TUDelft

Delft University of Technology

Safety science, a systems thinking perspective
From events to mental models and sustainable safety

Blokland, Peter; Reniers, Genserik

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Safety Science, a Systems Thinking Perspective: From Events to Mental Models and Sustainable Safety

Peter Blokland [1,*] and **Genserik Reniers** [1,2,3]

1   Safety & Security Science Group (S3G), Delft University of Technology, 2628 BX Delft, The Netherlands; G.L.L.M.E.Reniers@tudelft.nl
2   Center for Corporate Sustainability (CEDON)—KULeuven—Campus Brussels, B-1000 Brussels, Belgium
3   Department of Engineering Management, Faculty of Applied Economic Sciences (ENM), University of Antwerp, B-2000 Antwerp, Belgium
*   Correspondence: P.J.Blokland@tudelft.nl; Tel.: +32-496-236-222

check for updates

Abstract: In the past one hundred years, concepts such as risk, safety and security have become ever more important and they represent a growing concern in our society. These concepts are also important subjects of study to enhance sustainability. During the past fifty years, safety science has gradually developed as an independent field of science. In this period, different concepts, theories, models and research traditions have emerged, each with its specific perspective. Safety science is now focused on finding ways to proactively achieve safety versus reaching safety in a reactive way. We think this increasing awareness and search for proactiveness can be found and presented when viewed in light of the systems thinking iceberg model, where increasing awareness and proactiveness can be seen as digging deeper into this systems thinking iceberg, discovering the levels of systems, structures and ultimately the mental models that are "below the waterline". It offers a way forward in understanding, and proactively managing, risk, safety, security and sustainable performance, in organizations and ultimately in society as a whole.

Keywords: sustainability; safety; security; risk; performance; systems thinking; mental models

1. Introduction

1.1. Purpose of This Paper

Every element and part of the universe can be regarded as a system. When humans and technology are involved, one can also look at these elements as being socio-technical systems (STS). These systems range from very simple to ultra-complex. Hence, the purpose of this paper is to look at safety from a systems thinking perspective and to indicate how such a perspective can be used to consider safety and sustainability issues today.

1.2. Systems

Meadows describes a system as a set of things—people, cells, components, molecules, and so on—interconnected in such a way that they produce their own pattern of behaviour over time. A system may be buffeted, constricted, triggered, or driven by outside forces. But the system's response to these forces is characteristic of itself and seldom simple. As such, a system can be defined as "an interconnected set of elements that is coherently organised in a way that achieves something" [1]. Consequently, a system consists of three kind of things: "elements", "interconnections" and a "function" or "purpose".

1.3. Systems Thinking

Arnold and Wade state that "systems thinking" was coined by Richmond in 1987 [2]. However, the origins and use of the term "systems thinking" are to be found much earlier. Richmond defines systems thinking as the art and science of making reliable inferences about the behaviour of systems by developing an increasingly deep understanding of their underlying structure [3]. Others say that systems thinking is a method of considering reality in a way that helps in getting a better understanding of systems and that allows to work with systems to influence the quality of life [4]. Many other definitions exist, but it is essential that system thinkers look at whole systems and their elements. This allows them to increase their ability to understand the "elements", see "interconnections" between these elements and ask "what-if" questions about possible future behaviours of systems. This insight in systems then permits them to be resourceful about system redesign [1]. Hence, systems thinking provides options to transform systems in such ways that they will behave as desired, generating wanted outcomes and creating intended value.

Systems thinking and systems engineering are different concepts. Systems engineering focuses on how to design, integrate, and manage complex systems over their life cycles, while systems thinking is a way of looking at reality, a language and a set of tools to understand systems. As such, at its core, systems engineering utilizes systems thinking tenets. Furthermore, systems thinking recognizes that relationships between a system's components and the environment are as important (in terms of system behaviour) as the components themselves. It looks at, and determines, feedback loops, considers emergent properties and complexity, takes into account hierarchies and self-organisation, and tries to understand the dynamics and their (un)intended consequences [5].

Systems thinking is rooted in general systems theory (GST) and systems dynamics (SD). General systems theory was introduced by von Bertalanffy more than 50 years ago [6]. It is an interdisciplinary practice that describes systems with their interacting components, applicable to biology, cybernetics and many other fields of science. GST is a solution to problems posed by an increasingly more complex and connected world, where traditional modes of thinking fail whenever large numbers of elements and processes interact [7]. Systems dynamics, on the other hand, was founded in 1956 by Jay Forrester at MIT. He recognised the need for better ways of understanding social systems, to improve them in the same way people use engineering to improve mechanical systems [8]. As such, systems thinking has been applied to a wide range of fields and disciplines, as it has the ability to solve complex problems that are not solvable using conventional reductionist thinking [6].

1.4. Four Levels of Thinking

Systems thinking comprises different levels of thinking. First, there is the level that deals with the basic information or data of events. It concerns, for instance, "what happened", "where", "when", "how" and "who". These are the directly observable facts, the collectable data, concerning an event. However, a richer picture can be drawn from a deeper level of thinking when the data are combined across a larger time frame, revealing patterns and trends of events. Searching for common cause effect relationships can then help in trying to explain what the causes behind these patterns and trends are. Nevertheless, a much deeper level of thinking is to reflect on how the interplay of different factors brings about the outcomes that can be observed. The critical issue at this level of thinking is to understand how these factors interact, to see the system and its elements behind the patterns and how interactions are structured. Still, there's another, much deeper level of thinking that hardly ever comes to the surface. This represents the "mental models" of individuals and organisations that influence why things should or do work, or should not or do not work [9].

Mental models reflect the beliefs, values and assumptions that individuals or organisations hold. They underlie the reasons why and how things are done. However, mental models generally remain obscure, limiting the collective understanding of issues and, hence, when not aligned, reducing meaningful communications and the development of shared visions and common action [10].

Each level of thinking has its importance and provides answers to questions. Nevertheless, it is possible to see these different and increasingly deeper levels of thinking as an indication of the ever-increasing awareness and understanding of (socio-technical) systems.

1.5. Mental Models and the Ladder of Inference

The concept of mental models goes back to antiquity, but the phrase was coined by Scottish psychologist Kenneth Craik in the 1940s. It has since been used by cognitive psychologists, cognitive scientists, and gradually by managers. In cognition, the term refers to both the semi-permanent tacit "maps" of the world which people hold in their long-term memory, and the short-term perceptions which people build up as part of their everyday reasoning process. According to some cognitive theorists, changes in short-term everyday mental models, accumulating over time, will gradually be reflected in changes in long-term deep-seated beliefs [11].

Reality, as one perceives it, is just an image in one's mind. It is the result of a whole set of mental models that consciously and subconsciously influence what one observes. Reality, as it is, is significantly more complex than what the human mind can process and comprehend at any given moment.

In their paper "Neural substrates of cognitive capacity limitations", Buschman et al. state: "Despite the remarkable power and flexibility of human cognition, our working memory—the "online" workspace that most cognitive mechanisms depend upon—is surprisingly limited. An average adult human has a capacity to retain only four items at a given time." [12].

Therefore, in order for the human mind to deal with reality, individuals must conceptualise reality by using more abstract notions. This leads to concepts that can have different meanings and that can apply to different situations. Many of these thoughts are learned at an early age and become taken for granted. They turn out to be obvious and concrete, not abstract and questionable. It leads to the so-called "ladder of inference" (see Figure 1). This is what all human beings go through in order to make sense of their world and in order to act [11,13,14].
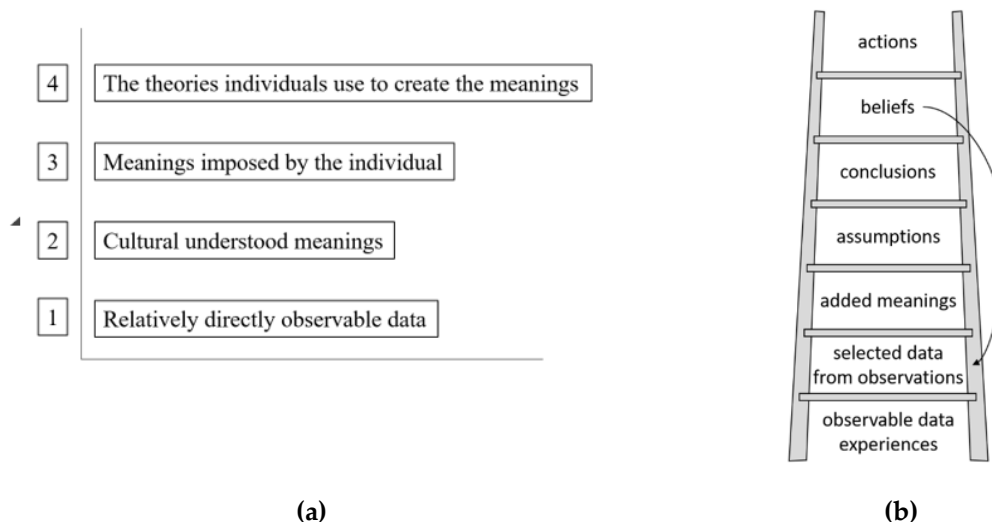

(a)   **(b)**

Figure 1. Ladder of inference according to (**a**) Argyris [13,14] and (**b**) Senge [15].

The ladder of inference (Figure 1) is a common mental pathway of increasing abstraction, often leading to misguided beliefs. The only visible parts are the directly observable elements, which are the data at the bottom of the ladder and the actions resulting from decisions at the top of the ladder. These actions are the result of self-generating beliefs which remain largely untested. One adopts those beliefs because they are based on conclusions which are inferred from what one observes, added to past experiences [11]. The whole process of added meaning, building assumptions, inferences and beliefs, shapes the mental models one holds. It also determines which added meaning is retained, what assumptions are used, how conclusions are drawn and what beliefs will govern one's actions.

As such, the ladder of inference is a reinforcing feedback loop (as shown in Figure 1), enforcing the mental models (meanings) at each level of the ladder. As a result, mental models are internal and incomplete, inconsistent representations of an external reality. They are context-dependant, can be very individual or shared and can change over time through experience and learning [16,17].

Safety, whether at an individual level, in organisations or at a societal level, is a result of the many decisions and actions taken at each moment in time. These decisions and actions are the result of the existing ladders of inference that are present at the decision-making level (individual, organisational and societal). This is why mental models influence and shape the systems that are at the origin of wanted and unwanted events happening for individuals, organisations or society as a whole.

1.6. The Systems Thinking Iceberg and Increasing Levels of Awareness

To provide for an inclusive (re)view of the increasing awareness among scientists regarding the concept of safety, and to understand the evolution of the leading perceptions in safety and security science over time, it is suitable to approach these concepts from a systems thinking perspective. As indicated in Section 1.3., systems thinking comprises four levels of thinking, each level digging deeper in the understanding of the behaviour of systems and the outcomes they produce. This evolution in increasing awareness is also to be found in the evolution of the concepts governing the ideas and approaches used in the pursuit of safety.

As a starting point for this systems thinking approach, the image of an iceberg, as shown in Figure 2, is helpful. It is often used as a metaphor for a system and its consequences.
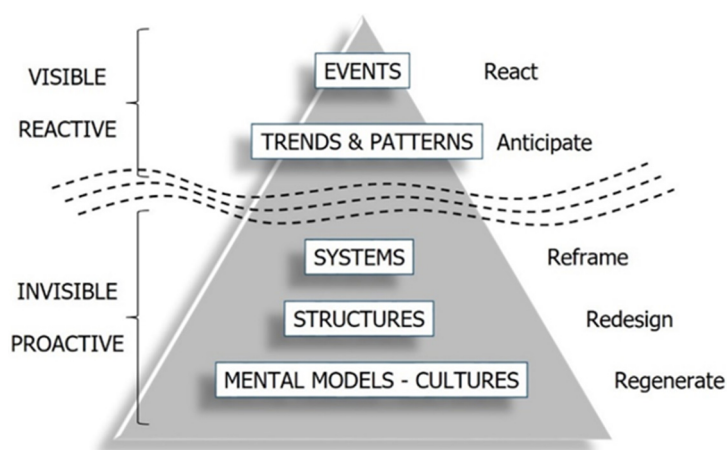

Figure 2. Systems thinking iceberg based on Bryan et al. [18].

At first, one can see the tip of the iceberg. This represents the events that occur day-to-day as a result of things happening or not happening. These visible cues at the "event" level might include failures, losses, but also the signs of created value and successes. In the case of unfortunate events, the normal reaction is to address these problem-related events on the spot and fix the issues one confronts (i.e., those one can see). Unfortunately, the effort, energy, and resources used in directly reacting to events do little or nothing to fix the fundamental causes of the problems that lie under the surface, at the base of the figurative iceberg. If one cannot address the root causes of these events, a cyclical process occurs, whereby the same problems continue to emerge regardless of how often one addresses them. The systems thinking approach suggests that interventions should be made at the root-cause level instead of just dealing with the events and symptoms one observes [19,20].

It is unclear who first adopted this iceberg metaphor in systems thinking, and different representations have been used in the past. A representation of the model that can help in getting this general view is the systems thinking iceberg model proposed by Bryan et al. [18]. This model provides a visual representation of a layered methodology to discover deeper levels of understanding and awareness. It also proposes ways on how to act at each level of awareness. In essence, the purpose of

managing risk and safety is to create socio-technical systems that generate wanted events and eliminate or avoid unwanted events. It is what is immediately visible, and these are the discernible symptoms of the (sub)system(s) at work.

2. Increasing Awareness in Managing Safety

The following sections are a concise overview of how safety science has evolved over the past 100 years, where scientists and safety practitioners discovered ever deeper layers of awareness and understanding of symptoms, patterns of events, systems, structures of systems, and finally of the mental models that are responsible for the creation of (un)wanted events.

2.1. The Objective of Profit (Creation of Value) and Preventing Loss of Value—Symptoms of Systems at Work

Both risk and safety are related sciences and started as fields of interest because systems produced outcomes that held the prospect of being very valuable, but at the same time they were also very vulnerable to losses. In both cases systems aimed at profit. For instance, trade overseas (risk and insurance) and industrialisation (health and safety) were very uncertain regarding the realisation of the wanted events (profit) and a high probability of loss (of investment and human lives) was present. Lost cargo or lost workforce were the visible events of the systems of overseas trade or industrialisation and these events required solutions in order to maintain desired profit levels. When these events are noticed, it may become clear that they repeat themselves in some way. As a consequence, it is possible to discover patterns and trends related to these events and one can act anticipating these events. This is the easiest level to work with and it is also at this easy-to-perceive level of awareness of systems that safety science and risk management originated. For risk in overseas trades, the answer was insurance. For safety in the early industries, for instance, the response was that less "accident-prone" people were required. One took actions that only dealt with the symptoms, but at that time, no solutions for the more fundamental causes of the unwanted events were envisioned. Still today, some risk and safety practitioners are driven by the visible facts that are directly observed and that are gathered in statistical data.

In their search for safety in the beginning of the past century, specialists tried to find and understand these symptoms, looking for trends or delineating recognisable patterns related to the events observed, in order to discover causal relationships and produce better predictions, allowing to prepare for these negative effects or prevent the events from happening again.

2.2. Trends and Patterns

When events are observed over a longer period of time, trends become visible when these events reproduce themselves in similar ways. At that time, the common approach to safety was to look at events such as loss of life, injury, harm, damage, or any other event generating negative effects, looking for ways to understand, predict and prevent these bad things from happening by analysing trends of unwanted events and their negative effects. As an example, the accident proneness theory [21], Heinrich's accident pyramid [22], or Rasmussen's skill- and knowledge-based approach [23] can be seen as a result of that approach and its level of awareness and understanding.

2.3. Systems

A further step in increasing awareness regarding accidents is becoming aware of the systems that produce reoccurring unwanted events. It is what every accident investigation tries to achieve, i.e., reach understanding regarding how bad things came about, in order to find ways to prevent them from happening again. When the system is understood, it is possible to proactively alter the system, so it does not produce or cause the same unwanted events again. Through history, scholars have been searching for ways to explain why unwanted events happen and how disasters can be predicted, trying to discover and describe the system(s) that is (are) behind the occurrence of unwanted events and major accidents. One of the first to develop a theory on accident causation was Heinrich with his domino

theory [24], providing categories for the elements of the system that are involved in the creation of accidents, using the metaphor of domino blocks to represent the subsystems. As such, triggered by higher values of objectives, scholars were digging deeper in the iceberg and becoming more aware of accident causation. Furthermore, Perrow's Normal Accident Theory (NAT) regarding tightly coupled and complex systems [25] and Rasmussen's Brownian movements model [26,27] can be seen as such efforts.

2.4. From Systems towards the Elements and Structures of Systems

Other, more encompassing concepts have also left their hallmark on safety science, looking at safety from a broader perspective, touching on the systems, structures and even mental models of socio-technical systems. This is done not necessarily to explain and predict but aiming at the intended outcome of resilience and high reliability to thwart misfortune.

Resilience theory is a multifaceted field of study that has already existed for many decades in different fields of science. It has been addressed by social workers, psychologists, sociologists, educators and many others. In short, resilience theory addresses the strengths that people and systems demonstrate that enable them to rise above adversity [28]. When applied to people and their environments, "resilience" is fundamentally a metaphor. Its roots are in the sciences of physics and mathematics, as the term was originally used to describe the capacity of a material or system to return to equilibrium after a displacement [29].

High reliability organisations (HROs) exhibit a strong sense of mission and operational goals, stressing not only the objectives of providing a ready capacity for production and service but an equal commitment to reliability in operations, and a readiness to assure investment in reliability enhancing technology, processes and personnel resources [30]. As such, HROs are a way to determine which elements in a system can help in becoming safer. Most of the time, these elements will also allow to increase the performance. In a sense, HROs are in fact high performing organisations (HPOs).

However, in their article "Beyond Normal Accidents and High Reliability Organizations: The Need for an Alternative Approach to Safety in Complex Systems", Marais et al. [31] stated that:

> "The two prevailing organizational approaches to safety, Normal Accidents and HROs, both limit the progress that can be made toward achieving highly safe systems by too narrowly defining the problem and the potential solutions", and that they believe that "a systems approach to safety would allow building safety into socio-technical systems more effectively and provide higher confidence than is currently possible for complex, high-risk systems".

Therefore, in order to be able to predict or to obtain more understanding and control vis-à-vis the systems involved in bad things that are happening, scientists dig ever deeper in the systemic iceberg in their efforts. This time, to determine the structure of the systems involved, aiming at getting a clearer view of the structures and the dynamics that are at the genesis of accidents.

2.5. Structures of Systems

The structures of systems are to be understood as being the elements of a whole together with the types of connections that exist between those elements. Recent years have seen a whole range of models that try to model and determine the structures of systems that generate unwanted events, searching for the elements that populate the system and trying to determine and understand how these elements interact, thus creating the dynamics that generate accidents. Some examples are the Systems-Theoretic Accident Model and Process (STAMP) [32], the Functional Resonance Analysis Method (FRAM) [33] or more recently the SAfety FRactal ANalysis (SAFRAN) method [34]. More widely used, there is also the Swiss Cheese model [35,36] and the models that build on the same human factor approach, for instance, the Software-Hardware-Environment-Liveware-Liveware (SHELL) model [37], or the Human Factors and Classification System (HFACS) [38], just to name a few. Furthermore,

theories regarding "safety culture" could also be promoted as an example of looking at systems, how they are structured and how systems generate safety or not.

2.6. Mental Models

Finally, scientists and practitioners also aim at developing an understanding of how the mental models that generate systems and their structures can be controlled and managed, in order to design methods with which systems proactively emerge, consequently generating safety and preventing unwanted events. Mental models are often described as paradigms, mindsets, beliefs, assumptions, cultural narratives, norms, expectations, or simply perceptions [20]. An example of mental models capable of generating safety is the concept of "just culture" [39,40]. This concept aims at restoring the mindset of trust and accountability in organisations. Trust and accountability are two powerful mental models, facilitating communication, reporting and a proactive behaviour towards safety and performance, that are capable of generating systems that are different from the systems where these characteristics are not shared on an organisational scale [41].

In fact, only insight in and knowledge of the mental models present in a system provide the basis for understanding how fundamental changes can be made in order to proactively obtain more safety and performance [15]. Regeneration of mental models, redesign of structures and the reframing of systems are the needed changes for improving socio-technical systems towards desired and sustainable results [42].

3. A Fundamental Understanding and Approach for Safety in Socio-Technical Systems

Safety is often defined as a dynamic non-event and mostly explained by the events that violated that state of dynamic non-events [43]. The problem with this approach is that it only covers the domain of unsafety and leaves any interpretation of safety open. When safety thinking is linked with dynamic non-events, it solely focusses on preventing bad things from happening. But is this the right approach in pursuing safety? The paradox in this approach is that something needs to happen before action is taken. When safety increases, and safety is defined in a negative way, a safe situation will lead to nothing significant happening. When this situation continues, the efforts to maintain safety at a desired level will come under pressure, as the effort will not be justified by things happening. Absence of happenings will result in a reduction of efforts, which will then lead to a reduction in resilience and reliability, making something unwanted happening more likely [44].

So, is turning away from unsafety the same as aiming for safety? When one considers a situation of 100% safety, is this a situation where nothing is happening? This seems an impossible assumption. There will always be something happening, events and consequences (positive effects on objectives) one desires and events and consequences (negative effects on objectives) one does not want. Both are important from a modern safety perspective. Moreover, understanding and managing the governing mental models about risk, safety and performance is important. So, what is particularly needed for safety to emerge, exist and persist?

A Modern Perspective on Risk, Safety and Performance

A modern perspective on risk, safety, and performance looks at the whole picture. It starts with whatever people, organisations and societies want. What are the objectives that are valued, needed and important to be safe? Safety should in the first place be concerned with making sure these valuable, needed and important objectives are pursued, achieved and secured. It is making certain that excellent performance is attained when pursuing and safeguarding objectives, and that health and wellbeing will be assured in all circumstances [45]. Hollnagel [44] refers to this as Safety I and Safety II, where Safety I is the traditional approach of avoiding losses due to the factors negatively affecting objectives. Safety II, on the other hand, is related to the variability in performances when pursuing or protecting objectives. In our view, this is how safety and safety science should evolve. It is about both the absence of losses (low unsafety), the presence of pursued, achieved and safeguarded

objectives (high safety) [45] and the mental models that are at the genesis of the systems that generate wanted and/or undesirable outcomes.

Leveson indicates that "safety" is an emergent property of systems, not a component property [32]. It means that safety is not inherent to systems, but rather something that needs to be consciously and persistently pursued, achieved and safeguarded. In a systemic perspective, systems are parts of larger systems and consist of smaller subsystems. As such, a component of a system is to be considered as a system itself, having a specific purpose different from the objective of the overarching system. Each of these (sub)systems is subjected to a specific set of risk sources that can affect those more individual objectives that also need to be safe to maintain the performance of the higher-order system(s).

A way to represent this approach and sense of risk, safety and performance, is by depicting a set of risks related to a particular objective of a (sub)system (Figure 3). Risk, as defined by ISO 31000, is "the effect of uncertainty on objectives" [46,47], and it is also stated that an effect is "a deviation from what is expected" [48]. In general, people have expectations regarding objectives and they rest on the mental models they carry, reflected in the ambitions and the attributed values these prospects carry. As Slovic states: "Risk does not exist out there, independent of our minds and cultures, waiting to be measured. Instead, human beings have invented the concept risk to help them understand and cope with the dangers and uncertainties of life. Although these dangers are real, there is no such thing as real risk or objective risk" [49]. From Slovic's statement, one can understand that risk is an individual construct, differing from one person to another because different people have different objectives and possibly value the same objectives differently. However, when one can determine the objective of a system component and make an abstraction of individual expectations and their attributed values, one could start considering a more objective approach to risk and safety that is tied to a well-determined objective, independent of the beholder and their expectations. Although this seems to be aligned with a traditional view on risk and the management of risk, it is not. Because a traditional approach will only look at the negative effects that can affect the concerned objective, judged from the governing individual, corporate or societal mental models. This is also true for objectives one is not aware of. On the other hand, with the proposed approach, risk still needs to materialise as a construct. Because only when an expected or desired value and its corresponding (perceived) likelihood are attached to an objective will it become real. In essence, this is the risk one is willing to take. It is the positive value attributed to that objective.
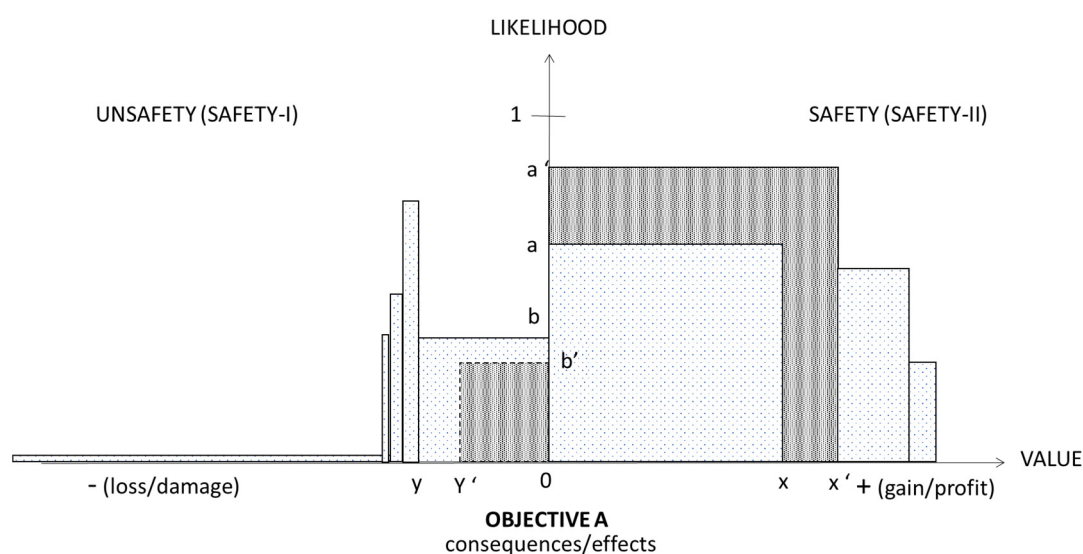

Figure 3. Risk, unsafety and safety [46].

According to one of the most used risk management standards, risk is the effect of uncertainty on objectives (ISO 31000). It is stated that this effect can be positive, negative or both. In practice, when considered over a period of time, objectives mostly undergo both kinds of effect. When an objective is clear, it is possible in some way to determine an associated value, and an according likelihood of achieving this value, to it. In Figure 3, these expected or desired values are related to "x" and "a". At this point, there's no 100% certainty regarding achieving this value, so risk is involved. When taking the risk of pursuing this objective, which means taking action aiming to achieve the objective, the expected value (level of safety or gain) "x" related to objective A is supposed to have a likelihood of "a" to be achieved. At the same time, a negative value (level of unsafety or loss) of "y" is also directly to be expected with a likelihood of "b" for the risks run, i.e., the negative effects that can directly prevent the achievement of the objective and their associated loss. At a first glance, the decision to pursue and maintain the objective will then be determined by the relative sizes of the two surfaces "ax" and "by". When this balance is positive, it is considered worthwhile and safe to proceed and act. Most of the time, this balance and action concerning objectives is based on the individual mental models (values and convictions) of the decision-makers involved when pursuing objectives individually, in organisations or society as a whole.

Traditionally, risk and safety practitioners, as well as scientists, have looked at the left-hand side of the graph (Safety-I), where it is important to reduce the likelihood b to b' and reduce the consequences y to y'. With a better understanding of the system(s) that can influence the safety of objective A, traditional risk and safety would also be interested in finding as much as possible about all the other possible negative effects on the objective A, trying to eliminate and/or reduce these effects and their associated likelihoods. Discovering the various risk sources that, alone or in combination, can create such effects is then paramount. Most of all, the high impact, low probability events (HILP, or Type II) have stimulated safety scientists to search for the holy grail of finding the systems and structures that allow preventing these catastrophes from happening. However, very limited research is performed concerning the mental models that create and govern such systems.

With new insights in managing risk and safety, expressed in modern standards and theories such as ISO 31000 and Safety-II, the right-hand side of the graph also becomes involved. Managing risk and increasing safety then also means to take those measures that increase the value x to x' and likelihood a to a', while also decreasing the value y to y' and the corresponding likelihood b to b'. As such, these alterations can be seen as primary effects of uncertainty on objectives.

Systems always act in larger overarching systems and are comprised of smaller sub-systems, each of them with a specific purpose and outcome. Even the smallest systems can fail and produce unwanted outcomes, creating adverse effects for the larger encompassing systems. This is also an effect of uncertainty on objectives.

4. Discussion

4.1. Taking Risk and Running Risk

As the history of safety science shows [45], in one's mind, connections are easily made between both sides of the risk and safety graph. Scholars pursue ways to improve safety but they have always been concerned with the reduction of unsafety. This happens because both sides of risk (negative and positive) and safety (Safety-I and Safety-II) exist from the moment a certain value or importance is attached to an objective. When the total value of that objective is perceived to be uncertain, it is considered as taking risk when pursuing this objective. Taking risk, therefore, actually signifies trying to create or obtain the expected or desired value attributed to that objective.

However, in taking action to pursue an objective—and not taking action is also to be understood as taking action—the objective and its value will also be linked to different risk sources that can generate positive and negative effects, adding or subtracting value of the objective. These are the risks one runs. Taking risk is active, deliberate and in the pursuit of value, but running risks is passive, un-deliberate

and with the prospect of loss. Because certain people are afraid of these losses, they immediately think of the possible losses when thinking of taking risk, they focus on the loss instead of considering the possible gains when the outcome of an endeavour is uncertain; they have a so-called risk-averse attitude. Other people, with a focus on gain, have a rather risk-seeking attitude, even if most people have a so-called risk-tolerant attitude. These attitudes result from the mental models (meanings, assumptions, conclusions and beliefs) generated through the ladders of inference one holds.

4.2. Proactively Generating Safety

Risk, safety and performance are related to objectives (what is valuable and valued by someone) at a given time. For risk, this lies in the future, for safety it is about a present situation and performance is about what has already happened and lies in the past [45].

Taking risks is all about pursuing uncertain objectives. This can be done by expressing an ambition and determining the level of risk one expects to take (corresponding to surface x.a in Figure 3), by articulating an expected level of value (x) and an expected level of likelihood (a) regarding achieving or maintaining that objective. In that way, a basic level of safety at the component level is established. Running a risk (surface b.y in Figure 3) is the cost or losses likely to be incurred when pursuing an objective, which can be considered a basic level of unsafety [45]. The establishment of these basic levels of safety and unsafety provides a starting point allowing to identify risk sources and their effects—both positive and negative—related to this specific objective. Combined with the associated likelihoods of those possible effects, a more comprehensive level of safety (creation of value) and unsafety (loss of value) can be determined, reflected by the additional surfaces on the graph in Figure 3.

Managing the level of safety can then typically consist of taking measures that improve the basic level of safety by increasing the performance regarding the concerned objective, by adding value and increasing likelihood on the safety side or reducing loss and likelihood on the unsafety side of the graph. This can happen by increasing the actual value ($x \rightarrow x'$) and/or likelihood of realisation ($a \rightarrow a'$) of the positive effects on objectives, which will then increase the level of safety. One can also take action in reducing the value ($y \rightarrow y'$) and likelihood ($b \rightarrow b'$) of the negative effects related to the concerned objective, decreasing the level of unsafety.

However, even for a single objective many risk sources and their associated effects can affect the end result of an actual case, which can be seen as the actual performance of the system in pursuing, achieving or safeguarding its objective. Each action taken (and no action is also considered a type of action) will create a new basic level of risk, safety and performance, adding, subtracting and changing the different risk sources and their associated effects influencing both sides of the graph. These actions (as a result of generated mental models) create a new level of safety with a new balance between the positive and negative effects of uncertainty on the concerned objective. When the actual performance of the system meets the desired level of performance, an objective can be considered as being safe or safeguarded. However, when the desired safety and unsafety levels are not reached, the objective is to be considered as a failed objective.

4.3. The Illusion of Cause-Effect Prediction and Accident Prevention

In safety science, many scholars are looking for ways to predict and prevent accidents. Most certainly this is the case for the high impact and low probability (HILP) events that can have a huge impact on society. It is the consequence of Safety-I thinking, where negative effects of uncertainty on objectives are seen to be related to cause-effect relations, looking for fixed systemic connections between risk sources and consequences. It is a line of thinking (mental model) aiming at the capability for accidents to be forecasted and to prevent them from happening. It is part of the "barrier" and "scenario thinking" methods that are still popular ways to deal with safety in organisations today. While it is possible to think of possible and plausible cause-effect relationships in simple and complicated systems, this is no longer possible for the complex socio-technical systems in today's society. Complex and chaotic contexts are unordered—there is no immediately apparent

relationship between cause and effect, and the way forward is determined based on emerging patterns. Furthermore, a complex system has the following characteristics [50–52]:

- It involves large numbers of interacting elements;
- The interactions are nonlinear, and minor changes can produce disproportionately major consequences;
- The system is dynamic, the whole is greater than the sum of its parts, and solutions cannot be imposed; rather, they arise from the circumstances. This is frequently referred to as emergence;
- The system has a history, and the past is integrated with the present; the elements evolve with one another and with the environment, and evolution is irreversible.

Though a complex system may, in retrospect, appear to be ordered and predictable, hindsight does not lead to foresight because the external conditions and systems constantly change.

Unlike in ordered systems (where the system constrains the agents), or chaotic systems (where there are no constraints), in a complex system the agents and the system constrain one another, especially over time. This means that one cannot forecast or predict what will happen [51].

There are cause-and-effect relationships between the agents in complex systems, but both the number of agents and the number of relationships defy categorization or analytic techniques. Therefore, emergent patterns can be perceived but not predicted [52].

Consequently, as indicated earlier, safety is an emergent property of complex systems and needs to be achieved over and over again. It is not a static situation but a set of dynamic circumstances, where the objectives of sub-systems need to be aligned and ever again achieved to create the desired level of safety. Therefore, the only way to envisage the future correctly is to shape it in the way it is looked-for, by adapting and aligning mental models through dialogue and learning on an individual, corporate and societal level, creating actions that generate safety and safeguard objectives on a continuing basis. It is the only way to achieve and maintain an adequate and sustainable level of safety in complex socio-technical systems.

4.4. Total Respect Management

A modern systems thinking perspective on safety that fits with this representation and a model to generate and achieve sustainable safety in a proactive way that is proposed is what is called Total Respect Management (TR^{3}M). Respect, in the way this word is intended to be used for this model, is an expression originally derived from the Latin word "respectus". In its turn, "respectus" comes from the verb "respicere", which means "to look again", "to look back at", "to regard", "to review", or "to consider someone or something". In other words, the original meaning of the word "respect" holds the connotation of giving someone or something one's appropriate and dedicated attention in order to have a better view on the matter or give it some thought, particularly to come to a better understanding. When used in the context of Total Respect Management, this is exactly how the word "respect" needs to be understood [50,52]. As much as possible, the corresponding sub-systems of socio-technical systems and their objectives need to be known and understood by giving them the appropriate level of attention. For TR^{3}M, this means respecting people (individual objectives and mental models) by developing leadership, respecting profit (corporate objectives and corporate mental models) by managing risk and respecting the planet (societal objectives and societal mental models), aiming for excellence and sustainability. As such, TR^{3}M aims at promoting a systems thinking approach, developing leadership to align mental models, implementing risk management according to ISO 31000 and a continuous improvement of inadequate sub-systems. It builds on the generation of trust and accountability in order to proactively influence safety and performance, as trust is an important precondition to facilitate the process of proactively changing and aligning mental models in organisations [53].

4.5. The Swiss Cheese Metaphor Revisited

The reason for this "respect" for systems and their sub-systems is the conviction that there is no common structure to (all) "accidents" (See also Section 4.3). They cannot be predicted when complexity is such that not all cause effect relations can be understood and managed anymore. Or, at least, it is the conviction that such a complex structure is beyond the actual means of comprehension of mankind. Alternatively, at best, for less complex systems, this notion exists at a very general level, not very well suited for an unquestionable prediction of future accidents. This is also a way in which one has to look at the Swiss Cheese metaphor. When picturing some Swiss cheese, people imagine a block of cheese with holes in it. In this metaphor, in the way one can see it, the whole block of cheese is a reflection of reality and of the performance of a socio-technical system and its sub-systems, created and governed by its associated mental models. The cheese itself can be understood as everything that goes well. It relates to the objectives that have been achieved and which are safeguarded. Hence, the cheese stands for the objectives where the value achieved and safeguarded in the Safety-II domain of the graph in Figure 3 is dominant. The cheese therefore represents the achieved and safeguarded objectives. On the other hand, the holes in the cheese are the sub-systems of which the objectives have not been achieved or safeguarded, and this mainly concerns those objectives where the loss of value, represented by the unsafety (Safety-I) side of the graph in Figure 3, prevails. These are the objectives that have failed. As such, these are the different reasons which contribute to phenomena going drastically wrong when they become connected (thus, the holes in the cheese represent Safety-I). The model's hypothesis is that one can never know with complete certainty which sub-systems will fail at a given time or why and how these failed objectives will become connected in a cause-effect relationship at a given time, thus causing a catastrophic failure of the whole overarching system.

Each hole in the cheese is to be regarded as a "failed" objective and it therefore creates unsafety. To understand how to view this model, it is important to remember that reality is dynamic and that conditions change from one moment to another. It is supposed that holes grow and shrink and this will happen at random. This process means that this Swiss cheese is dynamic. One has to picture a Swiss cheese of which the holes constantly change in positions and dimensions and this does not happen in a predictable way. In his book *Managing the risks of organisational accidents* [35], Reason states:

> Although the model shows the defensive layers and their associated 'holes' as being fixed and static, in reality they are in constant flux. The Swiss cheese metaphor is best presented by a moving picture, with each defensive layer coming in and out of the frame according to local conditions. [...] Similarly, the holes within each layer could be seen as shifting around, coming and going, shrinking and expanding in response to operator actions and local demands."

Although the idea of layers of protection, by putting barriers or fences around the holes, is useful to a certain extent, it only works for Safety-I. Therefore, the TR^{3}M model also focusses on successful performance as an element of safety. The aim of performance is to achieve objectives and maintain, as much as reasonably possible, the objectives safeguarded. As such, performance stands for the whole cheese, or the whole concerned socio-technical system and the aim is excellence (safe performance), where the holes are as little and few as possible and value is maximised.

The way TR^{3}M approaches the Swiss Cheese metaphor is by stating that each of the holes is a subsystem of which the specific objectives have failed (latent conditions) and that failures of achieving or maintaining these objectives can be seen as accidents on their own. One could say it is about slicing the cheese in little pieces and consider each piece as an entire block of cheese, whereby a hole through the cheese is considered an accident. These "accidents", or failed/failing objectives (unsafe performance), result from systems (risk sources) shaped and maintained by the mental models existing in, or surrounding, the concerned socio-technical system. In fact, it is just the level of importance and number of objectives involved that differentiate the catastrophic "accidents" from these incidents and their corresponding, apparently "less important", holes. In general, it is only when the holes

represent important objectives they are seen as real accidents and considered worth investigating. However, each one of the holes (objectives in a broad sense) is meaningful and needs one's respect. Hence the name Total Respect Management [54].

4.6. Quality of Perception

The key to achieving safety proactively depends on one's quality of perception, where the quality of perception should be understood as the level of deviation (gap) between reality and the perception (mental model) of that reality. When this gap is small, there is a high quality of perception, but when ignorance makes the gap bigger, there is a lower quality of perception. It means that increasing the awareness, knowledge and understanding of objectives involved in a socio-technical system, its subsystems and encompassing systems (context) is crucial in generating safety. How are the objectives possibly impacted by events or conditions? What are the consequences of failure or success of these objectives? How well are these objectives aligned? These are essential questions to be asked. When the quality of perception is high, this will have its consequences on the mental models present in the system, as these mental models will be better aligned with reality, allowing for better decisions, increased safe performance and less unsafe performance.

4.7. A Broad Perspective on Safety

In a traditional context, risk and risk management are more concerned with finance and profit, where safety is more worried about health and injury and security is troubled by possessions, opposing interest, power and human life. However, many different categories of objectives exist and every one of them has to be taken into account to determine and generate safety proactively. An aircraft crash or a similar important accident, in the processing industry for instance, has an impact on many important objectives. These objectives range from very specific and individual goals to very general and societal aspirations. They also cover different dimensions, for instance, individual objectives regarding life and health, or financial objectives of the organisation regarding profit and continuity of operations, or objectives regarding our environment on a societal level, and so on. The more objectives that are impacted by the negative effects of uncertainty and the more valuable they are, the more they will be perceived as an accident. One only has to think of the countless individual, organisational and societal objectives which were impacted by the Fukushima disaster. It is the range, the number and the importance of those goals that caught everybody's attention and which will remain in the memory of mankind. Nevertheless, this accident only happened because of some latent conditions that existed in the nuclear plant and in the organisation operating the plant; the protective measures, equipment and procedures in place did not achieve the objectives they were intended or designed for. These were the existing holes in the system nobody was aware of, or no one bothered to take action against, due to the governing mental models in this socio-technical system. They did not seem important until they were joined by the circumstances and the failure of other objectives, that proved to be crucial at that moment [54].

Risk, safety or security, in this sense, are nothing more than possible conditions of one's objectives and the performance regarding those objectives at a given (future) time, representing a reflection of a possible (future) reality. The better this future reality can be imagined, the more it can be shaped to the desires and needs of the beholders by taking the right decisions at the appropriate time. This can only come about when different viewpoints can be paired and aligned with the common objectives to create common and aligned mental models that fit with reality and generate systems producing safety proactively.

4.8. Some Examples of Mental Models Influencing Safety, Security and Performance

A first example of a mental model that positively impacts safety and performance at a very general level is a concept (or an idea, conviction, mindset, and so on) that is at the core of aviation safety. It generates a high level of detailed reporting of incidents and is responsible for huge efforts made

in accident investigation. It is the very simple belief that it is crucial to learn from the mistakes of others because one will never live long enough to make all mistakes oneself. This mental model works, since mistakes in aviation often have a deadly outcome impacting many important objectives. It is only because of these huge efforts in reporting and analysing incidents and accidents that knowledge and understanding increase in such ways that it allows to reach the very high standards of safety and performance currently achieved in the aviation industry.

Another idea (or belief, concept, conviction, and so on) is the mindset focused on always holding the stair railing when ascending or descending stairways or ladders. In some organisations, often in the petrochemical industry, this mental model is so ingrained that it has become an attitude and even a habit. It is impossible to know how many accidents have been prevented by this very simple organisational, yet individual, mental model. However, thousands of people get injured every year, and some killed, by either careless behaviour, or by faulty architectural features that contribute to these accidents [55].

Mental models play a role in causal relationships, as can be found in investigations of notorious accidents. In their paper "The accident of m/v Herald of Free Enterprise", Goulielmos and Goulielmos state the following:

> "This was for the ship to do the crossing at the least possible time reducing time at the port as a rule. Moreover, two written instructions were issued in 1986 showing a pressure for the earlier sailing of the ship from Zeebrugge (as ships delayed in Dover as a rule), by 15 min, management asking for exerting pressure, especially on Chief Mate, for this end. The company had passed the culture of the urgency of sailing without this being supported by any marketing survey and as happened at the expense of safety."

Knowing that the ship sailed off without verifying the closure of the bow doors, which directly caused the ship to sink, it can be argued that a different mental model, not putting the emphasis on time pressure, but on safety and following procedures instead, could have created a different system, with a different outcome [56].

Similar flaws can be discovered in relation to the Fukushina nuclear disaster. Hasegawa [57] states, amongst others, the following reasons for this tragedy:

> "A series of "unexpected situations" which the executive members of TEPCO have sought to explain should be re-termed a series of "underestimations" by the company and the government. First, the height of the tsunami was underestimated by TEPCO and the Japanese Nuclear Safety Commission (JNSC). Even though some researchers gave scientific warnings of a 15.7-m tsunami in May 2008, both TEPCO and JNSC neglected this warning. So, the plant remained designed to withstand a 5.7-m tsunami according to the 2002 estimation. On 11 March 2011, a tsunami with a height of 14–15 m at its maximum hit and flowed beyond the 10-m elevation of basement, flooding to a depth of over 5 m. Second, Japanese power companies and government agencies had not expected any possibility of lengthy "station blackouts," the total loss of AC power of the station, especially caused by a large-scale natural disaster, earthquake, tsunami or flood. But the blackout occurred with the result that all cooling functions were lost. Authorities expected that in the case of a station blackout, external power would be recovered within 30 min. This expectation was formed without sufficient basis thereby dismissing the need to prepare for the possibility of a lengthy period of no AC power."

It is clear that "unexpected situations" is a sign of possibly flawed mental models of TEPCO and the government agencies, indicating at least a low quality of perception, not taking into account scientific warnings. Or even worse, it could also be an indication of a deliberate underestimation of possible events, with a priority for profitability as a governing mental model, leading to the latent failed objectives of protection against flooding and the availability of AC power. What would have

happened if other governing mental models had been in place? For instance, the mental model that it is wise to listen to scientists and heed their advice.

When mental models are different in different groups of people (societies), they can generate conflicting objectives. When this happens, security issues will arise. As an example, one only has to think of the governing mental models in a democratic open society, where equal rights, gender equality and women's empowerment are very important aspects of life that possibly conflict with differing viewpoints ruling different societies. When these conflicts are very outspoken, terrorist action can be expected, as has been the case with Al Qaeda or the Islamic State of Iraq and Syria (ISIS). Likewise, there are the different mental models that go along with political systems, which can also cause deliberate the harmful action of one group of people against another one, not adhering to the same values and belief systems (mental models). Understanding these governing mental models and their associated objectives provides information and directions on how to deal with these related security issues.

An example of the influence of one simple mental model on the performance of a complex socio-technical system can be found in the book "The Power of Habit". In this book, Duhigg tells the story of the Aluminium Company of America (Alcoa) and how in 1987, at a time when Alcoa was struggling, a new CEO was appointed. Paul O'Neal, the new CEO, drastically changed Alcoa's performance by focussing on one simple idea regarding safety. This mental model was "zero injuries". His ambition was to make Alcoa the safest company in America by reducing Alcoa's injury rate. At that time, this was something unheard of in the corporate world and at the beginning this approach was met with a lot of scepticism. But within a year, Alcoa's profits would hit a record high and by the time O'Neal retired in 2000, the company's annual net income was five times larger than before he arrived [58].

4.9. Mental Models and a Sustainable World

Safety is closely linked with sustainability. In a certain way they are almost synonyms. Safety is about achieving, maintaining and protecting what is valuable and important. How can something be safe when it is not sustainable or how can something be sustainable when it is not safe? Societal mental models governing sustainability are to be found in domains such as climate change, corporate social responsibility or world peace. Examples can, for instance, be found in the United Nations Global Compact principles (specific mental models) for corporate sustainability. One can read on the UN website: "By incorporating the Ten Principles of the UN Global Compact into strategies, policies and procedures, and establishing a culture of integrity, companies are not only upholding their basic responsibilities to people and planet, but also setting the stage for long-term success" [59].

5. Challenges and Further Research

While the idea of working with mental models seems easy and simple, it is not. As indicated earlier, mental models result from individual ladders of inference and are different from one person to another. These individual mental models govern the behaviour of people and, as a consequence, also the behaviour of organisations and ultimately society as a whole. Alignment of individual mental models in organisations and society is therefore an important challenge that needs to be addressed when aiming to implement organisational or societal mental models. This aspect of working with mental models is something to address in another paper and certainly is a possible way forward in the search for proactive and sustainable safety.

6. Conclusions

In our ever more complex and connected world, the safety of systems depends on the interactions and performance of the much smaller sub-systems. A proactive way to reach safety of systems is therefore to focus on the performance of the sub-systems at ever deeper levels of detail within the concerned system. It would therefore be interesting to study how mental models and these smaller

subsystems relate and how they determine risk, safety and performance in the concerned socio-technical systems. Discovering appropriate empowering mental models, as well as relevant harmful mental models, would then allow to work with mental models that generate safety and eliminate unwanted events. This should be made possible at an individual, corporate and societal level, to create an environment where people can be safe in any aspect of life.

To be able to do so, it is of the utmost importance to organise and structure dialogue to create and disseminate the corresponding mental models that generate and allow for this dedicated focus and attention to detail (the role of leadership). At the same time, it is important to discover how the sub-systems interact and create value, or produce unwanted events that can be avoided (the role of risk management). As such, it is necessary to simultaneously consider risk, safety (and security) and performance of even the smallest sub-systems and aim to reduce the number of failed objectives by continuous improvement, creating and maintaining safety in a sustainable way (the role of excellence).

References

1. Meadows, D.H. *Thinking in Systems: A Primer*; Chelsea Green Publishing: White River Junction, VT, USA, 2008.
2. Arnold, R.D.; Wade, J.P. A definition of systems thinking: A systems approach. *Procedia Comput. Sci.* **2015**, *44*, 669–678. [CrossRef]
3. Richmond, B. System Dynamics/Systems Thinking: Let's Just Get on with it. In Proceedings of the International Systems Dynamics Conference, Sterling, Scotland, 11–15 July 1994.
4. Kim, D.H. *Introduction to Systems Thinking*; Pegasus Communications: Waltham, MA, USA, 1999; Volume 16.
5. Monat, J.; Gannon, T. Applying Systems Thinking to Engineering and Design. *Systems* **2018**, *6*, 34. [CrossRef]
6. Monat, J.P.; Gannon, T.F. What is systems thinking? A review of selected literature plus recommendations. *Am. J. Syst. Sci.* **2015**, *4*, 11–26.
7. von Bertalanffy, L. *General System Theory: Foundations, Development, Applications*; Braziller: New York, NY, USA, 1969.
8. Aronson, D. Overview of Systems Thinking. 1996. Available online: http://resources21.org/cl/files/project264_5674/OverviewSTarticle.pdf (accessed on 19 November 2018).
9. Cavana, R.; Maani, K. A Methodological Framework for Integrating Systems Thinking and System Dynamics. In Proceedings of the 18th International Conference of the System Dynamics Society, Bergen, Norway, 6–10 August 2000; pp. 6–10.
10. Bosch, O.; Maani, K.; Smith, C. Systems thinking-Language of complexity for scientists and managers. In *Improving the Triple Bottom Line Returns from Small-Scale Forestry*; The University of Queensland: Brisbane, Australia, 2007; Volume 1, pp. 57–66.
11. Senge, P.M.; Kleiner, A.; Roberts, C.; Ross, R.B.; Smith, B.J. *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organization*; Currency, Double Day, Random House, Inc.: New York, NY, USA, 1994.
12. Buschman, T.J.; Siegel, M.; Roy, J.E.; Miller, E.K. Neural substrates of cognitive capacity limitations. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11252–11255. [CrossRef] [PubMed]
13. Argyris, C. The executive mind and double-loop learning. *Organ. Dyn.* **1982**, *11*, 5–22.
14. Argyris, C. Interventions for improving leadership effectiveness. *J. Manag. Dev.* **1985**, *4*, 30–50. [CrossRef]

15. Senge, P. *The Fifth Discipline: The Art and Practice of Organizational Learning*; Currency, Double day, Random House, Inc.: New York, NY, USA, 1990.

16. Jones, N.A.; Ross, H.; Lynam, T.; Perez, P.; Leitch, A. Mental models: An interdisciplinary synthesis of theory and methods. *Ecol. Soc.* **2011**, *16*, 46. [CrossRef]

17. Doyle, J.K.; Ford, D.N. Mental models concepts for system dynamics research. *Syst. Dyn. Rev. J. Syst. Dyn. Soc.* **1998**, *14*, 3–29. [CrossRef]

18. Bryan, B.; Goodman, M.; Schaveling, J. *Systeemdenken*; Academic Service: Cambridge, MA, USA, 2006.

19. Testa, M.R.; Sipe, L.J. A systems approach to service quality: Tools for hospitality leaders. *Cornell Hotel Restaur. Adm. Q.* **2006**, *47*, 36–48. [CrossRef]

20. Stroh, D.P. Leveraging grantmaking: Understanding the dynamics of complex social systems. *Found. Rev.* **2009**, *1*, 9. [CrossRef]

21. Farmer, E. The method of grouping by differential tests in relation to accident proneness. *Ind. Fatigue Res. Board Annu. Rep.* **1925**, 43–45.

22. Heinrich, H.W. *Industrial Accident Prevention. A Scientific Approach*, 1st ed.; McGraw-Hill Book Company: London, UK, 1931.

23. Rasmussen, J. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Trans. Syst. Man Cybern.* **1983**, 257–266. [CrossRef]

24. Heinrich, H.W. *Industrial Accident Prevention. A Scientific Approach*, 2nd ed.; McGraw-Hill Book Company: London, UK, 1941.

25. Perrow, C. *Normal Accidents: Living with High Risk Systems*; Princeton University Press: Princeton, NJ, USA, 1984.

26. Rasmussen, J. Risk Management and the Concept of Human Error. *Joho Chishiki Gakkaishi* **1995**, *5*, 39–70. [CrossRef]

27. Rasmussen, J. Risk management in a dynamic society: A modelling problem. *Saf. Sci.* **1997**, *27*, 183–213. [CrossRef]

28. Van Breda, A.D. *Resilience Theory: A Literature Review*; South African Military Health Service: Pretoria, South Africa, 2001.

29. Norris, F.H.; Stevens, S.P.; Pfefferbaum, B.; Wyche, K.F.; Pfefferbaum, R.L. Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *Am. J. Community Psychol.* **2008**, *41*, 127–150. [CrossRef]

30. La Porte, T.R. High reliability organizations: Unlikely, demanding and at risk. *J. Contingencies Crisis Manag.* **1996**, *4*, 60–71. [CrossRef]

31. Marais, K.; Dulac, N.; Leveson, N. Beyond Normal Accidents and High Reliability Organizations: The Need for an Alternative Approach to Safety in Complex Systems. In *Engineering Systems Division Symposium*; MIT: Cambridge, MA, USA, 2004; pp. 1–16.

32. Leveson, N. *Engineering a Safer World: Systems Thinking Applied to Safety*; MIT Press: Cambridge, MA, USA, 2011.

33. Hollnagel, E. *FRAM, the Functional Resonance Analysis Method: Modelling Complex Socio-Technical Systems*; Ashgate Publishing, Ltd.: Farnham, Surrey, 2012.

34. Accou, B.; Reniers, G. Developing a method to improve safety management systems based on accident investigations: The SAfety FRactal ANalysis. *Saf. Sci.* **2019**, *115*, 285–293. [CrossRef]

35. Reason, J. *Organizational Accidents: The Management of Human and Organizational Factors in Hazardous Technologies*; Cambridge University Press: Cambridge, UK, 1997.

36. Reason, J. *Managing the Risks of Organizational Accidents*; Routledge: Abingdon, UK, 1997.

37. Hawkins, F.H. *Human Factors in Flight*; Gower Publishing Company: Brookfield, VT, USA, 1987.

38. Shappell, S.A.; Wiegmann, D.A. *Reshaping the Way We Look at General Aviation Accidents Using the Human Factors Analysis and Classification System*; Ashgate Publishing Company: Aldershot, UK, 2003.

39. Dekker, S.W. Just culture: Who gets to draw the line? *Cogn. Technol. Work* **2008**, *11*, 177–185. [CrossRef]

40. Dekker, S. *Just Culture: Restoring Trust and Accountability in Your Organization*; CRC Press, Taylor & Francis Group: Abingdon, UK, 2017.

41. Conchie, S.M.; Donald, I.J.; Taylor, P.J. Trust: Missing piece (s) in the safety puzzle. *Risk Anal.* **2006**, *26*, 1097–1104. [CrossRef]

42. Blokland, P.; Reniers, G. Total Respect Management (TR³M): A Systemic Management Approach in Aligning Organisations towards Performance, Safety and CSR. *Environ. Manag. Sustain. Dev.* **2015**, *4*, 1–22. [CrossRef]

43. Weick, K.E. Organizing for transient reliability: The production of dynamic non-events. *J. Contingencies Crisis Manag.* **2011**, *19*, 21–27. [CrossRef]

44. Hollnagel, E. *Safety-I and Safety-II: The Past and Future of Safety Management*; Ashgate Publishing, Ltd.: Farnham, UK, 2014.

45. Blokland, P.; Reniers, G. An Ontological and Semantic Foundation for Safety and Security Science. *Sustainability* **2019**, *11*, 6024. [CrossRef]

46. International Standards Organisation (ISO) 31000: 2009. *Risk Management–Principles and Guidelines*; International Organization for Standardization: Geneva, Switzerland, 2009.

47. International Standards Organisation (ISO) 31000: 2018-02. *Risk Management–Guidelines*; International Organization for Standardization: Geneva, Switzerland, 2018.

48. International Standards Organisation (ISO) Guide 73: 2009. *Risk Management—Vocabulary*; International Organization for Standardization: Geneva, Switzerland, 2009.

49. Slovic, P. The risk game. *Reliab. Eng. Syst. Saf.* **1998**, *59*, 73–77. [CrossRef]

50. Blokland, P.; Reniers, G. *Total Respect Management–Excellent Leidinggeven voor de Toekomst*; Lannoo: Tielt, Belgium, 2013.

51. Snowden, D.J.; Boone, M.E. A leader's framework for decision making. *Harv. Bus. Rev.* **2007**, *85*, 68.

52. Kurtz, C.F.; Snowden, D.J. The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Syst. J.* **2003**, *42*, 462–483. [CrossRef]

53. Sharkie, R. Trust in Leadership is Vital for Employee Performance. *Manag. Res. News* **2009**, *32*, 491–498. [CrossRef]

54. Blokland, P.; Reniers, G. *Safety and Performance: Total Respect Management (TR3M): A Novel Approach to Achieve Safety and Performance Pro-Actively in Any Organisation*; Nova Science Publishers: New York, NY, USA, 2017.

55. Atlas, R. What Is the Role of Design and Architecture in Slip, Trip, and Fall Accidents? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*; SAGE Publications: Los Angeles, CA, USA, 2019; Volume 63, pp. 531–536.

56. Goulielmos, A.M.; Goulielmos, M.A. The accident of m/v Herald of Free Enterprise. *Disaster Prev. Manag. Int. J.* **2005**, *14*, 479–492. [CrossRef]

57. Hasegawa, K. Facing nuclear risks: Lessons from the Fukushima nuclear disaster. *Int. J. Jpn. Sociol.* **2012**, *21*, 84–91. [CrossRef]

58. Duhigg, C. *The Power of Habit: Why We Do What We Do and How to Change*; Random House: New York, NY, USA, 2013.

59. United Nations Global Compact. Available online: https://www.unglobalcompact.org/what-is-gc/mission/principles (accessed on 30 April 2020).