspects. Again I refer to the extended energy barrier perspective where secondary and tertiary SCFs were introduced. I claim that the information processing perspective primarily related to these two levels in the SCF hierarchy. Since it is hard to establish binary cause and effect relations between the information processing aspects and the performance of the SCFs a risk influence modelling (RIF[9]) approach seems reasonable. As an example I outline an approach for a specific case. Inspection programs are implemented to ensure integrity of structures on oil and gas production platforms. These programs are established based on theoretical models e.g., corrosion growth, but also on plant specific knowledge related to e.g., temperature, pH content in the fluid etc. During inspections failure progression may deviate significantly from the expected value. This often leads to further investigation regarding what are the causes. Such an investigation will usually result in a modified inspection program for that installation. Often such information is not distributed to other installation which may face the same problems. In this context the information processing aspects are of vital importance. Given such knowledge, the quantitative modelling by means of SCFs and RIFs is rather straight forward. Elements of such an approach would be to identify the importance (weights) of the RIFs and the status (scores) for the RIFs

The decision-making perspective

In this perspective the main focus is on the handling of conflicting objectives. Other aspects of this perspective deals with situations where activities are migrating toward the boundary of acceptable performance. A third aspect would be covered in the notion of distributed decision-making which seems important in an IO setting.

Two control questions are put forward for this perspective, the first one reads: *"Is an organization's ability to make decisions concerning risk strengthened or weakened?"*. We note that the control question addresses decision making in a general context, rather than focusing only on handling of conflicting objectives. Decision making occurs on all levels related to the hierarchy of SCFs, but need different treatment in a risk analysis setting. We discuss an example with increasing backlog in the maintenance. To catch up with the maintenance both economical recourses are required, but also execution of maintenance will cause production disturbances. One aspect when studying decision making is to investigate the governing documents regarding how to control handling of the maintenance backlog. If these documents are very precise this will be a god basis for efficient decision making. However, if the governing documents are unclear, this will make decision making difficult. Another aspect would be to check how decision making actually is conducted. By such an approach we may both measure premises for good decision making, but also actual decision making practice. Next we need to link decision making practice to the risk. For the maintenance backlog this would be rather easy since it is relatively easy to establish a link between the risk level and the size of the backlog. Bad decision making would usually result in lack of control with respect to the backlog.

The second control question reads *"Is the likelihood for drift towards the limit for acceptable risk increased or reduced?"* This control question is linked to secondary and tertiary SCFs. Especially tertiary SCF focusing on identification of new problems, and the control of existing problems. The maintenance backlog example exactly points at such challenges. The maintenance backlog represents a situation where the hazard is known since we know that a

[9] See e.g., Risk_OMT Risk Modelling - Integration of Organizational, Human and Technical factors, www.preventor.no/omt

long maintenance backlog represents higher risk. In order to develop a risk model we need to take at least the following factors into account: What measures are implemented to reveal "new areas of drift"?, How well are drift monitored?, and How well is the organization dealing with drift problems? Quantification follows standard RIF modelling.

The theory of Normal Accidents

Normal Accident theory focuses on two important system characteristics, high interactive complexity (non-linearity) and tight coupled systems. Systems with high interactive complexity are difficult to control, not only because they consist of many components, but also because the interactions among components are *non-linear*, e.g., feedback loops. The control question of relevance here is "*Does the IO development create complex interactions and tight couplings between system elements?"* I argue that complexity as such is not a threat to a system. Therefore an approach would be to analyse the system, and the risk scenarios to be explicit on where complexity is seen as a real threat. Next we need to be explicit on what are the negative consequences of complexity. In order to make a precise discussion I recommend to split into complexity as seen from the analyst, i.e., he is not able to foresee relevant combination of factors that forms a threat, and complexity as seen from the operator in the sharp end. Problems in the sharp end translate to lower success probabilities in those models that are used to describe essentially crisis handling. Vatn discusses complexity as seen from the risk analyst in an example with failures introduced during maintenance as a result of more opportunity based maintenance[10].

The theory of High Reliability Organizations

This theory is grounded in intensive studies of organizations that have demonstrated an outstanding capacity to handle fairly complex technologies without generating major accidents. Important concepts from this research tradition are organizational redundancy and a capacity of organizations to reconfigure in adaptation to peak demands and crisis. This is also reflected in the following control questions "*Is organizational redundancy strengthened or weakened?*" and "*Is the ability to change operational modus in crisis situations strengthened or weakened?*" Note that organizational redundancy not only points towards redundancy in terms of e.g., two units performing the same function. Organizational redundancy also relates to the situation where two or more units interacts to increase reliability of the function by e.g., utilizing different competence in order to "solve the problem better". This means that if we try to *measure* organizational redundancy we cannot only focus on counting units, we need to understand interaction between units. As for the case of complexity we do not believe that organizational redundancy as such is critical. The objective of the risk analysis would be to identify areas where organizational redundancy is important. One aspect to "measure" would be to which degree the organization allows asking "silly questions", like "I am responsible for maintaining this unit, but I do not really know how to do it". This again indicates that when part of the organization is in doubt, it is possible to ask for a second opinion, a key feature of organizational redundancy.

The second control question relates to the ability the organization has to change operational modus in crises situations. Many risk analyses are rather static, hence it is hart do incorporate change of operational modus in these analyses. However, if dynamic analyses are developed it is possible to model explicit the change of operational modus. Two important aspects need to

[10] Vatn, op.cit.

be covered, the first one is whether the organization realize that a change is needed. The term 'shared situational awareness' is often used in this context to underline that need for a change must be recognized by more than one part of the organization. Next we need to consider how "well" the organization is able to operate in a changed modus. I will not go in further details here, but once again emphasize that in order to incorporate the last control question we need dynamic risk analyses.

The theory of Resilience Engineering

The term 'resilience' is often used to indicate that a system is able to maintain operation in a new state when disturbances occur. Another similar term would be 'robust', whereas robust more emphasizes the system capabilities to withstand disturbances, and not so much focus on ability to maintain operation if disturbed. A second aspect of resilience, or more specific in the emerging discipline of resilience engineering, is captured by the following control question: *"Is the ability to expect and be prepared for the unexpected strengthened or weakened?"* This expands the original definition of resilience to also cover the ability to operate if *unexpected* disturbances occur. Three elements seem reasonable to include if resilience is to be translated to risk. The first element relates to safety management elements described by tertiary SCF, or more specific to which degree the organization is able to keep on identifying new threats during the entire life cycle. This relates to the *awareness* of the organization. The second aspect would be permanent preparedness measures like supply vessels *just in case* a man falls overboard. The final aspect would be the capability of the organization to adapt or respond to deviations, or variation in normal operation. As for many of the other perspectives resilience is not necessarily an important feature of an organization. In order to quantify the effect of resilience it is necessary to structure the scenarios where resilience is believed to be a key factor. And in particular which SCFs are really dependent on resilient behaviour.

Conclusions

I have been investigating eight control questions that are used to capture important aspects of various perspectives. Although these control questions have their origin from different theoretical perspectives, the control questions may be seen as "stand alone" questions. In my opinion these questions may be seen independent of the theoretical background. Obviously it will be valuable to investigate the corresponding theory to get more insight, but may be not critical. In this essay I have studied these control questions in order to investigate whether they points at issues which may be incorporated in quantitative risk analysis. Although I do not present a comprehensive theory on how this may be done, I indicate elements that need to be put in place if such an approach is desired. Whether it is of value to incorporate these control questions in a quantitative risk analysis will as always depend on the need for decision support. The reason why we generally want to asses the impact on various factors or conditions on risk is the realization that even though there are positive effects, there are also negative effects. To be explicit on uncertainties involved is then of value in my opinion.

Evaluation of safety methods and tools supporting decision making in IO - based on HRO and Resilience Engineering

Siri Andersen

Department of Industrial Economics and Technology Management
Norwegian University of Science and Technolgy, Trondheim, Norway
siri.andersen@iot.ntnu.no

Abstract. Due to changes introduced with Integrated Operations (IO) it is possible that methods and tools for risk and safety assessment are also affected. This assay uses IO properties as well as criteria based on High Reliability Organisation (HRO) and Resilience Engineering (RE) to suggest a framework for evaluating methods and tools. Methods and tools are described as supplementary decision information, aiming to give valuable safety information regarding risks involved in decisions and actions. The overall evaluation criteria made are, for HRO; organisational redundancy, reconfiguration and mindfulness, and for RE; adjustability.

Introduction

There are a great variety of methods and tools in the 'toolbox' for risk and safety assessment used in the oil and gas industry today, some of them founded in regulations and standards while some are company specific. Experience and methods well adapted and developed for handling safety threats in traditional operation have during the years resulted in stable conditions with few major accidents and personnel injuries. A major goal would be to not jeopardize this stable picture when integrated operations (IO) are introduced. It is therefore necessary to check whether the toolbox will still be valid, and to decide what changes to make to keep activities at an acceptable safety level.

The aims of this essay are to give some criteria for how the theories of High Reliability Organizations (HRO) and Resilience Engineering (RE) can be used to evaluate methods and tools for risk and safety assessment in IO. The essay starts by making a brief categorization of IO properties/elements to be used in the method evaluation. Then relevant actors are given and the system is described and delimited. The essay uses Rasmussen's (1997) socio-technical system as a starting point for delimiting IO actors and system description, and to shows how methods for risk and safety assessment are related. Finally, indications are given for how HRO and RE can be used in the evaluation of methods/tools, and a brief discussion of their relevance and use in further work is given. It is a pre-assumption that the theories of HRO and RE are relevant for IO.

Properties/elements describing IO

With the transition from traditional operation to IO there will be changes in the properties/ elements that describe the system. One can talk about both old (deleted) elements, unchanged elements and new elements (figure 1). Together the unchanged and new elements will constitute the whole package of properties/elements describing the socio-technical system in IO. These elements/properties will (as illustrated in figure 4) be an important basis for the assessment of methods and tools. The deleted, unchanged and new elements are described as follows:

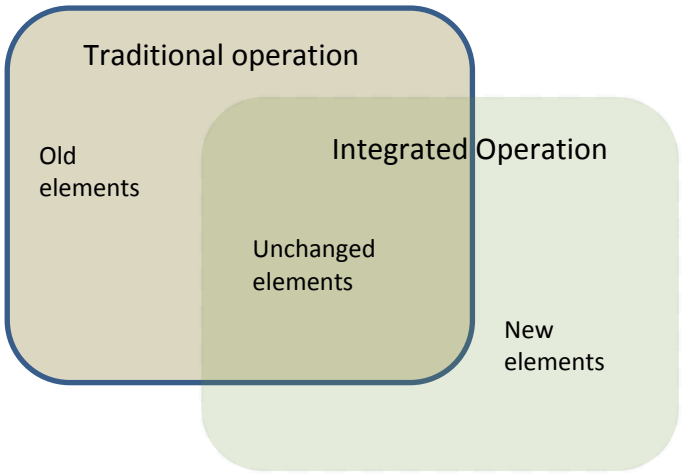


Figure 1. Old (deleted), unchanged and new elements in IO

- ***Unchanged elements:*** even though IO introduces many new elements to design and operation there are still a lot of unchanged elements (compared to the traditional fields). Unchanged elements are those which to a little degree are affected by the IO changes. These unchanged elements will still need to be treated regarding risk and safety, and for these it is assumed the tools in the traditional toolbox will still be valid. Examples of such elements are: the offshore structures (e.g. pipelines and templates) and protection of these, operators working and performing maintenance at the installations (close to hazards), technical installations offshore and safety barriers in these (e.g. pumps, valves, engines, compressors etc).
- ***New elements:*** these are elements/properties introduced by IO that were not present in traditional operation, i.e. as a new element in itself or an element that affects the traditional systems. These elements are especially interesting and critical since they might not be coved by methods and tools at all, and further result in threats we are not able to uncover. Examples of such elements are: access to more real time data, better access to expertise, changes in communication, more operational information available onshore, more actors involved in decision making (operators, contractors and subcontractors) etc, (Grøtan & Albrechtsen, 2008).
- ***Old elements:*** these are elements that are present in traditional operation but not in IO. Such elements are treated by the traditional toolbox, but do not need to be covered by the IO toolbox. No examples are given since I am not even sure there will be such elements in IO.

When evaluating methods/tools the new and unchanged elements introduced by IO will be studied.

Approach towards actors involved

As a starting point Rasmussen's (1997) description of socio-technical systems is used. Rasmussen divides actors into six main categories (Work, Staff, Management, Company, Regulators and Government) and highlights how these are all involved in decision-making in

Risk Management to control the hazardous processes. Low risk operation depends on proper co-ordination of decision-making at all levels.

Hazard scenarios with particular hazard sources and control requirements are given at the bottom. Further the diagram puts focus on information flow among actors and how these are used as input in judgment and decision making. Information about actual state of operation is propagating upwards, information about objectives propagating downwards. In addition there is information supplied sideways. This is additional safety related references and criteria communicated from external sources, such as technical descriptions from design, safety information and preconditions for safe operation from risk analyses.

Rasmussen's socio-technical system is useful to illustrate the approach taken in this essay toward actors and their involvement in decision processes in IO. See highlighted areas in figure 2.

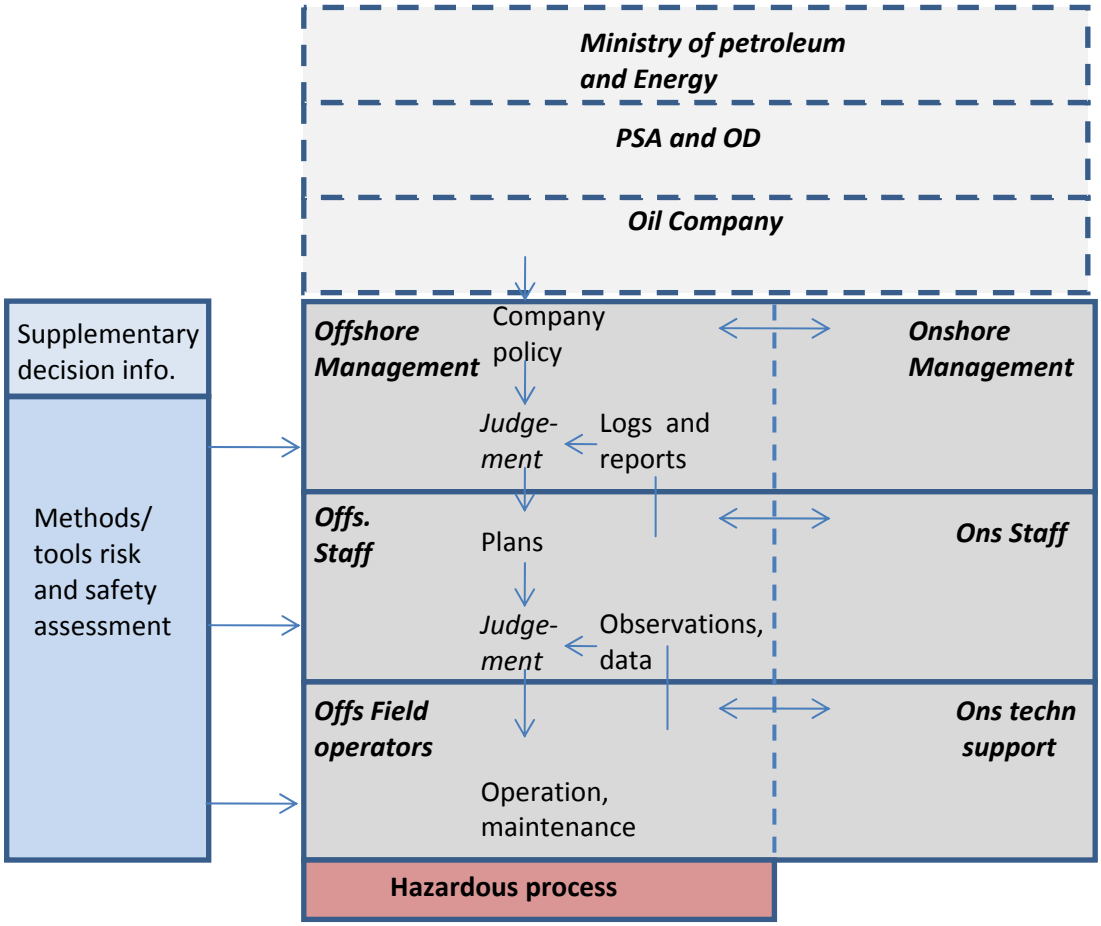


Figure 2. Actors and decision processes in IO. Adapted from Rasmussen (1997)

Both the offshore and onshore organization is included, since they in IO will be more integrated and involved in daily decision making. Information flowing upwards, downwards and sideways will be more shared among these two. Methods/tools for risk and safety assessment are chosen as focus area for information added sideways. Mark that even though

onshore operators are becoming more involved in operational decisions, it is still only the offshore operators that are directly exposed to the hazardous processes.

System approach for assessment of methods and tools

The previous section highlighted actors, information flow, decision making in IO and how safety methods and tools are supposed to support decision making. A major component of collaboration and decision making in IO are the collaboration centres, which connects onshore and offshore on a more continuous basis. Figure 3 includes the collaboration centres and gives a final system description of actors and processes involved.

When evaluating methods and tools it is important to keep in mind that these should give useful and supportive safety information to decision makers. Not only about hazards directly related to the offshore technical processes, but hazards impost by processes in the socio-technical system, including both onshore and offshore. *The goal is therefore to evaluate whether methods/tools are able to give valuable safety information and promote knowledge of risks involved in decisions and actions (for all actors).*

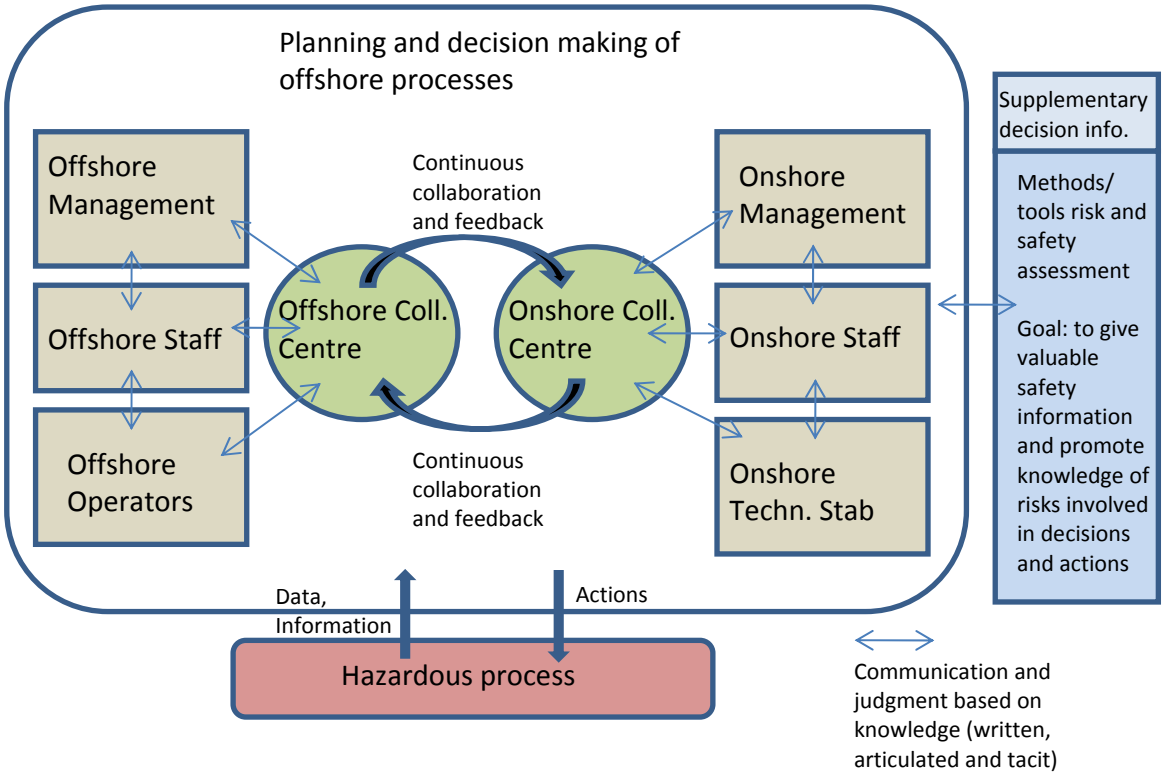


Figure 3. System description for evaluation of methods and tools

Figure 3 is general but can be extended and further detailed depending on the actual case studied. It is possible to include more collaboration centres (both operator and contractor), and to specify actors and their involvement in the collaboration centres.

Relevance of HRO and RE

There are several established research perspectives on risk and accidents. The two chosen here, HRO and RE, are particularly developed to describe systems like the ones introduced by IO. They can therefore give valuable contributions to the evaluation process of methods and tools suitable for IO. Figure 4 illustrate how they are used as input in the evaluation. Note however that the actual relevance of HRO and RE depends upon whether properties/elements describing IO are similar to system descriptions in HRO and RE.

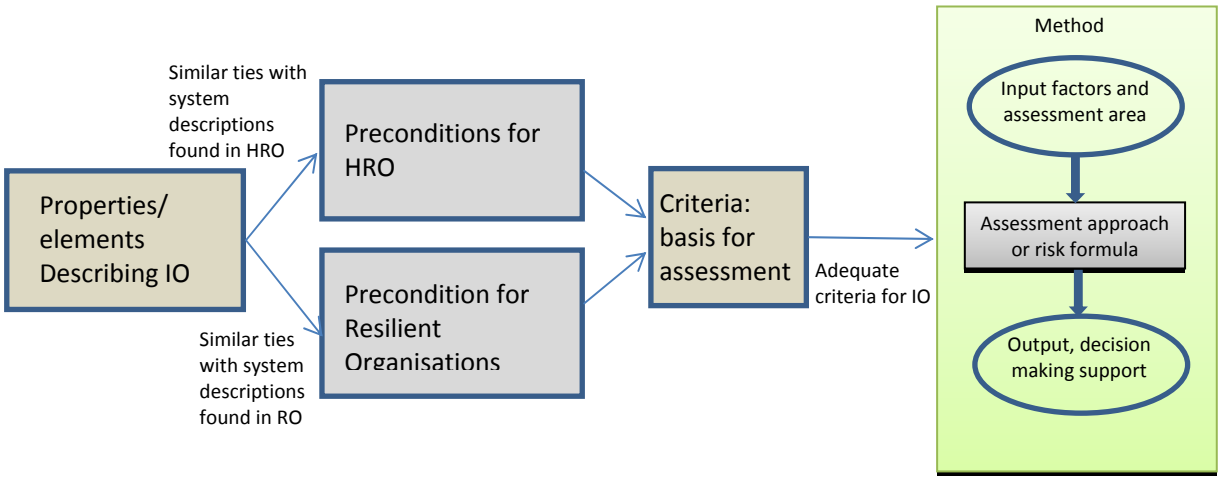


Figure 4. Approach towards assessment of methods and tools based on HRO, RE and properties of IO

High Reliability Organisations

The idea behind HROs is that some organisations manage to stay stable and avoid accidents even though they operate in a complex environment with the potential to cause major accidents. LaPorte and Rochlin (1991) explained such organizations by active use of redundancy to recover errors and derive reliable performance. Rosness et al. (2000) termed this capability '*organizational redundancy*', and divided it into two parts: 1) structural/instrumental preconditions and 2) cultural preconditions. LaPorte and Rochlin (1991) also described HROs by the ability to '*reconfigure spontaneously*' during demanding situations and crisis, by changing into a more flexible and resilient pattern. Weick and Sutcliffe (2007) introduced '*mindfulness*' as a prominent characteristic. A mindful organization is able to detect errors and contain them at an early stage. They track small failures, resist oversimplification, remain sensitive to operation, and takes advantages of shifting locations of experts. The three main evaluation criteria used for HRO are: organisational redundancy, reconfiguration and mindfulness.

Resilience Engineering

In resilience engineering failures and successes are treated as closely related phenomena. Both failures and success are the outcome of normal performance variability, and safety cannot be achieved by constraining that. A resilient system is defined by its ability to efficiently adjust its functioning prior to, during or following changes and disturbances. In order to be in control (be resilient) it is necessary to know what has happened (the past), what happens (the present) and what may happen (the future), as well as knowing how to respond. Hollnagel (2009) defines the four cornerstones of resilience, each representing an essential system

capability: 1) the *actual* – being able to respond to disturbances, regular and irregular threats, 2) the *critical* – being able to flexibly monitor what is going on, including the systems own performance, 3) the *potential* – being able to anticipate future disruptions, pressure and consequences, 4) the *factual* – being able to learn from experience, both successes and failures. Woods (2006) explained resilience by thorough monitoring of system boundaries to avoid making decisions that fall outside the safety boundaries. The following contributing factors of resilience were given: buffering capacity, flexibility/stiffness, margin and tolerance. Woods also points to the four 'I's for RE: being independent, involvement, informed and informative. The main evaluation criteria used for RE are adjustability (responding, monitoring, learning and anticipating).

The criteria developed for HRO and RE are summarized in the table 1. A cross is used to mark areas where they are *partly* coinciding. The table gives a *theoretical* basis for important criteria in HRO and RE, which can be selected to the final criteria basis for evaluation of methods/tools (illustrated in figure 4). The final selection will however be based on actual relevance for IO.

Table 1: HRO and RE assessment criteria for method evaluation

RE \ HRO		Organisational Redundancy		Reconfiguration	Mindfulness	
		Structural/ instrumental precondition	Cultural precondition		Detection	Containment
Adjustability	Responding			x	x	x
	Monitoring	x	x		x	
	Learning	x				
	Anticipating				x	

Discussion

Prior to table 1 two separate tables (one for HRO and one for RE) were made, each of them describing more detailed the factors of each assessment criteria. These detailed tables were used to indicate coinciding elements (crosses in table 1). In parallel to this essay an extended article is being developed that also include the two more detailed tables, and show the relationship to table 1.

HRO and RE uses both coinciding and differing factors to explain respectively high reliability and resilience. HRO put emphasis on persons and team's possibility and capability to cooperate, observe and intervene, as well as organizational flexibility, redundancy and error/symptom attention. RE does not put that much emphasis on cooperation and overlapping tasks, but includes several issues not directly found in HRO. Such as: awareness of system performance and boundaries, continuous and revised monitoring of performance and changes, extended and continuous learning (both successes and failures), available resources both for predefined events and future emergent threats etc.

In further work it is possible to use the above criteria for HRO and RE as a theoretical basis for selection of final evaluation criteria for methods/tools, both those already in use in IO and other possible methods/tools. The final criteria should however be selected based on particular relevance for IO. A premise is also that the properties/elements describing IO are similar to system descriptions found in HRO and RE.

It is probably correct to assume that there is no single method that nicely and efficiently supports all criteria for HRO and RE. It is neither not a goal to find this ‘one method’, but rather to evaluate which methods/tools are able to support parts of the properties of HRO and RE. And to see whether is possible to find a package of tools that together support reliability and resilience in IO.

Further, reliability and resilience in organisations are probably not possible to achieve by ‘just’ performing a package of methods/tools. The follow up, implementation of results, distribution and presentation of information in the organisation, and involvement of actors are probably just as important.

And finally, it can be assumed that there are properties in HRO and RE that are not at all possible to embed in method/tools, but need to be handled in a more holistic way outside the toolbox.

Characteristics of decision-making processes within integrated operations, and implications on risk management

Bodil Aamnes Mostue and Eirik Albrechtsen

SINTEF Technology and Society, dep. of Safety Research, Trondheim, Norway

bodil.aamnes.mostue@sintef.no
eirik.albrechtsen@sintef.no

Abstract. Integrated Operations (IO) changes decision settings and decision making processes that might influence the major accident risk. This essay gives examples of decision settings and decision-making processes with IO according to level of authority and proximity to hazards. The essay discusses boundary surfaces and responsibilities, proximity to hazards, information overload and interaction between decisions. A series of questions regarding major accident risk and decision-making are raised

Introduction

Better, safer and more efficient decisions are one of the major expectations of the IO development. It can be argued that integrated operation lead to better and safer decisions by real-time data and detailed understanding of situations; interdisciplinary teams involved in the decision making; access to expertise; and parallel activities (Ringstad and Andersen, 2007). Decision-makers and decision-making processes are influenced by various constraints in the decision settings. Will IO generate contexts that influence the major accident risk in a *negative* way? In this essay we highlight some possible negative impacts on decision processes that influence the major accident risk, by raising a series of questions. The elaboration is related to a typology of decision setting defined by Rosness (2000).

A typology of decision settings

Rosness (2000) distinguishes between different decision settings based on two underlying dimensions: 1) proximity to the hazard and 2) level of authority, see Figure 1. Proximity to hazard is conceived both in physical terms and in causal terms, i.e. the number of intervening links in the causal chains between the decision maker's actions and potential accidents. We can thus distinguish between actors at the sharp and the blunt end.

Description of the decision settings (Skjerve, Kaarstad and Rosness, 2009:13):

- *Operations* refer to sharp end settings such as control rooms.
- *Business Management* refers to settings for high level decision-making in enterprises, such as company boards, and decisions made by executives and senior managers.
- *Administrative and Technical Support Functions* refer to settings towards the blunt end with limited formal authority, such as design, risk analysis and handling of routine cases by a regulatory authority.
- *Political Arenas* are assigned the task of handling decision-making involving conflicting interests. Political arenas include local councils, governments, parliaments, and EU institutions.
- *Crisis Handling* refers to settings where important values, such as human lives, are exposed to an imminent threat which requires prompt action.

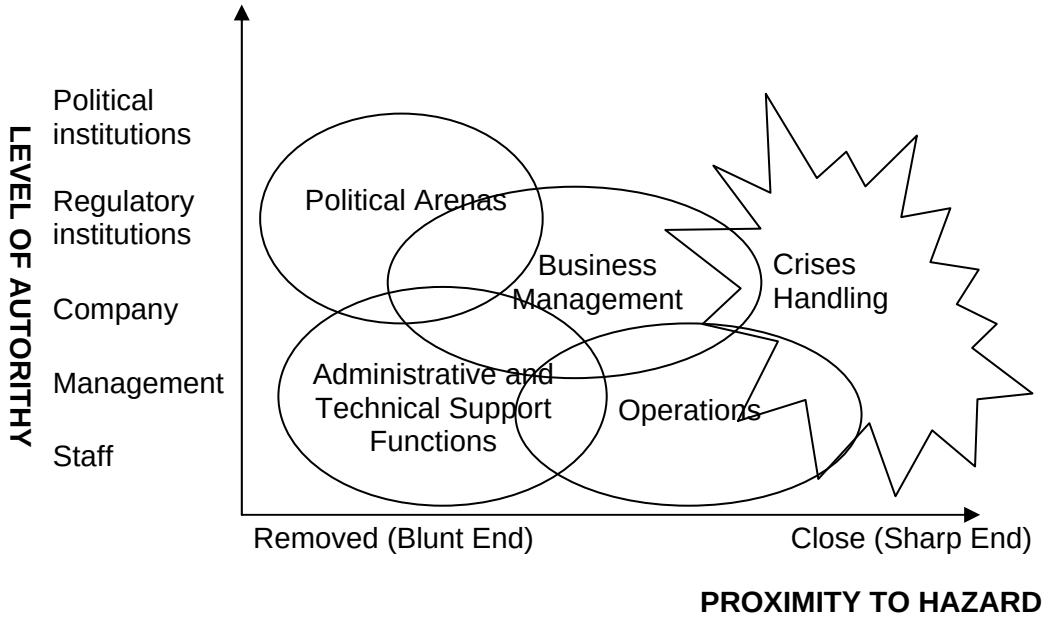


Figure 1. A typology of decision settings (Skjerve, Kaarstad and Rosness, 2009).

Table 1 Characteristics of five decisions settings (Skjerve, Kaarstad and Rosness, 2009).

Decision Settings	Dominant Constraints	Dominant Decision Criteria	Potential Problems
Operations	- Workload - Limited situation awareness	- Smooth and efficient operations - Acceptable workload	- Slips - Missing warnings - Local rationality, ignorance about side effects - Safety margins may erode
Business Management	- Information processing capacity - Dependence on information filtered by subordinates	- Optimise profit (or other KPIs (key performance indicators)) - Avoid trouble - Efficient decision-making - Ensure commitment or compliance	- Recycling of ineffective solutions - Reliance on simplistic indicators - May face strong incentives to run a risk
Administrative and Technical Support Functions	- Limited hands-on-knowledge - No authority to enforce decisions	- Comply with rules and standards - Consistency - Optimise a single attribute	- Unrealistic assumptions. - Unrealistic models.
Political Arenas	- Conflicts of interest - Changing constellations of power	- Robust consensus - Secure status of decision-maker	- Inconsistency over time - Decisions not followed up by actions. - Safety margins may erode in the absence of strong watchdogs.
Crisis Handling	- Stress - Time to obtain information and act	- Avert catastrophic outcomes - Avoid extreme stress levels	Defence mechanisms may lead to defective coping if danger materialise

Decision settings with integrated operations (IO)

IO makes it possible to make decisions over long distances, between companies, over different cultural boundaries and across time zones. What are the IO-implications on the different decision settings, as described in figure 1 and table 1? We will in the following limit our discussions to a low authority level, i.e. the decision settings 'operations', 'administrative and technical support functions' and 'crisis handling'.

Figure 2, shows that in IO, there are many actors who contribute to and make decisions that can influence the major accident risk in the Norwegian oil and gas industry. The figure illustrates that in an IO context, there is not only one support centre involved in decisions, but several.

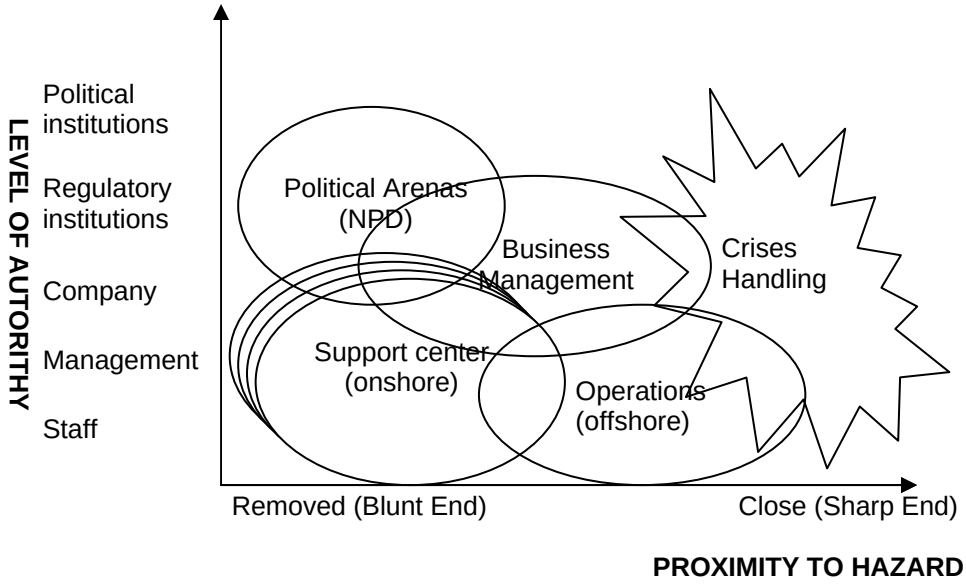


Figure 2. Decision settings with IO in the oil and gas industry (based on Rosness (2000)).

As a result of several actors involved in different decision settings, one should consider the figure as 3-dimensional. The horizontal axis can be divided into two different proximities: 1) proximity to hazard in casual terms (i.e. the number of intervening links in the causal chains between the decision maker's actions and potential accidents) and 2) physical proximity to the hazard. This is illustrated in figure 3. Rosness (2000) has a main emphasis on the first, interpreting proximity as whether the actor's decision can directly lead to an accident or not (i.e. the "z-axis" in figure 3).

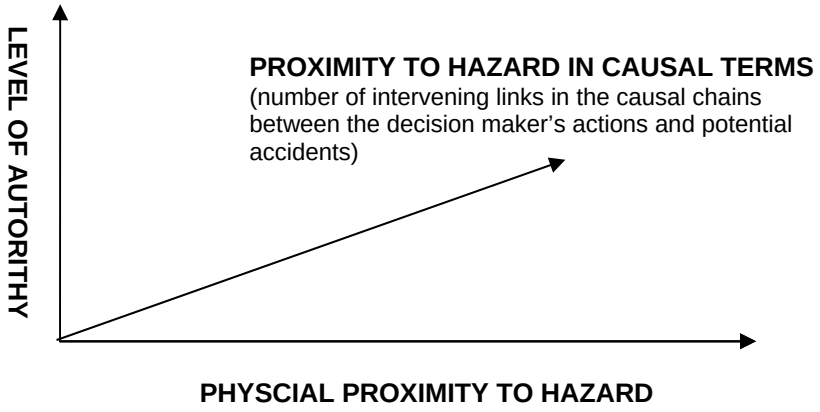


Figure 3. 3-dimensional diagram, showing two different proximities to hazard. One related to physical proximity to hazard, and one related to causal terms

Figure 4 shows how the proximity dimension can be divided in these two new dimensions. By these two dimensions, the importance of physical proximity comes into consideration as an addition to Rosness' figure. The black arrows illustrate interactions in decision-making processes. As seen in the figure, onshore support centers are physically removed from the offshore hazards, but can be both close and removed in terms of proximity in causal terms.

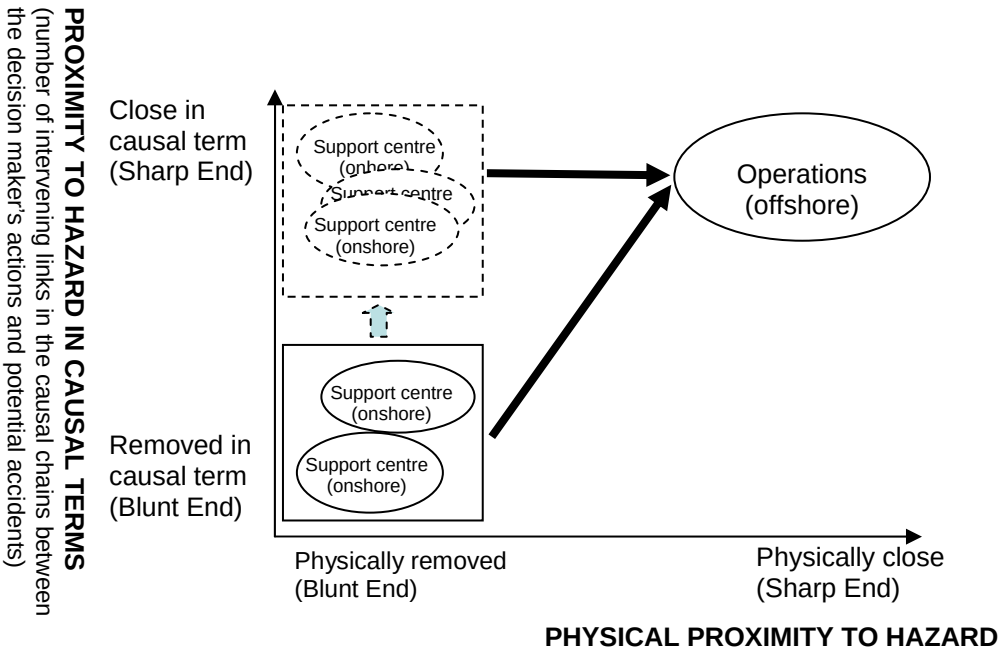


Figure 4. Decision settings, proximity to hazard in physical and in causal terms, and information lines in decision-making processes with IO. The figure shows actors at low level of authority.

Based on these figures, we will in the following discuss: boundary surfaces and responsibilities; proximity to hazards; information overload; interaction between decisions.

Boundary surfaces and responsibilities

Interaction between many actors both inside and outside an operating company might have positive effects on the major accident risk e.g. by improved awareness and access to expert support. At the same time team-based and distributed decisions might result in diffuse responsibilities. Fragmentation of companies and specialization might result in lack of people with knowledge on the overall system. Different actors can have limited knowledge of a system and information of a limited part of a problem. Many actors involved (onshore and offshore, departments in a company and operating company, contractors and subcontractors etc.) might have an impact on distribution of responsibility, communication and emergency preparedness (e.g. responsibility in connection to defined hazard and accident situations). The following questions of relevance to risk management can be raised:

- How do operating companies ensure unambiguous responsibilities and lines of command in a situation of many distributed actors?

- How is major accident risk managed when many actors are involved (e.g. operator and suppliers)?
- What kind of tools are needed to identify and assess risk related to unclear responsibility in boundaries between different actors?

Physically proximity to hazard

Reduced knowledge of local factors might influence both decisions made at distance from the drilling and production facilities and decisions made by "nomad life"-personnel i.e. use of semi-skill workers without belonging to any specific installation (contractor/subcontractor). Distributed decision-making processes will depend on situational awareness; the interaction between actors (collaboration, communication, learning, management); how individuals interact in networks (relations, power/influence and confidence); and culture (values, norms, knowledge, identity).

These elements are related to being familiar with persons, equipment, conditions etc involved in the decision-making process. Such knowledge is more difficult to get if there are many distributed actors to deal with, and easier to get through informal interaction such as small-talks.

Another aspect is difference in working hours onshore and offshore. "IO is functioning with "planned decisions" where we now solve problems in collaboration rooms where actors see each other, something that is well functioning. Related to "immediate decisions" IO does not function as long as the onshore support is not 24/7" (Petroleum Safety Authority, 2009:29).

- Does physical proximity to hazards still matter for decisions related to major accident risk in an IO context, or can all decisions regarding major accidents be made independent of location? Are there decisions where one needs to stop and think rather than emphasizing "faster, safer and better decisions"?
- Will differences in working hours onshore and offshore be a critical element regarding management of major accident?
- Do support centers obtain more decision-making authority (part of the command line) rather than supporting offshore decisions (illustrated by the stippled ellipses in figure 4)?

Information overload

Large amount of real time data has to be analysed and presented in a meaningful way at the operation centres. In addition, the persons on the installations have to make sense of this information. There is thus loads of information being transferred between sharp and blunt end. This possible information overload may result in absorption of uncertainty. Persons at the installations might spend less time close to the sharp end (physically) as a consequence of increased demand of shared information and communication from the drilling and production facilities to operation and support centres.

- How does one ensure that uncertainty and limitations in decisions is not accumulated in ways that influence the major accident risk? (e.g. limited awareness on uncertainty of information and advices from different sources)
- Can information overload be a threat to safety?

Interaction between decisions

One decision often influences another decision and big decisions are composed of a chain of decisions. Decisions might interact in ways that influence the accident risk. Examples of such a pattern of interactions are: distributed decision-making and local optimization, meta-decision (decisions on how to make decisions), absorption of uncertainty and normalization of deviance (Skjerve, Kaarstad and Rosness, 2009). More work performed in a parallel fashion might result in faster decision-processes and failure to see unwanted interactions. The safety margins might erode with increasing number of actors (hierarchy of companies, contractors/subcontractors).

Interaction between decisions has to be paid attention, and also decisions where many actors and trade disciplines are involved.

- How can we ensure that decisions made by different actors do not interact in an unexpected manner?
- How can we avoid meta-decisions to impose goal conflicts on lower level decision makers?
- How can we avoid absorption of uncertainty and possible consequences as uncertainty and value of flexibility to be underestimated and robust options to be ignored or rejected?
- How can we avoid that deviant events interpreted as normal?

Conclusion

By the introduction of IO, new settings for decision-making and decision-making support occur. The new decision-making context can influence the major accident risk. In this essay a series of questions regarding major accident risk and decision-making are raised. The most important of these seem to be:

- How do we ensure clear responsibilities and lines of command in a situation of many distributed actors?
- Does physical proximity to hazards still matter for decisions related to major accident risk in an IO context, or can all decisions regarding major accidents be made independent of location? Are there decisions where one needs to stop and think rather than emphasizing “faster, safer and better decisions”?
- How can one ensure that decisions made by different actors do not interact in an unexpected manner?

With regard to risk management a question relevant for all challenges related to decision-making occur: *do we have the adequate risk assessment methods and risk management approaches to cope with the new challenges related to decisions and IO?*

Mission IO: Change, supermodels and dreams of efficiency

Fred Størseth

SINTEF Technology and Society, dep. of Safety Research, Trondheim, Norway
fred.storseth@sintef.no

Abstract. The essay issues a warning and calls attention to the possible side effects of IO. By use of 'psychological goggles', the essay formulates an educated guess regarding how IO as a kind of organizational change strategy by use of Resilience Engineering may produce unintended and adverse side effects. These are by-effects that may hit at the operational level (personnel) and propagate to become a safety threat at organizational level.

Introduction

Integrated Operations (IO) creates opportunities and possibilities for the petroleum industry, no doubt. Although the selling points are compelling, we cannot forget that IO is business as usual. Core drives are faster, better, cheaper – all wrapped in, and justified by science based arguments of economy, decision making, technology, and safety. IO may in this way be viewed as an industry specific variation of what contemporary work life is all too familiar with: organizational change. Ah.. that old song I hear some of you say.

Before we open that can of worms however, let us in all fairness take stock of what we are looking at: For researchers and consultants alike, IO has been a steady paycheck for years already. Pros and cons are being elaborated and sought balanced, warnings are both issued and debunked (and so on). In parallel to this analytic activity, the suspense is by now long gone in some companies – as IO – in different shapes and sizes has been part of life for quite some time. On top of all this however, there is still enough "newness", and – dare I say uncertainty left, as IO for others is still a state of the future. In this picture, we find "followers" and "opponents" of IO. Although either-or is a caricature description, there seems to be differences in terms of how much IO is sought after; by researchers, consultants, and the industry itself. This diversity may certainly be seen as a good thing, as it seems to propel a continued search for ideas and solutions. One specific set of ideas that is mentioned in relation to IO is resilience.

Resilience Engineering (RE) (Hollnagel et al., 2006) sails up as a interesting set of ideas; perhaps particularly by the fact that RE seems to fit the IO scope – which is vast (!), e.g. system level premises, ICT-solutions, decision-making, power, social interaction. As IO faces challenges relating to collaboration and decision making across distance and organizational borders, the RE ideas regarding adaptation and intractable systems emerge as powerful analytical tools. Although the degree of RE ideas actually being "designed in" various IO-versions remains to be seen, it is here to be argued that it is due time for a WARNING in terms of possible side effects. This warning should apply both for the industry buying the recommendations as well as the research communities and consultants producing them.

Although it is hard to predict the future the current essay aims to provide an educated guess based on psychological research findings regarding possible side effects of organizational change. The simple premise for this guess is thus to consider IO and the possible implementation of RE principles into the industry as any given change strategy.

Side effects and perceptions of change

In psychological research, negative outcomes of organizational change are thoroughly documented. The obvious side effect that was documented early on was that job loss (due to downsizing) was associated with negative health outcomes. However, another robust finding relates to the harmful aftermaths of change for those who "survived" the downsizing round, i.e. the so-called "layoff survivors" (de Vries and Balazs, 1997). This survivor-effect illustrates how the *experience* and *perception* of organizational change may produce negative outcomes stretching far beyond the actual change process. In fact, there may not even be actual change going on. This is when some get irritated and points out that the only thing that is interesting is to look at actual, objectively true change, and that everything else is just subjective rambling.

The answer to this kind of reprimand is quite frankly: no. From a psychological perspective, what is interesting is precisely the subjective experience and perception of reality. Another thing is of course, and this should come as no big surprise, peoples perceptions are most often correlated with reality.

So, what is the point of bringing all this up? The point is that, from a psychological standpoint, the demarcation point of change is hard to pin down. Put differently, when a given change begins – and ends is hard to pin down as we are dealing with people's perceptions and experiences.

In order to formulate the announced educated guess, we have to elaborate on the issue of *perception of change*. This can be illustrated by looking at another protest against the psychological approach. In addition to the aforementioned critics that are calling for the objective truths of the world, there is another vein of opponents. Basically these are typically the ones that emphasize that change is constant and necessary, and that it is no value in trying to hamper this natural state of affairs. "Things change, deal with it!" This mantra is often thrown into discussions that bring possible negative side effects of organizational change to the corporate table. Shouting out that change is constant is an effective way to disarm any attempt to pinpoint that change management *may* produce outcomes beyond increase in production and revenues. It works well as a way to defuse because; in a sense they are correct in their observations. That is, (organizational) change is very much the ever present force. What is more: change is necessary, both for business organizations and for human beings. In pompous terms one may say that without change and dynamics there is no life. As the well-known opposite of life is death, the majority would probably stick with the former.

Sarcasm aside, it is fair to say that both of these lines of argumentation are highly theoretical, almost abstract in nature. It is here to be argued that it is in this sense that they become accurate. That is, from a distance and kept on a table-top strategic plane it becomes easy to stand by both of these banners (i.e. that only actual change should count, and that change is a constant). Moving in closer to organizational reality however, it may quickly become evident that boundaries of change are muddled and difficult to distinguish. Perhaps the most irritating for both opposing camps (the 'keepers of the objective truth', and the 'change is constant believers') is the fact that they both can be credited as being right!

In a way they are both using pile drivers on open gates. Of course it is relevant to look at actual change, and yes, the extent of change and restructuring in the work life of today most certainly justifies the label 'constant'. But: they are somewhat missing the point. The crux of the matter is that people's experiences and perceptions of change matter; and they seem to be

related to negative side effects. In psychological terms then, an individual may well experience and be affected by change well after the master plan has been rolled out and the consultants have left the building. Or, the opposite, the individual may experience and be affected by change before any actual change activity is implemented.

Basically, perceptions of change and the almost constantly perceived threat of another round of change has been linked to mental health complaints, stress/lack of control, and reduced well-being in general. What is more, and something that should be reflected on in relation to IO: These employee experiences and reactions has been found related to negative safety outcomes (e.g. Størseth, 2005). Simply put, the perception of low control and high demands makes for an unhealthy combination for both employee and the organization.

To recapitulate the argumentation, the above suggests that IO (as any form of organizational change) may, from a psychological perspective be linked to perceptions and experiences that are not easily kept and canned within the change project. This completes the first half of the announced warning.

The second half of the warning rests upon the option that RE becomes neatly packaged and marketed as a specific kind of 'supermodel' to handle every IO contingency. Again, the reader is invited to take the psychological side of it, and consider what RE may involve for the individuals within the resilient system.

Human cog wheels and resilient virtues

RE is an interesting theory or set of theories. Vivid imagery and metaphors creates powerful associations in the reader. Another very interesting aspect of RE is the mere scope of the theory. Although RE is primarily focused on system, it also encompasses group level and individual level elements. In fact, it is here to be argued that the RE emphasis on, and across all these levels or centers of gravity is a key strength. The way that RE suggests paths between organizational layers and levels and how various organizational focus points are intertwined is a strong point for the approach (Størseth, Albrechtsen & Rø, 2009). There is however, at the same time a "darker side" to this interlock of levels and layers. This darker side is tied to the (tacit) – demands that the drive towards a resilient system may place on the individual.

RE is seemingly unaffected by the 'level problem' in sociology. In fact RE leap rather casually and carefree between system, group, and individual level. The sociological level problem is basically the methodological concern regarding level distinctions and demarcations. In other words, differentiations between micro, meso, and macro-level – and how can they be studied. Anthony Giddens, a sociologist himself, call attention to these kinds of discussions and describe it as the sociological problem of dichotomies.

Giddens' point is that sociology must surpass this dichotomized way of thinking (Moe, 1994). In a way RE follows Giddens advice by the letter. RE as a theoretical approach, moves unconstrained between system-, group-, and individual level. In fact, this is a highly interesting "take" on organizational levels, in that it binds and shows connections between different layers, levels, and focal points in a system.

These "inter-layer connections" makes RE highly interesting in relation to IO as it may serve as a frame for important IO issues such as new technology, transferring functions from offshore to onshore, collaboration across organizational and cultural borders, and increased automation.

At the same time though, the resilient moves between different organizational levels and layers opens for the possible side effect of silently adding demands on the individual. In other words, there is a possibility that when resilient properties are to be engineered into "the system", they end up as a human performance demand feature. This possibility can be illustrated by looking at the adaptation process in RE. The adaptation process refers to the capacity to adjust and adapt and consists of Anticipation (knowing what to expect), Attention (knowing what to look for), and Response (knowing what to do) (Hollnagel and Woods, 2006). How can adaptation be engineered into the organization? And: what does it imply? It is here to be suggested that adaptation may be seen as a (i) function, (ii) premise, and (iii) characteristic. A key question becomes which one of these is to be focused on and effectuated? Without this being specified – there is an opening for (mis)interpretations or perhaps ignorance, leading to placing adaptation demands exclusively at the individual level. In other words, that adaptation becomes a human performance requirement.

The second part of the warning can be recapitulated and summed up in the following thesis: If a resilient system (organization) silently implies resilient personnel, the "resilient jumps" between organizational layers, segments, and focal points may result in the one-dimensional, misinterpreted remedy of placing responsibility on individuals. This is unfortunate because it omits the RE premise of the interlock between layers and levels; and suggests a way to tune and install resilience by demanding resilient virtues of people. And by this latter point, we come full circle on the psychologically based warning as these requirements, placed on the individuals reveal a worldview of cog wheels - a human cog wheel is.. still a cog wheel.

Coda

Although there are various ideas and versions of IO around, it is safe to say that 'mission IO' represent a change momentum in the industry. In this state of optimism, the psychological warning that has been issued here should be read as a reminder that often is forgotten – that change fueled by dreams of efficiency may have a hitch.

Automated control and supervision in eDrilling operations: an HMI-centred set of risk assessment criteria

Denis Besnard

Industrial Safety Chair, MINES ParisTech, Sophia Antipolis, France
denis.besnard@mines-paristech.fr

Abstract. In this essay, a new control and supervision system for offshore drilling (eDrilling) is studied, focusing on automation and human-machine interaction (HMI) issues. The output is a set of HMI-centred criteria for risk assessment in integrated operations

Introduction

Automation in complex systems does not remove humans from control and supervision tasks. In fact, the opposite is true: it make humans all the more important. This is one of the ironies of automation identified by Bainbridge (1987). Indeed, automated systems sometimes face exceptions and unforeseen states for which humans are needed. What is more, given that the missions assigned to automated systems are more and more business critical, it follows that more than ever before, humans are needed. However, the essential contribution of humans to the control of supervision of complex systems can be jeopardised by the system properties themselves. This is the topic that this paper will investigate.

Control and supervision refers the mental activity by which humans typically pilot processes, from flying and aircraft to controlling production chains. Roughly speaking, it is an activity where cues in the environment are used along with a technological artifact in order to assess the current state of a given process, decide whether something has to be done or not, when, how, etc. Consider the example of aircraft piloting: pilots use way points, altitudes, speeds, etc. to fly the aircraft, and use various flight controls and devices to do so.

Control and supervision became heavily computerised over time. The increased level of automation that Information Technology (IT) introduced was initially taken as a guarantee of reliability. However, some important issues started to surface by the late 1980s, when IT was becoming more and more heavily ported to aviation. Along with this technological shift, new forms of accidents appeared where pilots were losing control of the aircraft under normal operational conditions. These events sparked a reflection upon the role of automation in control and supervision tasks that is still going on today (see for instance Baxter, Besnard & Riley, 2004; Besnard & Baxter, 2006).

Because it was heavily affected and because it is a safety-critical industry, commercial aviation pioneered much of the research in automated HMI (human-machine interaction). Today, it is paramount, from car driving to control rooms in nuclear power plants. This leads one to wonder whether automation-related issues could affect other industries than aviation. This paper will address this precise question in the context of offshore drilling operations. Namely, a drilling control and supervision software system named eDrilling will be analysed against a series of automation-related dimensions. This exercise will be performed with the aim to determine a generic set of HMI-centred risk assessment criteria for eDrilling.

Some automation-related issues

Remoteness of control

The first issue is that modern interfaces tend to position operators away from a direct contact with the process. The physical contact with the process to control is lost: the activity becomes that of interacting with the control and supervision system. Also, the activity is more symbolic: the process is indirectly controlled from a set of parameters. Finally, the capacity of the operator to mentally represent the state of the process becomes crucial.

System authority

Interacting with an automated system means collaborating with a machine towards a task. As Sheridan (1992) pointed out, the two agents (human and technical) do not always have the same level of control over a given action. Whether the allocation of control is adjustable or not, and how this can be done, are design decisions. Adjustability is a key notion when humans need to free resources to resort to a strategic level of planning (thereby leaving the machine perform control actions). Conversely, adjustability is also important when an exception arises and humans claim full control over operations (thereby denying all privileges from the machine).

Change of control cues

Automation does not solely mean that a machine is performing a part of given a task. It also changes human work, and what the latter uses as control cues. In the case of aviation, for instance, the mechanical commands acting on control surfaces (e.g. flaps) have been replaced with electric motors and hydraulic actuators. This shift basically deprived pilots from vibrations coming from the wings and elevator; so much so that simulated vibrations modules had to be built into modern control columns. This problem of change of control cues is not unique to aviation. It has also been an issue for high-furnace operators when steelworks factories implemented process control rooms: operators lost direct access to temperature, colour and texture information about the molten metal. Later, the move from electro-mechanical to digital displays deprived nuclear plants control room operators from the clicks that dials used to make when a parameter was changing.

Opacity

Automation (especially due to computerisation) makes the controlled process difficult to track by the operator. If one takes the example of Babbage's Difference Engine, one can actually see the various cogwheels, levers and drums work the result of a computation together. However, once digitised into e.g. a pocket calculator or a computer, the process is not visible any more. Without prior knowledge of internal design, only the input interface (the keyboard) and display (the screen) interface are left to the user to infer how calculations are actually done. The same is true for any machine or system relying on digital technology: it has an impact on how clearly humans can represent the internal functioning of control devices.

Complexity

With automation, complexity increases: the number of functions is usually higher, and the amount of data exchanges between functions increases also. For instance, aboard modern cars, a huge amount of engine and environment-related information reaches the calculator (engine load, composition of exhaust gases, ambient temperature, etc.) which, in turn, adjusts the functioning of the engine (e.g. the fuel/air mixture). The same applies to the piloting of the ABS braking system: the latter uses several input sources of data and acts on several components of the car. These features are a source of efficiency under nominal conditions. However, in case of failure, exception or emergency, manual override is not always possible.

Mode confusion

A mode is a pre-programmed, selectable set of instructions that configures the behaviour of a device. In aviation for instance, when in Go-around mode, the flight systems will automatically reconfigure the aircraft so that it gains altitude and thrust as quickly as possible. However, as Sarter and Woods (1995) pointed out, it can be difficult to know how modes operate. For instance, modes can change indirectly, i.e. due to a change of a flight parameter (as opposed to a manual selection by humans). Such a behaviour introduces a mode awareness issue son the part of the crew. The crash of an A320 on Mont Sainte Odile in 1992 (METT, 1992) or the accident of an A300 in Nagoya (Ministry of Transports, 1996) are events where a mode was erroneously engaged. As a consequence, pilots could not understand why the aircraft was not behaving as expected. It is also interesting to note that both flights were total losses, whereas the aircraft behaved exactly as designed.

So far, we have addressed some automation-related issues for process control and supervision. Now, one might wonder: are these issues relevant to Integrated Operations (IO)? How can they inform risk assessment? These are the questions we are going to turn to in the next sections, using the case of eDrilling.

Automation in IO: the case of eDrilling

eDrilling is a (potentially distributed) software control and supervision system allowing remote parties to share information and contribute to drilling operations. It combines a computerised control interface with sensor information display, display sharing with remote collaborators, as well as a 3D model enabling visualisation and simulation.

The scope of this paper is not to describe the eDrilling system but identify HMI-related issues for risk assessment. For description purposes, interested readers might want to refer to Rommetveit *et al. (*2007), Rommetveit *et al. (*2008a) and Rommetveit *et al. (*2008b). An interesting aspect of these articles is that little is said about the interaction with the system; which is what the following sections will address. In other words, where should one look at to discover HMI-related risk assessment criteria for eDrilling?

Automation-related risk assessment criteria

We are now going to revisit the automation issues laid out in the first section, see how they apply to the eDrilling system. Because risk assessment needs criteria, the pitfalls are now going to be decomposed into basic assessment-focused questions, for which some material was reused from Abbott *et al.* (1996).

- *Remoteness of control*. Can the operator hold an accurate mental representation of what is happening at the actual drilling site? What are the cases where the operator's mental model can be discrepant to what is happening at the actual drilling site? What information is available to operators to revise their mental representation?
- *System authority*. Can the degree of authority of eDrilling be adjusted by the operator? Is full manual control available at any one time? Can there be operations that eDrilling performs that the operator is not informed of, or cannot prevent? Can the eDrilling system override a human action without feedback, or take an action that is not explicitly directed by the operator?
- *Change of control cues*. What is the set of analog control cues that operators use for the control of the physical drilling (noises, vibrations, variations in rotation speed, dial readings, mud properties, etc.)? What is the set of new control cues that operators have to use to control the eDrilling task? Is there any discrepancy between the above two sets of cues? If so, in which cases can this discrepancy have an impact on the operator's performance? Can analog cues be simulated at the eDrilling computer station?
- *Opacity*. Is eDrilling designed so that the operator can understand its internal functioning? Is the system transparent, as if built in a glass box?
- *Complexity*. Can operators understand the internal functioning of eDrilling? Is this understanding sufficient for operators to be able to predict the behaviour of the system? Is the system so hard to interact with that operators can become complacent? Can complexity mask situations that can develop into problems?
- *Mode confusion*. Are there modes in the software system? If so, is their functioning understood by the operator? Can modes change automatically (i.e. indirectly)? Can modes be mixed up? Can modes be disengaged at any one time?

Limitations and recommendations

Risk assessment involving human activity cannot be achieved without carrying field observations, data gathering and possibly experimental performance measurement. Also, it is important to note that the automation-related criteria listed in this paper need to be translated by a human factors specialist into concrete performance indicators before they can be used for risk assessment.

Also, the criteria discussed here must be completed by more traditional HMI dimensions such as ergonomic properties of the software and hardware interfaces, configurability, adaptability, error tolerance, workload, situation awareness, skills transfer from the analog task, ecological representativeness of the computerised task, operator training on their own failure modes, etc. Many more HMI topics can be found, for which references exist. These should be included in risk assessment exercise but fall out of scope of this essay.

Last, an important feature was not commented here. Namely, eDrilling anticipates events, can simulate a sequence, send early warnings of particular drilling condition ahead, etc. which can then be used to take an informed drilling decision. These are uses of IT technology that are supportive of error recovery and anticipative control and supervision.

Conclusion

eDrilling is a system that harnesses IO and its principles of distributivity and remote control. However, eDrilling embeds a complex interface which might contain some potential for failure. In response, this paper attempted to convert some documented HMI evaluation criteria to the field of IO risk assessment, using eDrilling as the case under study.

ICT: Paradoxes of a collaboration technology with a "complexity hyperdrive"

Tor Olav Grøtan

SINTEF Technology and Society, dep. of Safety Research, Trondheim, Norway
tor.o.grotan@sintef.no

Abstract. The essay discusses how ICT can support the management of complexity in general, and resilience as a specific concept addressing complexity. K. Weick's notion of sensemaking and model of organizations as interpretive systems are elaborated in terms of organizational dialectics and actor-network theory to point out that the re-presentational powers of ICT constitutes a "complexity hyperdrive" that cannot be dismissed or turned off, and to reinforce the urgency for facilitation of continuous sensemaking.

Introduction

The aim of this essay is to investigate whether Information and Communication Technology (ICT) as collaboration technology reduce or reinforce complexity in Integrated Operations. In a previous (RIO) essay by this author, complexity has been described as a *wildness-in-wait* embedded in apparent order (Chesterton, 1909), and as *un-order* comprising both directed (resultant) and un-directed (emergent) order (Kurtz and Snowden 2003). These descriptions provide a difference that makes a difference, that is, a richer basis for assessing systemic hazards, and for seeing the implications with respect to various options for risk management.

This essay employs two distinct foci for the discussion of the ICT contribution:

1. The ICT relation to complexity in general. This may be seen as an elaboration or extension of the previous essay, that is, on assessing the effects of using ICT with respect to complexity and systemic hazards.
2. The impact of ICT on resilience (understood as a risk management strategy for dealing with complexity). Resilience is among other things about coping, adaptation, anticipation, attention and monitoring across various levels of the organization. ICT is in that sense a powerful tool, individually as well as socially (group-wise). This essay dismisses what is seen as the naïve conception that the ICT unequivocally contributes to the mastering of complexity through facilitation of resilience. Rather, the question of interest here is whether there is also a "dark side" of ICT that complicates or "complexicates" the fulfillment of resilience.

Hence, the above foci are meant as a preventive barrier toward the possibility that a general euphoria on IO as an expression of the blessings of ICT - as a result of the present focus on complexity, systemic hazard and resilience - immediately transforms into an equally poorly comprehended stance of "IO is made for resilience". I will try to elaborate the view that such a strategy is not likely to succeed, by employing a perspective of organizations as interpretations systems, as conceptualized by Karl Weick (2001), and subsequently look at some implications of this for ICT as collaboration technology.

This essay is hopefully not understood to be in support of a discourse which Renn (2008) denotes the communication of fear by "doom-mongers who outdo each other in conjuring up doomsday scenarios". However, it is meant as a necessary counterweight to what Renn (ibid) denotes the communication of opportunity by IO enthusiasts "requiring nothing but optimism

in order to roll up their sleeves and get on with it, while blaming scapegoats who have "blocked social progress" with their pessimism and technophobia".

Complex organizations as interpretation systems

Weick (2001) asserts that while organizations are the most complex systems imaginable, being vast, fragmented and multidimensional, most empirical research about organizations assumes that they behave as static frameworks and mechanical systems.

Weick argue that the game of "20 questions" is a good metaphor to illustrate the point that organizations typically are in the situation of trying to find an acceptable answer before their time and resources run out, or before the situation reconditions so that another answer is needed. Daft and Weick (2001) thus argues that it is fruitful to see (complex) organizations as *interpretation systems*. Weick (2001) presents a model of the interpretive organization that is build on the following premises:

1. Organizations are open social systems that process information from the environment.
2. Organizational interpretation is something else than individual interpretation. The thread of coherence among managers is what characterizes organizational interpretations, enabling organizations to preserve knowledge, behaviors, mental maps, norms and values over time.
3. Although organizations can be conceptualized as a series of nested systems in which each subsystem may deal with a different external sector, upper managers bring together and interpret information for the system as a whole.
4. Organizations differ systematically in the mode or process by which they interpret the environment, and develop specific ways to know their environment.

Weick thus assumes a layered, hierarchical structure of the organization that does not necessarily correspond with the premises of resilience as a concept (e.g. the seamless connection of activities dispersed at different organizational levels). Neither does it correspond directly with the most "empowered" manifestations (Grøtan et al., 2009) of IO. However, IO manifestations vary a lot and can also be very Tayloristic in their approach, strictly separating "planners" from "doers". There is nothing about IO that precludes the position that Weick's interpretive view of the organization is foundational to how individuals or clusters of actors behave in the IO context. The potential lack of alignment between resilience[11] as a concept and Weick's above assumptions may just be interpreted to serve the purpose of illustrating some remaining key challenges of actually transforming "resilience" in terms of aspects or properties, into an organizational reality.

Interpretation in organizations occur according to Weick as a "second stage" in a broader process which is founded on scanning and data collection. Interpretation gives meaning to data. The organization experiences interpretation when a new construct is introduced into the collective cognitive map of the organization. The third stage, learning, encompass a new response or action based on the interpretation. Weick (2001:248) develops an *Interprêtation System Model* (ISM) based on the idea that organizations may vary with respect to (1) their beliefs about the environment, and (2) in their intrusiveness into the environment. Based on these dimensions, four categories of interpretation behavior may be described (Figure 1, left side).

[11] Resilience as a notion associated with the literature on Resilience Engineering (e.g. Hollnagel et al., 2006). Weick defines resilience somewhat differently (e.g Weick, 2001, ch 4, and Weick and Sutcliffe, 2007).

With reference to my previous essay ("How to address the complexity issue in IO Risk Governance"), it is striking to see how well the ISM corresponds with the Cynefin framework (Kurtz and Snowden 2003), indicating clearly that the organization as an interpretation system (also) deals actively with complexity. The Enacting (corresponding to the Complex domain in Cynefin) mode implies that organizations construct their own environments and gather information by trying new behaviors and seeing what happens ("probing" in Cynefin terms). The Discovering (corresponding to the Knowable domain in Cynefin) mode puts emphasis on detecting the correct answer already residing in an analyzable environment rather than on shaping the answer. The Conditioned Viewing (Corresponding to the Known domain in Cynefin) mode perceives the environment as objective and benevolent, hence the organization does not have to take unusual steps to learn about it. Finally, Undirected Viewing (corresponding to the Chaos domain in Cynefin) is a mode that cannot rely on hard, objective data. Rather, managers act on limited, soft information to create their perceived environment, corresponding with the Cynefin focus on stability focused intervention, enactment tools and an *act-sense-respond* style of management intervention.

		Assumptions about environment	
		Unanalyzable	*Analyzable*
Organizational intrusiveness	*Active*	ENACTING: Experimentation. Testing. Coercion. Invent environment. Learn by doing.	DISCOVERING: Formal search. Questioning, surveys, data gathering. Active detection.
	Passive	UNDIRECTED VIEWING: Constrained interpretations. Nonroutine, informal data. Hunch, rumour, chance opportunities.	CONDITIONED VIEWING: Interprets within traditional boundaries. Passive detection. Routine, formal data.

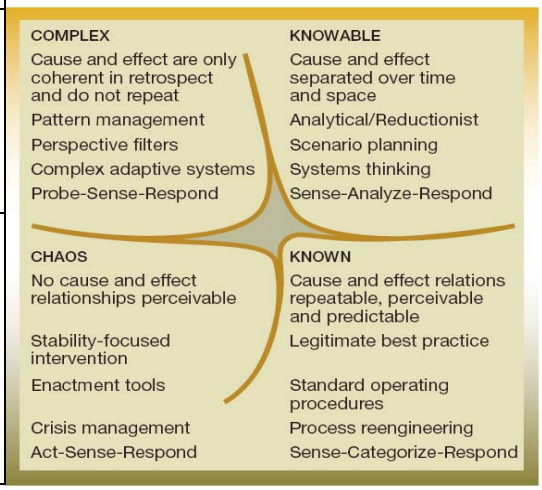


Figure 1. Interpretation System Model (Weick, 2001:248) compared to the Cynefin sensemaking framework (Kurtz and Snowden, 2003)

A *resilient* system could hence be paraphrased, according to ISM and Cynefin, as a system that is able to deal with complexity by dynamically and purposively shifting its assumptions about the environment as well as its own intrusiveness, in at least two distinct ways:

1. prepares to deal with anticipated, exceptional situations by enacting some perceived scenarios (inventing the environment), learning by the reaction and returning (safely) to a formal elaboration of the implications of what have been learned (visiting Complex from a Knowable base in Cynefin terms, Enactment with a subsequent Discovery in ISM terms)
2. is able to be in a permanent mode of enactment and invented environment due to an endured conception of the (real) environment as unanalyzable (intractable), and is able to deal with "chaotic" episodes by means of coercing its (re-)invented environment on to the "real", further stabilizing it so that it can be discovered by formal methods, or stabilizing it by imposing a well known pattern onto it (alternating between

Complexity and Chaos, with the possible option of controlled return to Known or Knowable in Cynefin terms, stretching the persistence of Enactment in ISM terms).

Differences in strategies of *equivocality*[12] *reduction* hints at what can be expected from ICT support. As asserted by Weick, *equivocality* will be most massive in Undirected Viewing (Chaos), in which it is reduced until a common grammar and a course of action can be agreed upon. In Enacting it's reuction will be more on the basis of taking action to see what works instead of interpreting events in the environment. Equivocality is substantially less in Conditioned Viewing and Discovering, as specialists will routinize data and perform systematic analysis and special studies. The data itself provide more uniform stimulus to managers, and less discussion is needed to reach a uniform interpretation. The less the equivocality, the more efficient the ICT support. Its *effectiveness* is however another issue.

Weick's (2001) idea behind the rather normative ISM is to establish a starting point from which to interpret the richness and complexity of organizational activity. Weick (2001:256) asserts that *"... people in organizations are talented at normalizing deviant events, at reconciling outliers to a central tendency, at producing plausible displays, at making do with scraps of information, at translating equivocality into feasible alternatives, and at treating as sufficient whatever information is at hand*". Weick makes the clear point that the outcomes and interpretation processes of real life are substantially less tidy than many have come to appreciate with simplistic models and assumptions about organizations. These human tendencies however also implies that organizations can build up workable interpretations from scraps that consolidate and inform other bits and pieces of data. One important aspect of the ISM model is thus to make sense of the "untidy" sensemaking that may be going on.

I would add that the *un-ordered* complexity argument derived from Cynefin in the previous essay, reinforces the need for such a (sensemaking) framework for organizational interpretation and sensemaking. E.g., the very act of assuming the environment as analyzable and interpreting it on that ground, *may* very well produce the unanalyzable (complex, chaotic) patterns in the environment[13]. In Cynefin terms, directed and un-directed order may in fact be inseparable in complex systems. Hence, neither Cynefin nor ISM have the normative capacity of defining which is which. The engine of the sensemaking process will stop if organizational life become ontologically "fixed" according to either scheme.

ICT contribution (to complexity and resilience)

Sense and nonsense in electronic contexts

The interpretive IO organization will be founded on the use of ICT. What happens at the screens when people act interpretively and seek to deal with complexity and equivocality? Weick (2001) address this topic by the premise that representations of events normally hang together sensibly within the set of assumptions that give them life and constitute a "cosmos", rather than a "chaos". Weick denotes the sudden losses of meaning that can occur when an event is represented electronically in an incomplete, cryptic form, as "*cosmology episodes*". Still according to Weick, this loss of meaning is due to loss of the connection to ways of knowing that preserve properties of events not captured by machines, e.g. feelings, contexts and sensory information: When people are forced to make judgments based on cryptic data,

[12] The extent to which data are unclear and suggest multiple interpretations about the environment
[13] Or, according to Chesterton (1909), the "wildness-in-wait".

and comparisons with other sources are nor available, people try to resolve their puzzlement by asking for more data, which in turn means that the majority of human processing capacity is used to keep track of unconnected details. On the other hand, when a person is able to connect the details and see what they might mean, processing capacity is restored. Meaning that can impose some sense on details typically comes from sources outside the electronic cosmos, sources like metaphors, culture, archetypes, myths, narratives and history. The electronic world makes sense only when people are able to reach outside of it for qualitatively different images.

Weick asserts that the fundamentals of sensemaking is about:
- *Effectuating*; e.g., people learn about events when they prod[14] them to see what happens
- *Triangulating*; e.g., convergence involves qualitatively different measures, not simply increasingly detailed refinements
- *Affiliating*; e.g., people learn about events when they compare with other's perceptions, and then negotiates something mutually acceptable
- *Deliberating*; e.g., people learn through slow and careful reasoning, mindless activity takes less time than mindful activity
- *Consolidating*; e.g., to consolidate bits and pieces into a compact, sensible pattern frequently requires a look beyond them to understand what they might mean

"*It is the very self-contained character of the electronic cosmos that tempts people, when data makes less and less sense, to retain assumptions rather than move to different orders of reasoning*". (Weick, 2001:448). Weick claim that people using ICT are susceptible to cosmology episodes because they act less, compare less, socialize less, pause less and consolidate less. This leads him to the following possible deficiencies when sensemaking is performed in front of the computer terminal:
- *Action deficiencies*; e.g., events are seldom confronted, prodded or examined directly
- *Comparison deficiencies*; e.g., as action as a source of comparative data is missing, single, uncontradicted data give a feeling of omniscience, but leads to non-adaptive action
- *Affiliation deficiencies*; e.g., if the ICT user is in a solitary setting, it degrades to "quantified narcissism disguised as productive activity"
- *Deliberation deficiencies*; e.g., whims and mixtures of feeling, thought and images are dumped into the analyzing process continuously, of which the sheer volume make it harder to separate figure from ground
- *Consolidation deficiencies*; e.g., the very self-contained character of information systems undercut their value.

Weick claims that surprisingly many of these problems can be solved if people simply push back from their terminals and walk around. That is not necessarily an option in the IO context. A more promising approach is the claim that information overload is not really the problem. Overload occurs when one get too much of the same kind of information, and declines with the increase of different *kinds* of information. As walking around is a limited option, access to not only different media, but with – as far as possible - different *kinds* of information would clearly be beneficial for IO. Another part of Weick's analysis is maybe even more applicable; people should *listen more and talk less*; the detail, specificity and concreteness that can be achieved by ICT is worthless until patterns are imposed on them. Moreover, people do not only need to edit. Each person should aim at absorbing uncertainty and transmit less than they

[14]Hence, enacting, or probing of complexity in Cynefin terms, is the "nominal" way of making sense

receive. There is always the danger of mishaps in doing so, but the even greater danger is to leave in too much, and paralyze oneself and others. And, as in the IO context in which people get more and more dispersed, faith and trust will be increasingly important.

Electronic organizations need to develop new respects for generalists, philosophers and artists who work with frameworks that provide context and meaning for the programs already in place. This final point of Weick (2001) make a bridge to points made by Carlsen et al. (2004), in which the fundamental distinctions between data, information, knowledge and understanding are highlighted. Such distinctions are fundamental for deriving powerful concepts like tacit knowledge, boundary objects, perspective making and perspective taking. These concepts are fundamental premises for a thorough understanding of human *communication and collaboration* within a *knowing as activity* framework. According to such a framework, knowledge representations are boundary objects, whether within or between groups, so to say formed "in anticipation of an answering world". The process of *re-presenting* experiences (stories, metaphors, procedures, images, practice terms) and their subsequent *enactment* (retelling, emphasizing, forgetting, re-interpreting, imagining) is something that cannot be transformed into the simplistic sender-receiver model of data communications, as it even moves beyond (still) simplistic dichotomies like tacit vs explicit knowledge.

All in all, this points in two main directions

1. In order for sound sensemaking to take place, the "IO organization" should seek to complement what Zuboff (1984) denotes the "electronic text" of data representations with other sources of information. With the reduced option of physically "walking around" in the IO context, video conferencing solutions should be comprehended not only as a specific mode of interaction, but also as an integrated part of the sensemaking process.
2. The interpretive capacity of the IO organization should be diverse, that is, accommodating both the specialist and the generalist, and even the "artist".

Dialectics of organizational resilience

The above practice-oriented implications for ICT-supported sensemaking can also be applied directly to the topic of resilience. Nathanael and Marmaras (2008) address organizational aspects of resilience in terms of the complex interplay between prescribed organizational provisions and emerging work practices. They argue that there will always be a mute *dialectic* between what is experienced – thus constituting practice - and what is prescribed. The mute dialectic between what is experienced (practice) and what is prescribed are so to say *springs* of resilient action. It is by monitoring and facilitating this mute dialectic that an organization can sustain and take advantage of these springs. If the dialectic is inhibited by merely trying to enforce practice into compliance with what is prescribed, an organization risk losing its ability to respond and adapt in the face of the unexpected, thus losing its access to the crucial repertoires deriving from these springs, and, by implication, losing the ability to be resilient. This makes an important distinction in relation to Weick's idea of the interpretive organization in which managers bring together and interpret information for the system as a whole: *the managerial (prescriptive) and the practice-oriented ways of sensemaking should not be aligned, but held at proper "dialectical distance"*!

It will however be tempting to make these repertoires for resilient praxis available (through ICT) as shared repositories for resilient action. However, even this is not as straightforward as

it may seem. Work practices differ from procedures or tasks with specifically defined finalities. According to Nathanael and Marmaras, the *core invariant* in repetitions of practice is not the recurrence of the same events as seen by an external observer; rather, it is a kind of shared or convergent *generative method* of who, when and how to act. Work communities progressively build some kind of rationalized representation of their proper practices. These are however not and cannot be a complete deciphering of their practices. They are "rationalized accounts" of "what they do" aimed to support (local, situated) sensemaking and consensus. They are thus "rational" in the weak sense meaning that they are acknowledged by members of *that specific* community as legitimate representations, but they do not carry any claim for universal validity (as they may be perceived by a "prescriptive eye").

Actor-networks: Taming the powers of re-presentations

Lilley et al. (2004) portrays ICT as a *re-presentation* technology. As re-presentations are both technical and social, representations also reveals their interconnection. By information the world becomes re-presented as a stock of parts that can be recombined at will. Objects (re-presentations) loose their essence and take on a mode that can be metaphorically exemplified by Lego or Meccano. Lilley et al. link these characteristic with power and the Actor Network Theory (ANT) approach, and argue that the appearance of detached representations is productive of a *new form of power* that enable a view of the world as a table top ruled by the human hand and eye. The attempt to trap all uncertainty tends to-wards an overarching and closed system. As a result everything is dragged closer together and made smaller, is displaced and abbreviated in order to facilitate remote control. The deployment of *ICT thus holds out the dream of grasping the uncertainty created by its own dispersal*. Hence, ICT has a Janus face: it is a booster of efficiency, but may also be effective of the risky propagation of an "artificiality" that may easily detach from reality, enable a new form of power that is un-evenly distributed, and that attempts to rescind the need for sensemaking of the "powerless".

ANT frames the implication of this by the notion of heterogeneous "worlds" that are too complex to be closed/ordered according to one single mode or logic. Hence, also IO will comprise partial orders which are interacting in different ways, to a much larger extent than the perspective of Nathanael and Marmaras (2008) addresses, in terms of interconnected and overlapping subworlds which are ordered according to different logics. By trying to make order in one subworld by imposing a specific logic, the same logic is making disorder in another.

Under such circumstances, the keywords are now multiplicities, inconsistencies, ambivalence and ambiguities, rather than alignment, stabilization and closure. Mastering this new world is not about achieving stabilization and closure, but rather more about ad hoc practices – "ontological choregraphy" of an ontological patchwork (Mol and Law, 2002). The re-presentational capabilities of ICT are indeed very productive, but may also develop beyond control if they are not conduced by an overall respect for the dialectics and sensemaking processes on terms of human understanding.

Mol and Law (2002) warns against the denouncement of simplification as being sheer violations towards an ontologically complex world (which then strikes back). This common rhetorical trope from quite a few complexity theory purists is flawed. Mol and Law argues, similar to Kurtz and Snowden (2003) for the recognition of *multiplicity* in knowledge practices. "*Various orderings of similar objects, topics, fields do not always reinforce the same simplicities or impose the same silences. Instead the may work – and relate – in*

different ways" (Mol and Law, 2002: 7). They argue that we should talk about *complexities* instead of complexity. For me, this is another reminder that un-order cannot be clinically separated from order. The wildness in *wait* will be *embedded* in apparent order, and a continuous sensemaking activity on dialectical premises is a prerequisite for resilience.

Conclusion

I have made an attempt to elaborate the "interpretive organization" perspective of Weick (2001) with other theories of practice-oriented communication, dialectical relations, the power-laden representative powers of ICT, and Actor-Network theory. All these are found relevant for ICT-supported management of complexity in general, as well as resilience as a distinct concept.

In line with e.g. Hanseth and Ciborra (2007) and Kallinikos (2006), I advocate the view that ICT may actually be an additional source of wildness, constituting a sort of Perrowian (escalation) effect not only in the "real" world, but also in the realm of information, interpretation, knowledge and sensemaking, individually and/or groupwise. According to Weick (1995), the "normal accident" may very well originate in this sphere, and proceed with a most dramatic impact onto the "real" world.

The "complexity hyperdrive" of ICT cannot be dismissed or turned off.

.

Improved safety performance by integrated operations and its implications for risk management

Eirik Albrechtsen

SINTEF Technology and Society, dep. of Safety Research, Trondheim, Norway
eirik.albrechtsen@sintef.no

Abstract. The essay shows how integrated operations (IO) can improve the safety performance and risk management approaches by IO concepts such as visualization, real time data and access to expert knowledge. Risk management needs to consider normal operations and successes in addition to unwanted incidents.

Introduction

Several of the actors in the Norwegian petroleum sector have stated that IO will lead to an improved HSE level in addition to business-related value creation (e.g. OLF, 2003, 2007 and Statoil homepage). This essay 1) identifies IO concepts that can contribute to an improved safety level; and 2) discusses implications for risk assessment/management.

Positive effects on safety by the IO development

An overview produced for the Norwegian Safety Petroleum Authorities (Grøtan and Albrechtsen, 2008) shows both negative and positive implications for the organizational accident risk in an IO context. By studying IO concepts in different perspectives on organizational accidents and resilient organizations, they look at both positive and negative influences on organizational accident risk. To fully understand risk in an IO context, we need to study both the positive and the negative influences on safety; however this essay is limited to studying the positive effects. The overview identifies the following positive effects on organizational accident risks:

A barrier perspective:

- Possibility to establish barriers according to the current risk picture by use of real-time data, information about future situations and access to expertise
- Faster detection (by monitoring) and improved ability for anticipation (by real-time data, sharing/visualization of data, expert support) of deviations
- Access to real-time data on safety conditions, and new ways for visualization and sharing of data makes it possible to follow-up of barriers more closely

An information processing perspective

- One of the key premises for new types of operation in the North Sea is real-time data and distribution of information. IO thus makes it possible to have a proactive safety mindset by access to real-time data available for many actors; simple presentations and processing of data; and access to expert knowledge.
- In collaboration rooms, many different actors from different organizations will see and have access to the same information thus having a potential for strengthening an interdisciplinary safety approach.
- Richer channels for communication by collaboration technology, implying that more actors see the same and more of the same

A decision-making perspective

- More efficient and better decisions are one of the key expectations of the IO development. It can be argued that integrated operation lead to better and safer decisions by real-time data and detailed understanding of situations; interdisciplinary teams involved in the decision making; access to expertise; and parallel activities (Ringstad and Andersen, 2007)
- Such decision-making contexts also makes it possible to involve safety personnel/experts and safety information in more decision-making processes.
- Decisions made on real-time data (and not only historical data), expert knowledge and interdisciplinary teams, provides a better basis for a proactive approach.

The theory of Normal Accidents

- If operations are well-planned, which is one of the main objectives of IO, the degree of complex interactions and tight coupling can be reduced by linear operation and few unforeseen interactions. (by increased insight)

The theory of High Reliability Organisations

- The ability for organizational redundancy is improved in IO as more actors can observe, participate and contribute in collaboration arenas between disciplines and organisations. Multi-disciplinary teams and different situational awareness can be an arena for creation and maintenance of both cultural and structural redundancy.
- The ability to change operation mode in crises can be improved in IO by faster detection of failures implying that the response time to the crisis is reduced. Additionally, more support from different actors can improve the handling of the crisis, e.g. bringing in experts on oil spills. Furthermore, integrated contractors can be given extended authority in crisis.

The theory of Resilience Engineering

- In general IO has a proactive focus, thus strengthening the ability to discover and be prepared for unexpected situations, by real-time data; integrated planning; and more actors creating a larger total situational awareness. More people can be trained for handling of complex situations, by “on the job training” as well as by improved real simulation training.

Implications for risk assessment and management

The implications of the IO improvement on the safety level can be divided in two: 1) IO as an enabler for new practical approaches to risk management/assessment; and 2) a change of mindset to capture “positive risk” and manage accordingly

IO influences on risk assessment/management

The IO development also influences important principles in safety management. OLF (2007) points out some aspects of IO that improves safety management systems:

- Real-time data on safety related conditions and new ways for visualization and simplified presentation of data, including possibilities to share safety-related information in larger parts of organizations.
- Improved continuity and follow-up of safety between shifts because of tighter integration between onshore and offshore
- Faster detection of failures and normalization due to real-time monitoring of systems and access to expert knowledge

- Improved possibilities to integrate suppliers, thus giving them better opportunities to influence safety management approaches
- New possibilities for work permits. Visualization can make work permits more comprehensible. The quality of the work permit system can also improve by using area categorizing and information on where people are located

As an addition to these improvements, three other potential IO improvements of risk management will be presented here: living risk assessments; risk visualization and safety support centers.

Living Probabilistic Safety Assessment

As described in, one of the main potential improvements by IO on safety, is real-time data, access to expert knowledge and sharing and presentation of information. How can these IO concepts be utilized for risk assessment?

Within the nuclear power sector, the concept Living Probabilistic Safety Assessment (LPSA) has been used for some decades. The International Atomic Energy Agency (IAEA, 1999:1) describes LPSA as: "*…a PSA of the plant, which is updated as necessary to reflect the current design and operational features, and is documented in such a way that each aspect of the model can be directly related to existing plant information, plant documentation or the analysts' assumptions in the absence of such information. The LPSA would be used by designers, utility and regulatory personnel for a variety of purposes according to their needs, such as design verification, assessment of potential changes to the plant design or operation, design of training programmes and assessment of changes to the plant licensing basis.*"

The LPSA is updated by real-time data in combination with manual updates, i.e. when changes occur in any aspect of plant operation or design, or improved understanding of processes or accident phenomenology, revised data or advances in analytical techniques (IAEA, 1999). For the real-time analysis, a monitoring system is used to determine the instantaneous risk based on the actual status of the systems and components.

There are clearly similarities between nuclear facilities and oil and gas installations, the potential for experience transfer between these sectors are thus in place. Without considering model, assumptions and pit-falls of quantitative risk assessment in general, there are clearly potentials within IO that can make risk assessments more "living":

- the increased capture and sharing of real-time data, in particular related to technological performance
- more and better information about underlying productive and administrative processes, which can be used to understand, improve and plan activities, including changes of operations and design
- access to expert-knowledge and use of interdisciplinary teams strengthen the understanding of e.g. technology and hazards, thus (continuously) improving the basic assumptions behind the assessment
- more actors can observe, participate and contribute in collaboration arenas between disciplines and organisation, which can provide a more realistic understanding of the phenomena to be assessed.

Risk visualization

Within the Center for Integrated Operations in the Petroleum Industry in Trondheim, there is an ongoing development of an application for risk visualization, called 'IO Map'. The project is now conceptualizing the model and application, and has started to test the application. IO Map shows how shared information surfaces and visualization technology can contribute to risk management in an IO context. The application is planned to highlight risks associated with different jobs and risk associated with combination of jobs (Skjerve et al., 2009), see figure 1. The application has been developed for use in long and short term maintenance and modification planning (Rindahl et al., 2008), and can be used in different stages in such planning processes, ranging from notifications (someone reports that a job needs to be done) to permission of work (a job that has been planned and evaluated with respect the HSE, criticality and feasibility)

Figure 1. Part of the IO Map visualization, the map shows the basement deck of an installation. Yellow area indicates explosion danger, while yellow area indicates no danger. Triangles show work permits (Rindahl, 2009)

Making results of risk assessment available and understandable for decision-makers and stakeholders have been pointed out as one of the weaknesses of state of the art risk management. Risk visualization can be one of the solutions to this problem. Risk visualization in combination with collaboration technology should improve risk communication in an IO setting, not least regarding sharing information on risk between distributed actors.

Safety support centre / operation centre

Most companies operating at the Norwegian shelf have onshore centres operating closely with offshore activities. It varies whether the centres are a staff function (support centres) or a line management function (operation centres). Common characteristics of theses centres are: interdisciplinary teams; access to real-time data and information; and use of collaborative technologies to communicate with offshore installations and other onshore experts. Some of these centres are covering particular processes, e.g. the Subsurface Support Centre at Statoil and the Operational Drilling Centre at ConocoPhillips. A safety support/operation centre is thus a possible IO solution.

What can the role and function of a safety support/operation centre look like with regard to risk assessments and decision-making?

- *Network of experts.* Function as a central node in a network for safety disciplines and operational units, including external collaboration partners, e.g. suppliers. Also establish and utilize network with external experts (authorities, researchers, consultants)
- *Pre-assessment.* Understand and keep an overview of hazards, vulnerabilities and opportunities
- *Information database.* Keep an overview over historical data (risk assessments, KPIs, audits, accident investigations); procedures; regulations
- *Method database.* Keep an overview of available methods for risk assessment, including strengths and weaknesses, and give advices in selection of methods and guidance in use of these
- *Risk communication.* Communicate and present results in an understandable way for decision-makers, hence contributing to a good risk understanding among different actors
- *Expert support.* Provide expert advice to different decisions on safety (ranging from design to operation)
- *Updating models, input and assumptions.* A responsibility for making risk assessments "living", by updating models and input according to changes and new knowledge
- *Interdisciplinary team.* The centre should consist of different disciplines: technical safety, operational safety, quantitative risk assessors, etc

Risk management mindset

Risk managers and risk researchers are mainly preoccupied with what undesired events that can occur and the probability and damage potentials of these undesired events. This is of course an important part of managing risk, which is rooted in standards and legislation in the risk management field. Major accidents at the Norwegian Shelf happen seldom; as a result we must assume that most of the time things work well. Focusing on what goes right has thus a great potential for maintaining and performing better and safer (one of theses success factors can of course be a well functioning risk management in terms of managing what can go wrong).

To manage "positive risk" we need to identify situations of success, and understand how and why these situations are successful. Resilience Engineering (e.g. Hollnagel et al., 2007) offer a framework for understanding why normal performance create safety. Størseth et al. (submitted) describe critical success factors for incident recovery and prevention. Tinmannsvik (2008) asks 'why doesn't it happen more accidents at the Norwegian contintental shelf?', and explains this by i.a. barriers, robust working praxis and improvisation. Based on these positions, the following questions are examples that can be asked to understand why normal performance succeeds:

- How do individuals and groups adjust their performance in order to cope with disruptions and disturbances?
- How are individuals and groups aware of current threats and opportunities and how do they adapt according to these?
- How do individuals and groups anticipate future development and threats, and how to they adapt to handle these future vulnerabilities and opportunities?
- How do individuals and groups learn the right lessons from the right experience?
- What are the technical and operational barriers that contribute to normal performance?

Learning from and adapting to successes should prove to be an important addition to risk reduction. However, if this new approach is to be part of the operational risk management, the approach should be operationalized in terms of a method or a tool. Resilience Analysis Grid (RAG) seems to be a promising tool. Check-list based tools that identify successes and provide insight into why successes happen can be another possible development. Additionally, monitoring safety perforamce by leading indicators can be another solution.

Summing-up

As risk managers and risk researchers we can not only pay attention to what might go wrong. In addition to what goes wrong, we also must pay attention to successes – what are the barriers and processes that produce successes and why do these succeed? This focus is not least important in an IO setting, where several stakeholders claim that IO will create an improved safety level. There is a need to develop methods that identifies the IO concepts that improves safety and study how and why they maintain and reduce major accident risk. Additionally, IO concepts can enrich risk management by e.g. real-time data; visualization and information sharing; and interdisciplinary knowledge.

References

Adams, J. (1995). *Risk*. Londonn: UCL Press

Albrechtsen, E. Mostue, B.Aa. & Hovden,J (2009) *Risk Governance of Integrated Operations*, SINTEF Technical Note

Atomic Energy Commission (1975). *Reactor safety study: An assessment of accident risks in U.S. Commercial power plants (WASH-1400)*. Washington, D.C.

Aven, T. & Renn, O. (2009). On risk defined as an event where the outcome is uncertain. *J. Risk Research*, 12, 1-11.

Aven, T. (2009). Perspectives on risk in a decision-making context – Review and discussion. *Safety Science*. 47(6), 798-806.

Aven, T. (2009). Perspectives on risk in a decision-making context - Review and discussion. *Safety Science*.47(6), pp 798-806.

Aven, T. (2010). *Misconceptions of Risk.* Chichester: Wiley

Beck, U. (1992). *Risk Society: Towards a New Modernity.* London: Sage Publ.

Bolman, L.G. & Deal, T.E. (1984). *Modern Approaches to Understanding and Managing Organizations.* San Fransisco: Jossey-Bass Publishers.

Chesterton, G.K. (1909). *Orthodoxy*. New York: Lane Press (Reprinted by Grenwood Press, Westport, 1974).

Cohen, M.D., March, J: G: & Olsen, J.P. (1972). A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly*,17(1),

Daft, L. D. and Weick, K. E. (2001). *Toward* a model of organizations as Interpretation Systems. In Weick, K.E. *Making Sense of the Organization*. Oxford : Blackwell

Davoudian, K., Wu, J.-S. & Apostolakis, G. (1994). Incorporating organizational factors into risk assessment through the analysis of work processes. *Reliability Engineering & Systems Safety*, *45*(1-2), 85-105.

Dougherty, E. M. Jr. (1990). Human reliability analysis - Where shouldst thou turn?. *Reliability Engineering and System Safety*, *29*(3), 283-299.

Etzioni, A. (1986). Rationality is anti-entropic. *Journal of Economic Psychology*, 7, pp. 17-36.

Giddens, A. (1990). *The Consequences of Modernity*. Cambridge : Polity Press

Giddens, A. (1991). *Modernity and Self-Identity*. Cambridge: Polity Press.

Grøtan, T. O., Albrechtsen, E. and Skarholt, K. (2009). How shared situational awareness influence organizational accident risk in the offshore oil industry. Proceeding of *ESREL 2009* conference. Taylor and Francis.

Grøtan, T.O. & Albrechtsen, E. (2008). *Risikokartlegging og analyse av Integrerte Operasjoner (IO) med foksus på å synliggjøre kritiske MTO aspekter*. [Risk identification and analysis of integrated operations concerning MTO aspects] SINTEF report STFA7085

Grøtan, T.O., Albrechtsen, E., Rosness, R., Bjerkebæk, E., (2008). The Influence on Organizational Accident Risk by Integrated Operations in the Petroleum Industry. In proceedings of *WorkingonSafety 2008*

Hale, A. & Hovden, J. (1998). Management and culture: The third age of safety. A review of approaches to organizational aspects of safety, health and environment. In A. M. feyer & A. Williamson (Eds.), *Occupational Injury: Risk, Prevention, and Intervention*. CRC Press.

Hanseth, O. & Ciborra, C. (eds.) (2007). *Risk, Complexity and ICT*. Northampton, MA: Edward Elgar

Hepsø, V. (2006). When are we going to address organizational robustness and collaboration as something else than a residual factor? Paper presented at the *2008 SPE Intelligent Energy Conference and Exhibition.* SPE 100712. Amsterdam, Netherlands

Hill, E. J. (1975). Sneak Circuit Analysis of Military Systems. Proceedings of the *Second International System Safety Conference.*

Hollnagel, R. (2009). The four cornerstones of resilience engineering. In Nemeth, C.P., Hollnagel, E. & Dekker, S. *Resillience Engineering Perspectives, Volume 2. Preparation and Restoration.* Farnham: Ashgate

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

Hollnagel, E. & Woods, D.D. (2006). Epilogue: Resilience Engineering Precepts. In Hollnagel, E, Woods, D.D, Leveson, N. (Eds.). *Resilience Engineering – Concepts and Precepts*. Aldershot: Ashgate.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method (CREAM).* Oxford: Elsevier Science Ltd.

Hollnagel, E. (2004). *Barriers and Accident Prevention*. Aldershot: Ashgate.

Hollnagel, E., (2008). *The changing nature of risks*. Ecole des Mines de Paris, Sophia Antipolis, France.

Hovden, J. & Rasmussen, J. (1999). On allocation of professor at Swedish universities in support of SRV's risk management research. SRV, Karlstad.

Hovden, J. (1999) *Sikkerhetsforskning. En utredning for NFR.* In Norwegian. NTNU report.

Hovden, J., Rosness, R., and Wallace, S.W. (2001): Exploring beliefs in modelling decision-making: Optimising and cost-cutting versus risk and vulnerability. Paper presented at *The 5th International Conference on Technology, Policy an Innovation*, Delft 2001

International Atomic Energy Agency , IAEA, (1999), *Living probabilistic safety assessment (LPSA).* IAEA-TECDOC-1106.

Kaplan, S. (1991). Risk assessment and risk management – basic concepts and terminology. In Knief et al. (eds.) *Risk Management: Expanding Horizons in Nuclear Power and Other Industries*. Boston, MA: Hemisphere Publ. Corp., pp. 11–28

Kets de Vries, M.F.R. Balazs, K. (1997) The Downside of Downsizing. *Human Relations, 50* (1), 11-50.

Kirwan, B. (1994). *A guide to practical human reliability assessment*. London: Taylor & Francis.

Kjellén, U. (2000) *Prevention of accidents through feedback control*. London: Taylor and Francis

Klinke & Renn (2001) Precautionary principle and discursive strategies: classifying and managing risks. *In Journal of Risk Research*. 4 (2), 159-173

Krimsky, S & Golding, D. (eds.) (1992) *Social Theories of Risk.* Westport: Praeger Publ.

Kurtz, C.F. and Snowden, D.J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal*, 42(3) pp. 462-483

LaPorte, T.R. & Consolini, P.M. (1991).Working in Practice but not in Theory: Theoretical Challenges of "HighReliability Organisations". *Journal of Public Administration Research and Theory*, 1(1)

Luhman, N., (1993). *Risk: A Sociological Theory.* New York: Aldine de Gruyer,

Lupton, D., (1999).*Risk*. London: Routledge

Maani, K. E. and Cavana, R.Y. (2007) *Systems Thinking, System Dynamics: Managing Change and Complexity*. North Shore, N.Z. : Pearson/Prentice Hall

Moe. S. (1994). *Sosiologi i 100 år. En veileder i sosiologisk teori.* In Norwegian [100 years of Sociology. A guide in sociological theory] Oslo: Universitetsforlaget,

Næsje, P. (2009): *Effects of Integrated Operations on Offshore Installations' Health and Safety Performance*, SINTEF-report A12025,.

Nathanael, D. and Marmaras, N. (2008). Work Practices and Prescription: A Key Issue for Organizational Resilience. In Hollnagel, Dekker and Nemeth (eds): *Remaining Sensitive to the Possibility of Failure*. Ashgate Studies in Resilience Engineering.

OLF (2003) *eDrift på norsk sokkel – det tredje effektiviseringsspranget.* In Norwegian [eOperation at the Norwegian continental shelf – the third efficency leap] Report from OLF The Norwegian Oil Industry Association

OLF (2007) *HMS og Integrerte Operasjoner. Forbedringsmuligheter og nødvendige tiltak.* In Norwegian [HSE and Integrated operations. Possible improvements and necessary measures] Report from OLF The Norwegian Oil Industry Association

Perrow, C., (1984; 1999). *Normal Accidents. Living with High-Risk Technologies.* New York: Basic Books.

Petroleum Safety Authority (2009): *"Risikonivå i petroleumsvirksomheten, RNNP. Hovedrapport, utviklingstrekk 2008, norsk sokkel"* (Risk level in the petroleum industry. Main report, progress 2008, Norwegian shelf), Stavanger.

Poucet, A. (1989). *Human Factors Reliability Benchmark Exercise - Synthesis Report* (EUR 12222 EN). Ispra (VA), Italy: CEC Joint Research Centre.

Rasmussen, J & Svedung, I. (2000). Proactive Risk Management in a Dynamic Society. Räddingsverket. Swedish Rescue Services Agency.

Rasmussen, J. (1997). Risk management in a dynamic society. A modelling problem. *Safety Science*. Vol. 27, No 2/3, 183-213.

Rasmussen, J., (1997). Risk Management in a Dynamic Society: a modeling problem. *Safety Science* 27(2-3), 183-213.

Reason, J. T. (1997). *Managing the risks of organizational accidents*. Aldershot: Ashgate Publishing Limited.

Renn, O. (2008). *Risk Governance. Coping with uncertainty in a complex world.* London: Earthscan.

Rindahl, G (2009), *Future Collaboration Environments*, Presentation at the IO center, October 2009

Rindahl, G Skjerve, AM, Fallmyr, O, Nilsen, S, Sarshar, S, Randem, HO, Braseth, AO (2008) *Studies on the effects on risk identification in team decision processes when applying Risk Visualisation Technologies to support Long and Short Term Maintenance and Modification Planning in Future Integrated Operations in the Petroleum Sector* White paper, IFE

Ringstad, J., Andersen, K. (2007): ”Integrated operations and the need for a balanced development of people, technology and organisation, *Paper at the International Petroleum Conference IPTC*, Dubai. 11668

Rosness, R. (2000): *”Om jeg hamrer eller hamres, like fullt så skal der jamres”. Beslutningsdilemmaer og sikkerhet.* (Between the devil and the deep sea. Decision dilemmas and safety.) SINTEF report STF38 A01413

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

Simon, H. A. (1947). *Administrative Behaviour. A Study of Decision-making Processes in Administrative Organisations.* Republished by Simon & Schuster Ltd (1997)

Skjerve, A.B., Albrechtsen, E., Tveiten, C.K. (2008). *Defined Situations of Hazard and Accident related to Integrated Operations on the Norwegian Continental Shelf*. SINTEF report A9123.

Skjerve, A.B., Kaarstad, M. Rosness, R. (2009): *Building Safety. Literature Surveys of Work Packeges 2 and 3: Decision Making, Goal Conflicts, Cooperation, IO Teamwork Training, Decision Support, and the impact on Resilience of New Technology*. IFE/HR/F-2009/1388. IFE (Institute for Energy Technology), Halden.

Skjerve, A.B., Rindahl, G., Randem, H.O, Sarshar, S (2009), *Facilitating Adequate Prioritization of Safety Goals in Distributed Teams at the Norwegian Continental Shelf* In proceedings of *the IEA (International Ergonomics Association) 17th World Congress on Ergonomics*, Beijing, August 2009.

Størseth, F. (2005). The perception of job insecurity – organisational antecedents, employee experiences and outcomes for health and safety. *Doctoral theses at NTNU* 2005:123.

Størseth, F. (2006). Changes at work and employee reactions – organisational elements, job insecurity, and short-term stress as predictors for employee health and safety. *Scandinavian Journal of Psychology*, 47 (6), 541-550.

Størseth, F., Albrechtsen, E., & Rø M.H. (2009). Resilient recovery factors: explorative study (*submitted*).

Tinmannsvik,R.K. (2008). R*obust arbeidspraksis. Hvorfor skjer det ikke flere ulykker på sokkelen?* [Robust working praxis. Why doesn’t it happen more accidents at the Norwegian contintental shelf?’]. Trondheim: Tapir Akademisk Forlag

Turner, B.A. & Pidgeon, N.F., (1997). *Man-Made Disasters.* Oxford: Butterworth-Heinemann.

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

Weick, K. E. (1995). *Sensemaking in Organizations.* Thousand Oaks, Calif. : Sage

Weick, K. E. (2001). *Making Sense of the Organization*. Oxford : Blackwell.

Woods, D.D. (2006) Essential characteristics of resilience. In Hollnagel, E, Woods, D.D, Leveson, N. (Eds.). *Resilience Engineering – Concepts and Precepts.* Aldershot: Ashgate.

Trondheim
Address: NO-7465 Trondheim, Norway
Telefon: +47 73 59 30 00
Telefax: +47 73 59 33 50

Oslo
Address: P.O. Box 124, Blindern, NO-0314 Oslo, Norway
Telefon: +47 22 06 73 00
Telefax: +47 22 06 73 50