Identifying the typical biases and their significance in the current safety management approaches

Teemu Reiman[a*], and Carl Rollenhagen[b]
[a]VTT, Espoo, Finland
[b]KTH, Stockholm, Sweden

Abstract: The aim of the article is to describe a set of biases in safety management practices and their possible consequences for safety. We will outline main biases of safety management in four thematic areas: beliefs about individual behavior, beliefs about organizations, safety models and safety management methods. A common theme underlying the biases is a lack of systems view on safety. A systemic safety management takes into account people, technology and organization and their interaction in equal terms. Furthermore, such an approach can shift focus from people to technology to organizational aspects depending on their current safety significance.

Keywords: Safety management, organizational factors, human factors, safety science.

1. INTRODUCTION

Safety management practices are based on underlying models or theories of organization, human behaviour and system safety. These theories are either explicit or implicit or a combination of both. An important function of theories and models of safety management is that they create expectations and suggest potential actions. The aim of the article is to describe a set of biases in safety management approaches and their possible consequences for safety. We have tried to extract what safety management professionals and researchers generally look for – and what they might miss. When people form expectations, they assume that certain sequences of actions/events are likely to happen. Such expectations and their associated assumptions are embedded into organizational practices, routines, norms and management strategies [56, p. 41]. Expectations guide our attention and search for evidence, thus making it easier to confirm the accuracy of our original expectations by neglecting contradictory information. Expectations can also undermine reliable and resilient performance because they encourage confirmation seeking, reliance on existing categories, and oversimplification. Consequently, organizations should continuously work to override typical human tendencies to seek confirmation and avoid disconfirmation [56].

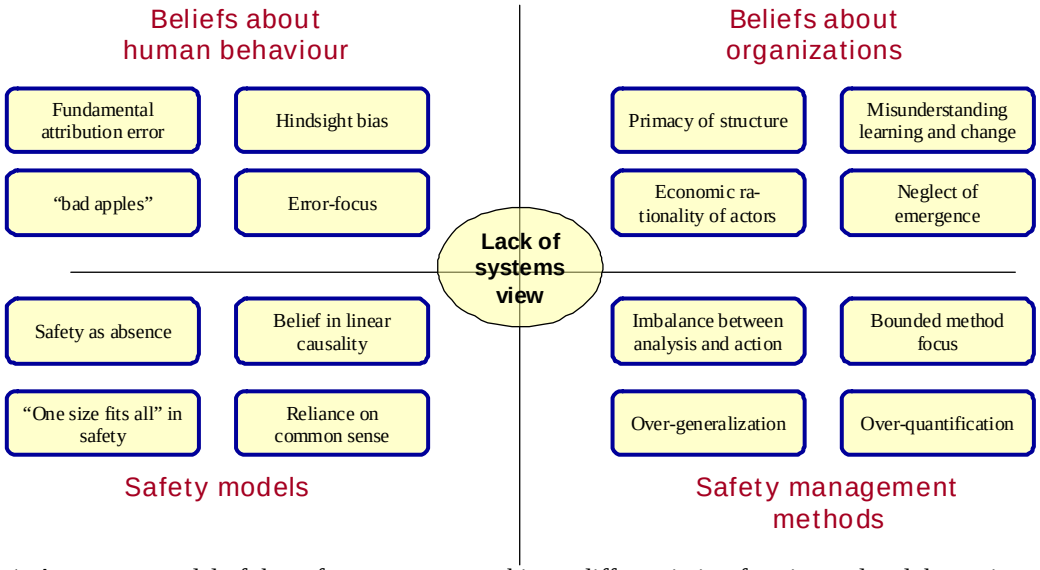

Figure 1. A systems model of the safety management biases differentiating four interrelated thematic areas

* teemu.reiman@vtt.fi

We will outline main biases of safety management in four thematic areas: beliefs about individual behavior, beliefs about organizations, safety models and safety management methods. A common theme underlying the biases is a lack of systems view on safety. Each of the thematic areas includes a number of biases that are elaborated below (figure 1). The list of biases is based on our experience in various research and development projects in different safety critical environments (nuclear, health care) as well as a review of relevant literature.

2. BELIEFS ABOUT THE NATURE OF HUMAN BEHAVIOR

Views about how humans contribute to safety often remain negative among practitioners as well as researchers [cf. 38, 27]. In addition, the practitioners' view on the "human factor" is much based on the idea of a single human being responsible for the accident [27] in contrast with adopting a more social and systemic framework. In this section we present four biases that shape this view.

2.1. Humans as fallible machines

It is often argued that over 80 percent of all accidents are caused by human errors or unsafe acts. This statement may seem reasonable at first sight, since humans design, construct, operate and maintain sociotechnical systems. This almost one hundred year old assumption still constitutes a basis for many safety initiatives [29]. The abundance of human error is often used as a justification for various "softer" methods such as behavioral programs or human factors training. In fact, in the e.g. nuclear industry, the entire concept of "human performance" is sometimes understood as basically being error prevention programs and techniques [40]. People are seen as a threat to safety because they may perform unexpected actions. This makes the *control of variation in human behaviour* one of the main challenges (though often visible only between the lines). The actions of human do not always fulfill intended goals - this phenomenon is commonly associated with the concept of "human error". The concepts of 'human error' and 'organizational factors' are but general designations for unique actions, measures or decision-making processes. They do not explain past incidents or predict the future any better than the term 'technical failure' explains or prevents disturbance situations [31]. The propensity to label negative outcomes as due to individual error tells more about general human tendencies in attributing causality than about the event itself. The attribution of error is a (social) judgment about human performance made with the benefit of hindsight [57]. A more fruitful starting point would be to treat human performance variability as a normal phenomenon behind both success and failure [19].

2.2. Fundamental attribution error

A typical human (and organizational) characteristic is a tendency to blame someone else's mistake on the basis of character (laziness, indifference, lack of ability), instead of situational or work conditions. However, people have been found to explain and justify their own behaviour differently than others' behaviour. The fundamental attribution error is a tendency for people to over-emphasize dispositional, or personality-based (internal), explanations for behaviours observed in others while under-emphasizing situational (external) explanations [11]. People have an unjustified tendency to assume that another person's actions depend on what "kind" of person that person is rather than on the social and environmental forces influencing the person. However, this same tendency does not apply to one's own behaviour when that behaviour is considered successful [11]. People claim more responsibility for successes than for failures. This bias also seems to operate on the level of social identities, which means that the successful actions of one's own group are merited to the group's characteristics, whereas failures are attributed to external conditions [40].

Studies show that the more serious the situation is (to the individual or society), the more disagreeable is the idea of the situation (e.g. an accident) happening due to pure chance. Chance also implies that the same incident could target or could have targeted me. This is why people so readily stress the fact that an incident could have been prevented and the person involved must have caused it [11, 31]. Locating the responsibility for incidents in individual decision makers also allows quick (and "dirty")

remedies such as firing, transferring or retraining the individuals [51]. A problem with these remedies is that they are seldom very effective in achieving the goal of increasing safety and long term productivity.

2.3. Bad apples – trait approaches to human behavior

People have a general tendency to perceive other people as having quite stable traits, and their behaviour as less dependent on context than their own [11]. Thus, errors and mistakes of others are perceived as being due to stable traits ("bad apples") rather than contextual reasons. Approaches having this bias often emphasize the importance of individuals' own attitudes to safety behavior. This is visible e.g. in some occupational safety campaigns where the explicit message is "you are all professional and you know how to do your work safely, thus it is a matter of attitudes whether you decide to work safely or not". Traditional behavior based safety approaches (BBS) essentially share the same underlying logic [22, 48].

Safety science research has argued for a shift to no-blame approach to safety [37, 9]. Adopting a "no-blame" culture is an idealistic approach worth discussing in more depth since the "no-blame" approach contradicts other goals and beliefs such as that people should be accountable for their actions. The basic issue concerns *under what conditions* a person should be hold responsible. This is however a much more intriguing and difficult question in comparison with the general quest for non-blame cultures. Reason [37], for example, talks about a "just culture" to highlight the trade-offs between reasonable and unreasonable blame.

2.4. Hindsight bias

In hindsight after an accident, many weaknesses existing in organisations are usually revealed. For example, it is rather common to detect various forms of "deviations" from rules and regulations. Some accident investigations practices equalize a deviation with a "cause": however, the fact that something did deviate from a prescribed rule is not necessarily a contributor to an accident or even an abnormal event. On the contrary, routine noncompliance to written procedures and local adaptations are often a norm rather than an exception [51,52]. Bourrier [17 p. 106] has argued, in the context of nuclear power plant maintenance, that "local adjustments to and re-arrangements of rules and, at times, even rule violations, are not only constant but necessary for organizations to effectively pursue their goals". Only when things go wrong are these adjustments considered negative.

Looking for human errors after an event is a "safe" choice, since one always finds them in hindsight. Looking and finding human errors makes it easier to find out who's guilty for the accident, who should be held accountable, and where preventative measures should be aimed. Unfortunately, the preventative measures are usually misaimed when "the cause" has been attributed to individual error. Accidents occur from a combination of many factors, which are not necessarily dangerous or erroneous in isolation but together or when their consequences combine, they may expose the organization to an accident. By blaming the individual, people can maintain the assumption (or an illusion) that the system is basically safe, as soon as it can get rid of the "bad apples" [cf. 9, 4].

3. BELIEFS ABOUT THE NATURE OF ORGANIZATIONS

Perspectives to the nature of the organization vary from those that emphasize the social and interpretive aspects of organizations to more rationalistic approaches focusing on structures and official routines [17]. Many theories of accidents and safety in industrial organizations are based on a static and rational model of an organization as elaborated in this section.

3.1. The primacy of structure

Researchers and practitioners have a tendency to view social phenomena such as safety management in individual terms and in terms of structure (e.g. an error) instead of process (e.g. performance

variability). Often, this tendency is due to the fact that individual phenomena and individual behavior is more apparent than social phenomena. In addition to that, structures are more stable (by definition) than processes and as such they have been easier to study.

The reality of organizational life is usually very different from that described in formal documents. This is natural in all social contexts and not dangerous by default. The search for deviations from the prescribed logic of the organization may camouflage reality since causes are attributed based on observed deviations rather than exposing contextual factors that unfold the reality of organizational activities. Antonsen et al. (4) point out that “one of the far-reaching consequences of such rationalistic approaches [that view safety as compliance with the official procedures] is that planning and pre-programming are separated from the people performing the work”. This can lead to a gap between work as imagined (by management and planners) and work as actually done (at the shop floor). Hackman and Oldham [15, p. 75], studying work motivation, present the following thought-provoking observation: "[t]he irony is that in many such significant jobs, precisely because the task is so important, management designs and supervises the work to ensure error-free performance, and destroys employee motivation ... in the process.”

3.2. The neglect of emergence

The phenomenon of emergence refers to patterns, structures or properties emerging at the system level (e.g., organization) that are difficult or impossible to explain in terms of the system’s components and their interactions [45, p. 4]. Emergent phenomena cannot be reduced to their components. On an individual level, mental properties may not be easily reduced to neurobiological properties. On the organizational level emergent phenomena include shared beliefs and practices (culture) as well as work climate [45]. These emergent phenomena affect the performance at the individual level.

For example, organizational culture has an influence on anyone working at the organization – the influence that is either positive or negative in terms of safety outcomes. Weick has emphasized that "strong cultures can compromise safety if they provide strong social order that encourages the compounding of small failures" [55, cf. 44] and further that "organizations are defined by what they ignore – ignorance that is embodied in assumptions – and by the extent to which people in them neglect the same kinds of considerations" [55, p. 74].

The climate of competition and acute cost awareness affects the way companies conduct business, but few accident investigations would blame such broad system factors (e.g. capitalism) when attributing causes for accident [cf. 21]. However, Rasmussen and Svedung [36] have for example argued that economical factors such as cost pressures in competitive environments are significant contributors to large-scale accidents. They do not, however, take any political stance rather they assume the economic environment as given and discuss the consequences for safety management in a “dynamic society”. The “political” dimension is often lacking in much safety science research and discussion [cf. 3].

3.3. Economic rationality of organizational actors

Rational economic man acts to obtain the highest possible well-being for himself. This homo economicus bases his choices on their utility; he seeks to attain very specific and predetermined goals to the greatest extent with the least possible cost. In other words, he seeks to maximise utility by weighing every choice against every other choice and choosing the choice that brings the most positive outcomes in terms of the goals that he has set. This scientifically unproven belief is alive in many management approaches. Still, according to the seminal work by March and Simon, decision-making in real life situations is controlled by limited or bounded rationality and the search for a satisfactory solution instead of a perfect one. Also, recent research has provided compelling evidence that decision-making in natural work situations is seldom synonymous with conscious selection between different alternatives. The available tools, the environment, people and the terminology used affect the perceptions and interpretations of individuals [31, 23, 26]. Furthermore, risk perception is influenced by the employee's duties, as well as his or her department and work role [1]. Thus, people

may observe risks in their organization in systematically different ways. In practice, organizations often engage in activities that seem non-rational: politics, power struggle and 'entertainment'. With hindsight, such activities may have led to useful new ideas or solutions to problems. At other times, organizations may face problems because they use methods and thought patterns that have traditionally worked well but are no longer suitable due to changes in the environment. Internal power conflicts and lack of focus on important issues can also cause safety consequences. This is why the 'non-rational', emotional and political sides of organizational activities should not be excluded from safety management approaches.

3.4. Misunderstanding learning and change

Some approaches to safety management are based on a static view of the organization. They aim at guaranteeing that nothing has changed, and that all the safety measures are still in place. They do not typically acknowledge the inherent change of sociotechnical systems and the fact that yesterday's measures may be today's countermeasures [40]. These approaches consider organizations as changing only when the management decides upon a new structure or process. Otherwise the organization is statically carrying out its tasks in the way the processes dictate, implies this approach. Weick [53, 54] has emphasized that instead of speaking of *organization*, we should speak of *organizing*. What we perceive as an organization is the (temporary) outcome of an interactive sense-making process [53]. Even heavily procedural complex sociotechnical systems adapt and change their practices locally and continually [cf. 7, 9]. Routines and practices develop over time even without any noticeable change pressure. People optimize their work practices, come up with shortcuts to make their work easier and more interesting, lose interest in the commonplace and reoccurring phenomena, and have to make tradeoffs between efficiency and thoroughness in daily tasks [19]. Sometimes change is needed in order to keep things stable in the organization, i.e. to counteract external change pressures and internal gradual drift of work practices. If everyday work requires too much adaptation and improvisation then the system is unstable and change management is needed to stabilise it.

Defining failure and success are social and political processes, and by reinterpreting the history, each can be turned into another [6, 54]. Sagan [44] reminds researchers and practitioners of "the resourcefulness with which committed individuals and organizations can turn the experience of failure into the memory of success". For successful organizations the danger is in developing a complacent attitude if the future is considered an automatic repetition of history. Learning is more than an accumulation of knowledge; it involves a continuous change and development of thinking (and action) in a specific operating environment. Learning also does not mean simply an accumulation of (work) experience. Long experience does not necessarily and automatically lead to more advanced models of thinking and action, but may rather result in restricted routines that are difficult to change. Connected to learning is the issue of certainty versus uncertainty. It is often emphasized in safety critical domains that one should know what one is doing or stop the work. This makes the handling of uncertainties a personal question linked to professional competence [31]. In terms of competence development and sense of control it is important to understand that uncertainty is never caused by an individual alone but is rather related to the object of work and to the characteristics of the complex sociotechnical system where the work is carried out [30, 31].

4. BIASES IN UTILIZING SAFETY MANAGEMENT METHODS

Often the available methods (and their implicit assumptions) dictate what to look for and analyze, instead of the phenomena dictating what kinds of methods one should utilize. Still, the methods for risk assessment and accident investigation typically develop more slowly than the sociotechnical systems they are supposed to help in managing. Thus, methods often "lag behind reality" [20].

4.1. Over-quantification

Being able to quantify a variable is usually (and rightly) perceived as a hallmark of control. The validity and reliability of various measures associated with risk and safety (and their associated

models) is a much-discussed subject in safety science. Naturally, numbers are no problem (or a solution) in themselves but rather how one makes sense of these numbers. In particular it is important, as far as possible, to make hidden assumptions explicit and to reveal the uncertainties associated with safety measures. Secondly, performance indicators that are easy to quantify may divert attention from more subtle but important issues - such as issues of power in organizations. Also many issues of subjective risk and "gut feelings" are difficult to monitor by quantitative indicators. People may experience that "something is wrong" or "missing" but without being able to clearly communicate what it is. If an organisation consequently dismisses such reports as reflection of general complaints rather than something that actually could be a vague perception of an existing risk, then subtle but existing risk might prevail although they in fact already have been detected. What we often call "intuition" is not some mystical faculty of the mind but rather a consequence of experience based pattern recognition that can express itself rather vaguely as a feeling of recognition [25]. Thirdly, without any underlying model describing the postulated *causal relations* among a set of performance measures it is indeed difficult to know why a change has occurred (learning) and what it implies for safety management strategies. One example of over-quantification relates to the "iceberg" folk models of safety where small incidents are counted in the hope of predicting when a more serious event will take place. Another bias involves a type of binary thinking where an existence (1 or 0) of a system or procedure takes precedence over its actual functioning [cf. 4], e.g. in auditing or development.

4.2. Imbalance between analysis, interpretation and action

When safety related problems are identified they have to be "solved" or coped with in some way or the other. Results from the problem finding processes should thus be transformed into reliable and robust solutions since there is of course no point in having an effective problem identification process if its output is not used as a basis for remedial actions. Also, it is not uncommon for organisations to collect masses of data but with no or little subsequent utilization. Turning data into information and action often presents a bigger problem than implied in safety management manuals. It is sometimes tacitly assumed that solutions are found more or less directly from the problem statements and that people engaged in accident investigations are therefore assumed to present a set of recommendations. However, it is often the case that accident investigation manuals give very little attention to problem solving activities [28] – standard recommendations such as more instructions and more training are not necessarily those that in the long run produce reliable solutions. This bias manifests also with safety performance indicators where a lot of time and effort is devoted to the collection of indicator data and much less effort on the subsequent interpretation of what the data is indicating [see 41].

Another bias associated with problem analysis is the emphasis on having "all the facts" before acting. This is related to the organization's way of dealing with uncertainty. This emphasis can paralyze an organization when it has to deal with issues where it is impossible to ever completely remove the uncertainty by collecting more data. Finally, the selection of countermeasures or corrective actions that match the level of perceived or analysed problem or threat is challenging. There is always the danger of over-reacting or under-reacting to the signals.

4.3. Bounded method focus

Sometimes there is an emphasis on safety management that everything has to be proceduralized. Correct performance can then be defined as strict adherence to rules and procedures. This same bias can apply to safety management tools and methods. For example, human error prevention and human performance enhancement tools often embed a narrow view on safety and human behaviour as being compliance to rules and visible safe behaviour. At their best, these tools can improve understanding and be a part of competence development if they are based on systemic view on safety and used in the right way. Unfortunately the underlying models of safety and safe performance of these tools are not always made explicit. Also some methods used for event investigation and risk analysis can prevent safety management development if they are not regularly checked for fitness to the actual situation and the changes in environment that might have occurred. When complexity increases in an organisation this could mean that the tools for event investigation, risk analysis etc, should be reviewed, perhaps

other methods are more suited as a result of enhanced complexity. Also ideas for organisational arrangements, including organisational innovations popular at the time, should be reviewed for fitness to the current situation. Unfortunately not all management innovation adopted for increasing productivity and efficiency are similarly suited applied in a risk management context.

Another side of the bounded method focus relates to abandonment of all concepts that cannot be operationalized and measured precisely. Dekker and Hollnagel [10] also warn against rejecting outright the human factors folk models. Some of the folk models may be able to generate useful empirical results, if only given time and opportunity to do so. What is important is that the falsifiability of the new explanations is larger than that of previous concepts and models.

4.4. Over-generalization

There is often an implicit assumption of context-free cures for analysed events in accident analysis manuals that provide "boxes" with text such as "suggestion of remedial action" or "follow up" as a context free chain of activities. The interface problems between the different stages in this process are typically not mentioned at all, or to a small extent. Moreover, the typical manual gives very little advice to the practitioner about how the identified weaknesses can be conquered and how the remedial actions should be selected. Non-contextual models serve the purpose of abstracting common themes among branches but they can be dangerous if they camouflage specific hazards or tasks. Another example of over-generalizing concerns different types of hazards; i.e., an increase in occupational accidents is postulated to mean that the risk of a serious process or production related accident has increased. These iceberg or pyramid models of safety postulate a causal relation between the causes of small injuries and causes of large accidents that scientific evidence does not corroborate [49].

Change management literature has many linear and non-contextual models for achieving change in organizational performance [42]. However, they seldom consider the fact that organizational changes in structures or practices are always interpreted and experienced in the light of the existing practices and culture of the organization. Their effect on performance is thus dependent on the meanings attributed to them by the personnel. This means that a structure or practice that has worked well in another organization or another industry can be interpreted and experienced totally differently in the existing culture. This has to be taken into account when borrowing solutions from another industries or organizations.

5. BIASES IN UNDERLYING SAFETY MODELS

Research has shown that concepts such as human factors, safety management, accident, or safety culture have different meanings, definitions and usages within the practitioners as well as within the research community [29, 9, 49, 50, 24, 28]. In this section we will deal with biases in safety models.

5.1. Reliance on beliefs and folk models

Steele and Pariés [50] point out that some of the common assumptions about aviation safety [prevalent in the field] are either false or do not hold under certain conditions. Examples of the kind of assumptions they refer to are: 'humans are a liability' or 'accidents occur as a linear chain of events' or 'following the procedures guarantees safety'. These assumptions were assumed to be 'truths' and were taken for granted without most people even being aware of them or considering them possible points for debate. Many of the models and methods currently in use are based on these assumptions. The authors conclude that the current methods do not meet the needs of the modern aviation industry and may in fact prevent further progress.

The accident models that personnel have incorporate beliefs about accidents and the human contribution to safety. When these beliefs are implicit, they might be "dangerous" in the sense that people might have misconceptions about safety. For example, people may believe that there is a direct one-to-one relation between the elimination of a "cause" and the elimination of a resulting "effect".

For instance, models suggesting causal influence of management decisions into work-conditions, producing human errors that lead to accidents, largely *neglect the reciprocal influence patterns* among objects and events both within and between different levels of explanation. This, in turn, may evoke false beliefs about the strengths of the remedial actions suggested as well as that the nature of the identified "causes". For example, simplified "human factor" solutions in terms of "more training and more instructions" frequently appear in event analysis reports - reports that otherwise might have strong technological biases in terms of identified causes.

Experience often narrows one's point of view to some "pet" theories or solutions that have worked well in the past. As experience accumulates, people learn what works well and what does not. These well-working solutions became personal preferences that are then applied to a wide range of situations. Relying on experience makes sense in many cases, but it has its drawbacks. Experience and the implicit beliefs formed with it have a strong influence on what the safety specialist subsequently pays attention to, what he considers important, and what he ultimately finds out. Dekker and Hollnagel [10] write: "The greatest risk of folk models is that they appear to make sense, even though statements and conclusions may not be falsifiable. They may therefore seem more plausible than articulated models, since the latter require an understanding of the underlying mechanisms." [10, p. 84].

5.2. Belief in linear causality

As noted by e.g. Hollnagel [18, 20] many accident models still share a linear philosophy in describing accident causation and fail to focus on the dynamic interplay among factors. Models of linear causality propose that the effect is proportional to cause; the larger the cause the larger the effect. This means that serious effects (e.g. accidents) are believed to be caused by serious, or big, causes (e.g. major negligence or ineptitude). Instead, in non-linear systems, according to systemic models, small causes can produce arbitrarily large effects [cf. 33]. The outputs are not necessarily proportional to the inputs. Rollenhagen [43] points out that when dealing with people, technology and organizations we are dealing with causally interdependent categories and it makes little sense to attribute (generic) causal primacy to any of the categories in safety models: "To depart from technology in itself without recognition of its interaction with human and organisation makes little sense, and departing from ''culture" in itself without understanding how technology and organisations shape beliefs, moral, values, attitudes and behaviours is also problematic" [43].

Causal explanations of incidents and accidents have implications for organizational control [51, 32]. Locating the responsibility for incidents in individual decision makers allows quick (and "dirty") remedies such as firing, transferring or retraining the individuals [51, p. 392]. These methods of "developing" organizational safety mask important systemic issues behind the incidents. An implicit belief in linear causality is also a common folk model. The same problem occurs in the "deeper" theories of organizational culture [46], where future behaviour is predicted as a repetition of past behaviour or manifestation of assumptions born out of past behaviour.

5.3. One size fits all in safety science

Barley and Kunda argue that since the dawn of systems theory in the end of the sixties "work has slipped increasingly into the background as organizational theory converged on the study of strategies, structures, and environments as its central and defining interests" [5, p. 76]. Rasmussen also points out that even in the safety critical area, "management theories tend to be independent of the substance matter context of a given organization" [35, p. 872].

According to many academic organizational researchers [16, 14, 8] the concept of safety culture has become a catch-all concept for psychological and human factors issues in sociotechnical systems. The critique e.g. expresses a concern that safety culture is not seen as a contextual phenomenon, but as a kind of general ideal model without adequate consideration of the work itself being carried out in the organization in question. Furthermore, the specific features of a safety culture inherent in a specific branch or task may be neglected in general safety management practices. But there is also a problem in

the opposite direction: various arenas of safety have produced more or less self-contained regulatory regimes. Depending on national laws and regulations, we can find that "occupational safety", "patient safety", "radiological safety" in fact becomes so context dependent that the common features among these different safeties becomes a problem. Various safeties may in fact "compete" with each other in the same organization – attempts to satisfy the demands in one area can lead to sacrifices in another area of safety. A more contextual approach is needed that emphasizes simultaneously the productivity, safety and health of the sociotechnical system, i.e. takes into account the core task of the organization [39]. Organizational core task denotes this work in context and its understanding is important for safety managers as well as other personnel. This apparent paradox between context free generalizations (e.g. general accident models, generic safety culture dimensions etc) and specific context oriented regulatory demands is by no means easy to handle. The safety management systems must be sufficiently fine grained to incorporate specific hazards, tasks etc. found in a various task domains but at the same time general enough to integrate various types of safeties found in many complex socio-technical systems.

5.4. Safety as absence of accidents, errors, deviations and uncertainty

It is surprising how often the definition of "safety" is taken-for-granted. In practice the different definitions of safety that are used explicitly or implicitly affect the safety management priorities and practices. Many implicit models embed an idea of safety as an absence of something or the missing inadequacy of something, e.g., the fewer the number of unplanned scrams or INES rated events, the higher the safety level at a nuclear power plant. Another bad example would the using the number of human errors to postulate the safety level, i.e. the fewer human errors the higher the safety level. It can be argued that safety should also refer to presence of something, not only absence [18, 43].

Many safety management approaches seem to depart from the often implicit assumption that reliability is synonymous with avoiding errors. People are perceived as a threat to safety because they may perform unexpected actions, neglect or bend the rules, forget things, miscalculate, act before thinking things through, and so on. This makes the reduction of variation in human performance one of the main challenges and goals of management. This is a problematic viewpoint. The variation, adaptability and innovation inherent in human activities enable complex organizations to carry out their tasks. More often than causing hazards, people probably carry out their duties exactly as they should, fixing defects in technology, compensating for bad tool or work design or stopping a dangerous chain of events based on intuition [31, 19]. This is why heavy constraints on normal human behaviour would most likely erode the activities of organizations and reduce work motivation. Safety critical organizations are so complex and deal with such technically difficult phenomena that it is unrealistic to think that all of the uncertainties could be removed. In our view, however, organizations do not discuss this aspect very much [cf. 32, 31]. The responsibility for dealing with uncertainties and doubts as to whether all of the consequences of activities are known is often left to the work groups and individuals in charge of actual work.

Grote [12] discusses two basic approaches to managing uncertainty in organizations: (1) minimizing uncertainty, and (2) coping with uncertainty. The first strategy is mainly based on a feed-forward control i.e. relying on planning and monitoring of plans. By and large, this strategy has been implemented by means of detailed rules and regulations intended to guide actors through complex task domains – little freedom is allowed for local initiatives. The second approach, associated with open system theories, focus on giving actor freedom to locally cope with uncertainty by means of feedback control. One of the assumptions here is that "local actors need to be given as many degrees of freedom as possible, achieving concerted action mainly through lateral, task-induced coordination" [12, p. 268]. Another way to frame these two strategies is in terms of the distribution of autonomy and control [13] interpreted as self-determination regarding rules and rules to follow (autonomy) in contrast with control "as the influence on a given situation allowing to reach goals which have been determined either autonomously or by others [12, p. 268]. Finding a proper balance among different safety management strategies associated with uncertainty is at the core of many safety problems: a system should be designed to allow for prediction and control but at the same time be flexible enough to

adapt, innovate and learn [34, 58]. One of our interests in this line of thinking which is of relevance for "biases" in safety management thinking concerns various beliefs about "the human factor": if human is dominantly portrayed as a risk factor, then it makes sense to limit variability to fight the "human error". But since human initiative and flexibility also provides the fuel for avoiding disasters, a safety management system must provide sufficient flexibility to support safe adaptation and learning.

6. LACK OF SYSTEMS VIEW

Concepts such as safety, reliability, or human factors are not absolute; rather, organizations construct their meaning and act in accordance to this constructed meaning [40]. For example, if the organization socially constructs a view that the essence of safety is to prevent individuals - the weakest links in the system - from committing errors, the countermeasures are likely to be targeted at individuals and include training, demotion and blaming. When the view is not based on scientific or empirical evidence and its functioning is not reflected in the organization we can speak of it being a bias.

The reciprocal causality of technology and the human elements of the system tend to be neglected in safety management field. People create technology, structures, and processes, which in turn influence how people think, feel, and act. The members of an organization assign meanings and beliefs to organizational elements (structures, systems and tools, others' behaviour) and these assigned meanings in turn influence the ways in which the members behave [48, 2, 54]. Even the technological solutions and tools are given meanings by their designers and users, which affect their subsequent utilization.

Conflicting goals are one of the sources of risky behaviour even in well-intentioned organizations. Safety and profitability, efficiency and thoroughness or occupational safety and production (or process) safety may conflict in real life. On the other hand, the long-term goals of the organization do not necessarily conflict, for example, in the sense that improvements to safety would always threaten financial profitability. Still, the goals do have different time perspectives in terms of return on investments. The challenge is not in determining whether safety and economy are in conflict, but in the ability to balance resources and focus attention on issues that most deserve it to ensure long term effectiveness - be they safety, production or personnel-related.

7. CONCLUSIONS

At worst, the biases depicted in this article can lead to an approach where people are treated as isolated and independent actors who make (bad) decisions in a social vacuum and who pose a threat to safety. Such an approach aims at building barriers and constraints to human performance and neglects the measures aiming at providing prerequisites and organizational conditions for people to work effectively. This reductionist view on safety management can also lead to a too strong separation of so called human factors from technical issues, undermining the holistic view on system safety. The five thematic areas, typical biases and their potential safety consequences are illustrated in Table 1.

A systemic safety management takes into account people, technology and organization and their interaction in equal terms. Furthermore, such an approach can shift focus from people to technology to organizational aspects depending on their current safety significance. More attention needs to be devoted to understanding why things are usually done well enough at the organization instead of looking at why things went wrong. Also in accident investigations it is important to inspect why people acted the way they did, and why did it make sense to them at that time [cf. 9, 19, 20].

Reflection is an important aspect of safety management. Safety managers, developers and safety scientists should reflect from time to time what assumptions they have concerning individual, organizations and safety. They should also reflect on the assumptions embedded in the methods they use; do these allow systemic issues to emerge or are they biased toward some type of phenomena.

Table 1: The five thematic areas and typical biases and potential safety consequences associated with those areas

Thematic area	Bias / phenomenon	Potential safety consequence
1. Human behavior	Fundamental attribution error	Divert attention from contextual factors, which may have safety significance. People are not treated as social beings, but rather as isolated actors making decisions in a social vacuum.
	Trait-based theories of humans – "bad apples"	People who commit mistakes are considered lazy, stupid or careless by character.
	Hindsight bias	Distorted view about factors that influence safety may lead to wrong priorities.
	Focus on human error	Strong and adaptive activities are not reinforced and context is ignored, which can lead to wrong remedial actions.
2. Organizational behavior	Causal powers of official structures	The effects of safety management actions are misunderstood and the actions have unexpected consequences.
	Misunderstanding learning and change	Failure to perceive the gradual change and optimizing of practices taking place all the time at the organization.
	Economic rationality of organizational actors	Failure to take into account the individual and collective motives that influence employees' behavior.
	Neglect of emergence	Failure to perceive the influence of emergent phenomena such as shared beliefs or norms on individuals and groups.
3. Methods	Bounded method focus	Strict adherence to rules and procedures may blind the person to issues not covered by the rules.
	Over-quantification	Qualitative phenomena, feelings, hunches are neglected.
	Imbalance between analysis and action	The organization does not turn its information into actionable knowledge and as a consequence is unable to develop its practices based on the information.
	Over-generalization	Using solutions that have worked for one problem or one situation to areas that they are unsuitable.
4. Models	Belief in linear causality	Failure to perceive the dynamics in the system and the interaction between people-technology-organization
	"One size fits all" – neglect of context and core task	Specific requirements of the given task and hazards are neglected. Also, the effect of safety initiatives aimed at e.g. occupational safety on e.g. process safety are not considered..
	Reliance on common sense	The employee does not question his/her beliefs since (s)he does not believe in beliefs / theories
	Safety as absence	Relying on absence may not reveal the underlying dynamics in systems that are essential for creating safety. Uncertainty management may be one-sided (reduction of uncertainty) without realizing the need for flexibility and adaptation.
5. Lack of systems approach		Systemic and emergent issues that affect organizational activity and can contribute to multiple events are neglected in the safety management initiatives.

Acknowledgements

The writing of the article was funded by the Nordic nuclear safety research (NKS). The authors would like to acknowledge the feedback from our colleagues Elina Pietikäinen, Pia Oedewald and Ulf Kahlbom.

References

[1] ACSNI. "*Organising for safety. Third report of the Human Factors Study Group of the Advisory Committee on Safety in the Nuclear Industry*", Health & Safety Commission, HMSO, 1993, London.
[2] M. Alvesson, "*Understanding organizational culture*", Sage, 2002, London..
[3] S. Antonsen. "*Safety culture and the issue of power*", Safety Science, 47, pp. 183-191, (2009).
[4] S. Antonsen, P. Almklov and J. Fenstad, "*Reducing the gap between procedures and practice – lessons from a successful safety intervention*", Safety Science Monitor, 12 (1), pp. 1-16, (2008).
[5] S.R. Barley and G. Kunda. "*Bringing work back in*", Organization Science, 12, pp. 76-95, (2001).
[6] P. Baumard, "*Presentation at the HRO 2007 conference at Deauville"*, France, May 29-31, 2007.
[7] M. Bourrier, *"Organizing maintenance work at two American nuclear power plants*", Journal of Contingencies and Crisis Management, 4, pp. 104-112, (1996).
[8] S. Cox and R. Flin. "*Safety culture: Philosopher's stone or man of straw*?", Work & Stress, 12, pp. 189–201, (1998).
[9] S. W. A. Dekker, "*Ten questions about human error. A new view of human factors and system safety*", Lawrence Erlbaum, 2005, New Jersey.

[10] S. Dekker and E. Hollnagel. "*Human factors and folk models*", Cognition, Technology & Work, 6, pp. 79-86, (2004).
[11] S.T. Fiske and S.E. Taylor, "*Social cognition. 2nd edition*", McGraw-Hill, 1991.
[12] G. Grote. "*Uncertainty management at the core of system design*", Annual Reviews in Control, 28, pp. 267-274, (2004).
[13] G. Grote, "*Autonomie und Kontrolle – Zur Gestaltung automatisierter und risikoreicher Systeme* (Autonomy and Control – On the design of automated and high-risk systems.)", vdf Hochschulverlag, 1997, Zürich.
[14] F.W. Guldenmund, "*The nature of safety culture: A review of theory and research*", Safety Science, 34, pp. 215–257, (2000).
[15] J. R Hackman and G. R. Oldham. "*Work redesign*", Addison-Wesley, 1980, Reading, Mass.
[16] A. R. Hale, "*Editorial. Culture's confusions*", Safety Science, 34, pp. 1-14, (2000).
[17] M.J. Hatch and A.L. Cunliffe, "*Organization Theory: Modern, Symbolic and Postmodern Perspectives. Second Edition*", Oxford University Press, 2006, Oxford.
[18] E. Hollnagel, "*Safety management - looking back or looking forward*", In E. Hollnagel, C.P. Nemeth and S. Dekker (Eds.), Resilience Engineering Perspectives, Volume 1. Remaining sensitive to the possibility of failure. Ashgate, 2008, Aldershot.
[19] E. Hollnagel, "*The ETTO principle: Efficiency-thoroughness trade-off*", Ashgate, 2009, Farnhamn.
[20] E. Hollnagel and J. Speziali, "*Study on Developments in Accident Investigation Methods: A Survey of the 'State-of-the-Art'*", SKI Report 2008:50, SKI 2008.
[21] A. Hopkins, "Lessons from Longford. The Esso gas plant explosion", CCH, 2000, Sydney.
[22] A. Hopkins. "*What are we to make of safe behaviour programs*?", Safety Science, 44, pp. 583-597 (2006).
[23] E. Hutchins, "*Cognition in the wild*", MIT Press, 1995, Massachusetts.
[24] P. Katsakiori, G. Sakellaropoulos and E. Manatakis, "*Towards an evaluation of accident investigation methods in terms of their alignment with accident causation models*", Safety Science, 47, pp. 1007-1015, (2009).
[25] G. Klein, "*The recognition-primed decision (RPD) model: Looking back, looking forward*", In C. E. Zsambok and G. Klein (Eds.), Naturalistic decision making. Lawrence Erlbaum, 1997, Mahwah, NJ.
[26] G.A. Klein, J. Orasanu, R. Calderwood and C.E. Zsambok, (Eds.), "*Decision making in action. Models and methods*", Ablex Publishing Corporation, 1993, Norwood, NJ.
[27] N. Korolija and J. Lundberg. "*Speaking of human factors: Emergent meanings in interviews with professional accident investigators*", Safety Science, 48, pp. 157-165, (2010).
[28] J. Lundberg, C. Rollenhagen and E. Hollnagel. "*What-You-Look-For-Is-What-You-Find – The consequences of underlying accident models in eight accident investigation manuals*", Safety Science, 47, pp. 1297-1311, (2009).
[29] F.A. Manuele, "*On the practice of safety. Third edition*", John Wiley & Sons, 2003, New Jersey.
[30] L. Norros, "*Acting under uncertainty. The core-task analysis in ecological study of work*", VTT Publications 546. VTT, 2004, Espoo.
[31] P. Oedewald and T. Reiman. "*Special characteristics of safety critical organizations. Work psychological perspective*", VTT Publications 633. VTT, 2007, Espoo.
[32] C. Perin, "*Shouldering risks. The culture of control in the nuclear power industry*", Princeton University Press, 2005, New Jersey.
[33] C. Perrow, "*Normal accidents: Living with high-risk technologies*", Basic Books, 1984, New York.
[34] J. Rasmussen, "*Risk management in a dynamic society: A modeling problem*", Safety Science, 27, pp. 183-213, (1997).
[35] J. Rasmussen, "*Human factors in a dynamic information society: where are we heading*?", Ergonomics, 43, 869-879, (2000).
[36] J. Rasmussen and I. Svedung, "*Proactive Risk Management in a Dynamic Society*", Swedish Rescue Services Agency, 2000, Karlstad.
[37] J. Reason, "*Managing the risks of organizational accidents*", Ashgate, 1997, Aldershot.
[38] J. Reason, "*The human contribution. Unsafe acts, accidents and heroic recoveries*", Ashgate, 2008, Farnham.
[39] T. Reiman and P. Oedewald, "*Assessment of Complex Sociotechnical Systems – Theoretical issues concerning the use of organizational culture and organizational core task concepts*", Safety Science, 45, pp. 745-768, (2007).
[40] T. Reiman and P. Oedewald, "*Evaluating safety critical organizations. Focus on the nuclear industry*", Swedish Radiation Safety Authority, Research Report 2009:12, SSM, 2009, Stockholm.
[41] T. Reiman and E. Pietikäinen, "*Indicators of safety culture - selection and utilization of leading safety performance indicators*", Swedish Radiation Safety Authority, in press, Stockholm.
[42] T. Reiman, P. Oedewald, C. Rollenhagen and U. Kahlbom, "*Management of change in the nuclear industry. Evidence from maintenance reorganizations*", NKS-119. Nordic nuclear safety research, 2006, Roskilde.
[43] C. Rollenhagen, "*Can focus on safety culture become an excuse for not rethinking design of technology*?", Safety Science, 48, 268-278, (2010).
[44] S.D. Sagan, "*The limits of safety. Organizations, accidents, and nuclear weapons*", Princeton University Press, 1993, New Jersey.
[45] R.K. Sawyer, "*Social emergence. Societies as complex systems*", Cambridge University Press, 2005, Cambridge.
[46] E.H. Schein, "*Organizational culture and leadership. Third edition.*", Jossey-Bass, 2004, San Francisco.
[47] M. Schultz, "*On studying organizational cultures. Diagnosis and understanding*" Walter de Gruyter, 1995, Berlin.
[48] T. A. Smith. "*What's Wrong with Behavior Based Safety*", Professional Safety, (1999).
[49] P. Smith, "*Hits and myths*", The Safety & Health Practitioner, 24, pp. 49-52, (2006).
[50] K. Steele and J. Pariés, "*Characterisation of the Variation in Safety Beliefs across the Aviation Industry*", 3rd Symposium on Resilience Engineering. Juan-Les-Pins, France, October 28-30, 2008.
[51] D. Vaughan, "*The Challenger launch decision*" University of Chicago Press, 1996, Chicago.
[52] D. Vaughan, "*The dark side of organizations: Mistake, misconduct, and disaster*", Annual Review of Sociology, 25, pp. 271-305, (1999).
[53] K.E. Weick, "*The social psychology of organizing. 2nd ed*.", Addison-Wesley, 1979, Reading, MA.
[54] K.E. Weick, "*Sensemaking in organizations*", Sage, 1995, Thousand Oaks.
[55] K.E. Weick, "*Foresights of failure: an appreciation of Barry Turner*", Journal of Contingencies and Crisis Management, 6, pp. 72-75, (1998).
[56] K.E. Weick and K.M. Sutcliffe, "*Managing the unexpected. Resilient performance in an age of uncertainty. Second Edition*", Jossey-Bass, 2007, San Francisco.
[57] D.D. Woods, L.J. Johannesen, R.I. Cook and N.B. Sarter, "*Behind human error: Cognitive systems, computers, and hindsight*", State-of-the-Art Report. SOAR CSERIAC 94-01. The Ohio State University, 1994, Ohio, Columbus.
[58] D.D. Woods and L.G. Shattuck, "*Distant supervision-local action given the potential for surprise*", Cognition, Technology & Work, 2, pp. 242-245, (2000).