Chapter 3
A Conceptual Foundation for Assessing and Managing Risk, Surprises and Black Swans

Terje Aven

Abstract This chapter presents and discusses some recent advances in the risk field, linked to the conceptualisation of risk and specifically addressing unforeseen events, surprises and so-called black swans. It shows how the traditional probability-based perspectives on risk are extended to broader ways of thinking about risk, which give due attention to the uncertainties and also draw on ideas from *inter alia* the quality discourse and organisational learning (collective mindfulness and its five principles: preoccupation with failure, reluctance to simplify, sensitivity to operations, commitment to resilience and deference to expertise). The main aims of the chapter are to point to this thinking and provide some reflections on how to use it for further developing the risk assessment and risk management fields.

Keywords Uncertainty · Risk perspectives · Risk concept · Risk assessment · Surprises · Black swans

3.1 Introduction

In recent years there has been an increased focus on uncertainties in relation to risk assessment and risk management. The traditional probabilistic approach to risk assessment and management is criticised for its narrowness in the way it looks at risk and how it copes with uncertainties (see e.g. [15, 52]). In response to the critique, several alternatives have been suggested. These can be grouped as follows:

1. Replacing the probabilistic approach with other quantitative approaches, for example based on interval probabilities (typically supported by possibility theory or evidence theory) [8].

T. Aven (✉)
University of Stavanger, Stavanger, Norway
e-mail: terje.aven@uis.no

2. Balancing alternative approaches, in particular the probabilistic approach and approaches that are effective in meeting hazards/threats, surprises and the unforeseen. For short, we refer to the latter approaches as "robust approaches".
3. Rejecting the probabilistic approach to risk assessment and risk management, and instead relying on robust approaches.

Robust approaches cover cautionary measures such as designing for flexibility (meaning that it is possible to utilise a new situation and adapt to changes); implementing safety barriers; improving the performance of barriers by using redundancy, maintenance, testing etc.; and applying quality control/assurance. It also covers concepts like resilience engineering, which is concerned with finding ways to enhance the ability of organisations to be resilient in the sense that they recognise, adapt to and absorb variations, changes, disturbances, disruptions and surprises [35]. We may also include the concept of antifragility [50]. According to Taleb, the antonym of fragile is not robustness and resilience, but "please mishandle" or "please handle carelessly"—to quote an illustration from Taleb when referring to sending a package full of glasses by post. The antifragile is seen as a blueprint for living in a "black swan world", the key being to love randomness, variation and uncertainty to some degree, and thus also errors. As our bodies and minds need stressors to be in top shape and improve, so do other activities and systems [9, 12].

In practice, alternative (2) normally applies. If we look at high risk industries such as nuclear and oil & gas, we find a mixture of the probabilistic and the robust approaches. It is acknowledged that the probabilistic approaches have limitations in managing risk, surprises and the unforeseen, and need to be supplemented by robust approaches. For societal safety and security contexts, we see this duality as even stronger, for example in relation to terrorism risk. Here probabilistic risk assessments are hardly used. The information provided by assigned attack probabilities is small in most cases [7, 21].

To confront risk and uncertainties, robust thinking is needed, to a varying degree depending on the situation. This is also true when alternative quantitative approaches (approach 1) are adopted. Using interval probabilities in place of specific probability numbers can allow for more balanced judgements reflecting what is known and what is not, but, as thoroughly discussed in [8], the interval probability approach is not easily implemented in practice—there are many challenges—and most importantly, it cannot replace the need for robust approaches. No analytical quantitative approach, probabilistic or not, can make robust arrangements and measures superfluous, as they will always be subject to some limitations. They can provide useful decision support, but not prescribe what to do.

This acknowledgement implies the need for a way of thinking about risk that sees beyond the probabilistic perspective. In engineering applications, risk is commonly seen as expected values (probability multiplied with loss) or as the combination of probability and consequences (loss) [5, 19, 38]. However, such a perspective is too narrow to adequately capture the balanced setting of the approaches of category (2), and in recent years we have seen many attempts to conceptualise risk to meet this broader scene. One example is the new definition of risk adopted by the ISO (2009): "risk is the effect of uncertainties on objectives". A key aspect with this def-

inition of risk is that uncertainty replaces probability. This may seem to be a rather minor change, but it has important implications. It will be thoroughly discussed in the following when reviewing recent developments in the risk field, paying special attention to the perspectives that can serve the assessment and management approaches (2) mentioned above. The review, which will be partly based on [5], shows that there are many frameworks and approaches suitable for (2), which give further substance and precision to the understanding of risk compared to, for example, the ISO risk definition. From this review we will discuss in more detail how surprises and the unforeseen (black swans) are coped with in these frameworks and approaches. The main purpose of the chapter is to provide new insights on the link between these concepts—risk, surprises and the unforeseen—and to give guidance on how to adequately think in this respect, first when it comes to conceptualisation, then next when it comes to the way we should assess and manage risk, surprises and the unforeseen.

3.2   Risk Perspectives, Brief Review of Historical and Recent Development Trends

The origin of the term "risk" is disputed in the literature, see e.g. [1] and [5], but there seems to be broad agreement between researchers that De Moivre's 1711 definition is one of the first formal definitions of risk used in a risk analysis context [25]. De Moivre defines the risk of losing any sum to be the product of the sum adventured multiplied by the probability of the loss, i.e. risk is defined as the expected loss. This definition is still being used in many contexts despite the strong arguments against it; see summary of arguments in [5], Appendix A. About 200 years later, expected loss is replaced by probability: risk is the chance of damage or loss [33]. This perspective is further developed to reflect the magnitude of the losses and consequences, and at the beginning of the 1980s the so-called triplet definition of risk was the dominating understanding of the risk concept, at least in engineering contexts, which covers events/hazards, the consequences of these and associated probabilities [38].

In the social sciences this probability-based approach to risk is challenged and alternatives are proposed, one of the most popular being Rosa's definition expressing risk as a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain [30, 53].

Many researchers from the social sciences refer to risk but do not distinguish between risk per se and how risk is managed. One illustrating example is the German sociologist Beck, who states that "Risk may be defined as a systematic way of dealing with hazards and insecurities introduced by modernization itself" [17, p. 21]. It represents a way of looking at risk which is in conflict with most other perspectives on risk, and it is difficult to justify. To use the words of [22, p. 151], in their investigation of Beck's work on risk:

> It is hard to think of a less adequate definition: risk is not a way of dealing with things... Beck's definition would make it impossible to ask: How are we responding to this risk?, as the response and the risk would be the same thing. Secondly, risk should not be so defined that it applies only to 'modernization', for there were of course risks before industrial society.

We also see that the concept of risk and risk perception is mixed. The perception dimension includes personal feelings and affections (for example dread) about the possible events, the consequences of these events and about the uncertainties and probabilities, and even judgements about risk acceptability. According to cultural theory and constructivism, *risk is the same as risk perception* [29, 37, 65]. Beck [17, p. 55] states that "because risks are risks in knowledge, perceptions of risks and risk are not different things, but one and the same". Also this way of thinking clashes with most professional and scientific risk perspectives, as they seek to distinguish between what is risk and what are feelings and affections, and value statements about risk (for example, what is acceptable risk).

In the economic environment, risk has also been associated with uncertainty, for example by Hardy as early as 1923 [32]. He states that risk is uncertainty in regard to cost, loss or damage. We also see this perspective today, and typically the uncertainty is expressed by the variance or standard deviation. The case that the risk is objective uncertainty is also commonly referred to today. This perspective is to a large extent based on Frank Knight's work from 1921 and his conceptualisation, where risk is used for the case that an objective probability distribution can be obtained (and uncertainty otherwise) [42]. This nomenclature has strongly influenced the risk area, and in particular the economic risk field. Referring to risk only when we have objective probability distributions would mean that we exclude the risk concept from most real-life situations, and this terminology is therefore avoided by most risk analysts and researchers (see relevant references and discussion in [5]). Clearly, if one adopts the subjective or Bayesian perspective on probability, Knight's definition of risk becomes empty. No objective probabilities exist.

The risk = uncertainty perspective is typically based on the assumption that the expected value is the point of reference and that it is known or fixed. The uncertainty is seen in relation to, for example, a historical average value for similar investments. Risk captures the deviation and surprise dimension compared to this level. Without such a reference level, the "risk = uncertainty thesis" does not work. Uncertainty seen in isolation from the consequences and the severity of the consequences cannot be used as a general definition of risk. Large uncertainties need attention only if the potential outcomes are large/severe in some respect; see example in [5].

This leads us to risk perspectives highlighting both losses/consequences and uncertainties. For short we refer to these as the (C, U) perspectives, C denoting the consequences of the activity and U the uncertainties (what will C be) [5, 13]. The ISO (2009) definition, where risk is being defined by the effect of uncertainty on objectives, can be seen as a special case of this definition, with the consequences linked to objectives [5]. The Petroleum Safety Authority Norway has implemented a new definition of risk in line with the (C, U) definition: For the area of health, working environment and safety, risk means the consequences of the activities, with associated uncertainties.[1]

---

[1]Risk definition of the Norwegian Petroleum Safety Authority, accessed 7th October 2015, http://www.psa.no/risk-and-risk-management/category897.html.

The consequences C can be seen in relation to objectives or some other references, for example some planned numbers, and the focus is normally on negative, undesirable consequences. There is always at least one outcome that is considered as negative or undesirable. The consequences are with respect to something that humans value (including life, health, environment, economic assets).

The consequences also cover events, such as a gas leakage or other hazardous events. The knowledge dimension enters the scene when we try to measure or describe risk. This is performed by specifying the consequences C and using a description (measure) of uncertainty Q (which could be probability or any other measure—measure here interpreted in a wide sense). Specifying the consequences means to identify a set of events/quantities of interest C' that characterise the consequences C. An example of C' is the number of fatalities. Depending on which principles we lay down for specifying C and the choice of Q, we obtain different perspectives on how to describe/measure risk. As a general description of risk, we can write

Risk description $= (C', Q, K)$, where K is the background knowledge on which C' and the assignment Q are based.

Adopting this risk perspective, it is argued in [3] that safety is the antonym of risk, and that the analysts conclude on safety being high or low, and being safe, by reference to the risk description (C', Q, K).

We see that such a way of understanding and describing risk allows for all types of uncertainty representations, and it could consequently serve as a basis of a unified perspective on uncertainties in a risk assessment context. The most common uncertainty representation is probability, but, as mentioned in Sect. 3.1, there are also others (including those based on interval probability, possibility theory and evidence theory) [15]. In all forms of risk descriptions, the knowledge dimension (data, information, justified beliefs) and the strength of this knowledge need to be seen as an integral part.

The risk descriptions are often based on modelling, for example using probability models and frequentist probabilities. Probability models and frequentist probabilities do not exist in general: they are model concepts that are meaningful only in some situations of repeatability.

A basic way of categorising risk is related to the distinction between conceptualisations of risk that see risk as an objective property of the world and conceptualisations that are based on judgements and knowledge of a person (i.e. are epistemological). We remember the well-known phrases used by Immanuel Kant (1724–1804), "Das Ding an sich" and "Das Ding fur mich". Risk (and probability) can be viewed as both an "an sich" property of the world and a "fur mich" concept; see discussions in [4, 60].

In their famous 1981 paper [38], Kaplan and Garrick also refer to risk as qualitatively defined as "uncertainties + damage", which can be seen as a version of the (C, U) type of definitions. However, the authors did not develop a theory for this perspective as shown above for the $(C, U) – (C', Q, K)$ approach (this theory is based on early work by, for example, [2, 11]).

In this chapter, we are especially concerned about how the risk perspective supports the integration of concepts like unforeseen events, surprises and black swans.

In a recent work a new perspective on how to conceptualise, assess and manage risk and the unforeseen, covering surprises and black swans, is presented [12]. This work builds on the (C, U) risk perspective and draws on ideas from the quality discourse and the use of the concept of collective mindfulness as interpreted in studies of High Reliability Organisation (HRO). It also provides a suggestion for a classification of black swan type of events. We will look more closely into this perspective in the following section.

3.3   Risk, Surprises and Black Swans

An event is commonly considered a surprise when it occurs unexpectedly and also runs counter to accepted knowledge [31]. There exist, however, many other definitions, for example: a surprising event may be regarded as one whose occurrence was not anticipated, or which has been allocated such a low probability that the possibility of its occurrence was effectively discounted [40, p. 69]. The literature includes many taxonomies for classifying surprises. Examples include the dichotomies between known (imaginable) surprises and unknown surprises, between unanticipated surprises and anticipated surprises, and between the unintended, imaginable and anticipated [31, pp. 37–41]. An imaginable surprise occurs when the event type is known but its occurrence was considered highly unlikely. If we know that something is going to happen, but not when and in what form, it is referred to as an anticipated surprise. As noted by [31, p. 40], a surprise cannot be registered in any meaningful way without an "expectation" in some sense, to create a deviation.

In 2007, Nassim Taleb defined and popularised the concept of black swans in his book, *The Black Swan* [49]. Taleb defines a black swan to be an event with the following three attributes: firstly, it is an outlier, as it lies outside the realm of regular expectations, because nothing in the past can convincingly point to its possibility. Secondly, it carries an extreme impact. Thirdly, in spite of its outlier status, human nature makes us concoct explanations for its occurrence after the fact, making it explainable and predictable.

Several alternative definitions have been suggested, including the one recommended by [8]: a black swan is to be seen as a surprising extreme event relative to the present knowledge/beliefs. Hence the concept always has to be viewed in relation to whose knowledge/beliefs we are talking about, and at what time. Aven and Krohn [12] build on this definition and distinguish basically between three types of such events:

(a) Unknown unknowns: events that were completely unknown to the scientific environment (unknown events to us, unknown to others);
(b) Unknown knowns: events not on the list of known events from the perspective of those who carried out a risk analysis (or another stakeholder), but known to others (unknown events to us, known to others);
(c) Events on the list of known events in the risk analysis but judged to have a negligible probability of occurrence, and thus not believed to occur.
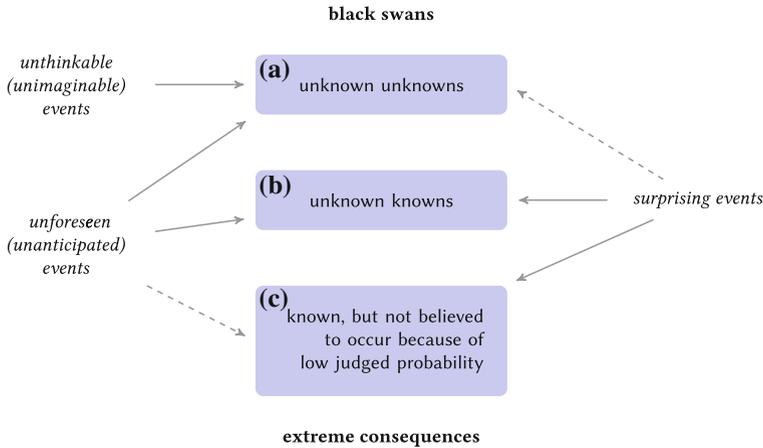
black swans

extreme consequences

Fig. 3.1 Schematic illustration of the concepts *black swan*, unknown unknowns, unforeseen events, surprising events and unthinkable events, based on the ideas presented by [12] and presented in [8]

The term "black swan" is used to express any of these types of events, tacitly assuming that it carries an extreme impact. See Fig. 3.1, which links terms such as black swans, surprising events and unforeseen events.

The first category (a) of black swan type of events is the extreme—the event is unthinkable—the type of event is unknown to the scientific community, for example a new type of virus. In activities about which there is considerable knowledge, such unknown unknowns are likely to be rarer than in cases of severe or deep uncertainties.

The second type of black swan (b) is events that are not captured by the relevant risk assessments, either because the analysts do not know them, or they have not made a sufficiently thorough consideration—the events are known to others. If the event then occurs, it was not foreseen. If a more thorough risk analysis had been conducted, some of these events could have been identified. The third category of black swans is events that occur despite the fact that the probability of occurrence is judged to be negligible. The events are known, but considered so unlikely that they are ignored; they are not believed to occur and cautionary measures are not implemented. Here is an example: an underwater volcano eruption occurs in the Atlantic Sea resulting in a tsunami affecting Norway. The event is on the list of risk sources and hazards but then removed as the probability is judged to be negligible. The occurrence will come as a surprise. The tsunami that damaged the Fukushima nuclear plant was similarly removed from the relevant risk lists due to the judgment of negligible probability.

A (unanticipated) surprising event (with severe impacts) is thus a black swan according to this logic. An event of category (b) and (c) will obviously come as a surprise, but it is not so obvious when we talk about category (a) unknown unknowns—whence the dotted arrow in Fig. 3.1. Considering an activity with deep uncertainties about the type of events that will occur and the impact they will generate, we may be completely free of "expectations" for what is coming. Hence, it may be questioned whether an unknown unknown is in fact coming as a surprise in such a situation.

Similarly, we may problematise what is an unforeseen event. If an event occurs which was judged to have negligible probability, was it then foreseen? Yes, in the sense that the fact that it could happen was anticipated, but not in the sense that it was not considered likely.

Think of a container of fluid; normally it is filled with water and daily people drink from it. One day Ole drinks fluid from the container and it turns out to be of a toxic type. We refer to it as black swan, a surprise in relation to his knowledge/beliefs (assuming that it carries a serious impact). It need not be a new phenomenon we are witnessing, or an unknown unknown to be labelled a black swan. In retrospect, we explain the incident easily.

A risk analysis could have identified such an event; nevertheless, it may be surprising for some people (Ole), in relation to their beliefs/knowledge. These are the type of events we are concerned about. Let us modify the example a little. Suppose a risk analysis has identified various types of toxic fluids that could fill the container in special situations, but excludes a dangerous form because of a set of physical arguments. But then, this scenario happens. The event was possible despite the fact that it was considered impossible (extremely unlikely) by the analysts. The real-life conditions were not the same as those that were the basis for the risk analysis, and the event came as a surprise even for the risk analysts. In retrospect it was, however, easily explained.

In relation to the classifications (a)–(c) above, the event is classified as belonging to category (c). For the first case (Ole), the basis is the beliefs that this person has, and it can then be placed under (b) or possibly (c).

Strictly speaking, it would make sense to say that an unthinkable event is an unknown unknown. We may, however, also argue differently. Viewed from a risk assessment point of view, an event that belongs to category (b) may also be judged as unthinkable provided a thorough analysis has been performed to uncover all relevant events. The question is: unthinkable for whom?

It is common to refer to the term "unexpected" when characterising surprises and black swan type of events. However, this term is problematic to use. Consider the following example. An event has three possible outcomes, 0, 50 and 100, with associated probabilities 0.25, 0.50 and 0.25, respectively. Hence the expected value in a statistical sense (the centre of gravity of the probability distribution), is equal to 50. The values 0 and 100 can thus be seen as unexpected; however, the probability of one of these outcomes occurring is 50 %, which cannot be viewed as surprising. Clearly for this term "unexpected" to make sense, we need to interpret it in relation to the probability distribution. What is "unexpected" needs to be understood more as outcomes not belonging to a sufficiently broad uncertainty interval $[a, b]$, such that the probability of the quantity of interest not being covered by this interval is small, say less than 5 %.

3.4   Assessing and Managing Surprising Events and Black Swans

It is a common perception that the challenge in dealing with surprises and black swans lies in the fact that they are beyond the sphere of probability and risk (e.g. [31]). However, in line with the new (C, U) risk perspectives, the issue of confronting such events is at the core of risk management. A suitable conceptualisation for this has been established, summarised in the previous section; now the challenge is to develop adequate types of risk assessments and management policies. Let us first reflect on how this can be done for risk assessment.

3.4.1   Assessment

Many types of traditional risk assessment methods address the issue of what can happen, for example HAZOP, HazId, fault tree and event tree analysis [66]. Using these methods, hazardous events and scenarios are identified, but as risk is an issue, uncertainties and likelihood related to these events and scenarios are also addressed, in a qualitative or quantitative way. Based on considerations of probability, some events and scenarios can be judged to pose a negligible risk and not be followed up any further. With an increased focus on surprises and black swan types of events and scenarios, we need to reconsider the use of these methods. We have to challenge the premises that the analyses are based on, the assumptions that the probability judgements are relying on, etc., as for example indicated in [6, 12]. It is beyond the scope of this chapter to present specific methods for how to carry out such extended analyses; rather, the aim is to point to some key challenges and provide some preliminary reflections on possible routes for the research and development within these areas.

Kaplan et al. [39] provide some ideas for new tools in this direction, the so-called anticipatory failure determination (AFD). This method is an application of I-TRIZ, a form of the Russian-developed Theory of Inventive Problem Solving. This method is particularly suitable for scenarios involving human error, sabotage, terrorism, and the like. The relevance of the TRIZ methodology to risk analysis is rooted in the fact that revealing and identifying failure scenarios is fundamentally a creative act, yet it must be carried out systematically, exhaustively, and with diligence [39]. Traditional failure analysis addresses the question, "How did this failure happen?" or "How can this failure happen?". The AFD and TRIZ go one step further and pose the question, "If I wanted to create this particular failure, how could I do it?" The power of the technique comes from the process of deliberately "inventing" failure events and scenarios [48]. See [8] for a simple application example of this method.

These methods can be supported by different types of analysis frameworks, for example, Actor Network Theory (ANT) [44, 48]. ANT seeks to understand the dynamics of the system by following the actors – it asks how the world looks through

the eyes of the actor doing the work. As highlighted by [27, p. 1630] and [48], through this approach issues emerge pertaining to the roles that tools and other artefacts (actors) play in the actor-network in the accomplishments of their tasks.

TRIZ provides one type of creative thinking; there are many others, as discussed by for example [57]. There is obviously a potential for further developments of analysis methods for black swan types of events using some of these. The first step would be to assess the suitability of the various techniques for such a purpose.

The area of scenario analysis can add valuable input to the risk assessments. Here scenarios are developed describing potential future conditions and events, using various techniques [23]. There is no search for completeness and characterisations of the uncertainties and risks, as in traditional risk analysis, but in the case of large uncertainties and a lack of accurate prediction models, the generation of such scenarios may provide useful insights about what could happen and possible black swans. The deductive (anticipatory, backwards) scenarios are of particular importance in this respect, where we start from a future imagined event/state of the total system and question what is needed for this to occur. System thinking, which is characterised by seeing wholes and interconnections, is critical if we are to identify black swans, as for example highlighted by many scholars of accident analysis, organisational theory and the quality discourse [28, 58]. Using for example techniques such as event trees to reveal scenarios has strong limitations, as the analysis is based on linear inductive thinking [35, 46].

The risk description of an activity could be strongly influenced by the available signals and warnings, and hence we need to question to what degree the risk assessments are able to reflect these signals and warnings, in particular attitudes such as awareness and mindfulness in relation to these. A number of techniques exists which can contribute to enhance such attitudes, for example the use of the Johari Window concept [63] and ideas from the collective mindfulness concept used for High Reliability Organizations [61]. Further developments are, however, required to better link signals and warnings, with the associated awareness and mindfulness concepts, to risk. Some initial reflections on the topic are provided by [12, 41].

Another method that may be useful in revealing black swans is *red teaming*, which serves as a devil's advocate, offering alternative interpretations and challenging established thinking [48]. For example, "businesses use red teams to simulate the competition; government organisations use red teams as 'hackers' to test the security of information stored on computers or transmitted through networks; the military uses red teams to address and anticipate enemy courses of action" [16, p. 136]. Red teaming challenges assumptions, generalisations, pictures or images that influence how we understand the world and how we take action, i.e. our mental models [54]. [26, p. 39] argues that Murphy's Law is wrong: everything that can go wrong usually goes right, and then we draw the wrong conclusions. Red teaming can help us to see why, by pointing to alternative scenarios and outcomes.

Probabilities and interval probabilities may be used to express uncertainties and degrees of belief, but equally important is the knowledge these are based on. Crude probability judgements can indicate that some events are unlikely, but such statements need to be supplemented with assessments of strength of knowledge that the

judgements are based on and in particular the assumptions that they are based on. Ideas for how this can be conducted is outlined in [5] and [6], for example using the concept of assumption deviation risk.

3.4.2 Risk Management

To manage risk, three major categories of measures/strategies are commonly referred to [14, 52]: risk-based approaches (use of risk analysis), cautionary and precautionary approaches (including strategies based on robustness and resilience), and discursive strategies. The risk-based approach can be used alone only in cases when the knowledge is very strong and the uncertainties small. In most situations, all three strategies are required. The challenge is to find the adequate balance between these approaches and strategies, in particular between the first two. When the stakes are high and the uncertainties large, we obviously need to highlight robust and resilient solutions and arrangements to be prepared in case some extreme unforeseen events should occur. Potential surprises and black swans call for robustness and resilience, and antifragility, as argued for by [50], see also Sect. 3.1 and [10].

Adaptive risk management also needs to be mentioned here. This is a type of management which seeks to treat risk by considering a set of alternatives and dynamically tracking these to gain relevant information and knowledge about the effects of different courses of action [34, 64]. Such a strategy is especially attractive in the case of large/deep uncertainties, as discussed in [24]. This reference also points to several other risk management tools that can be used to confront deep uncertainties in a risk context, including learning what to do by well-designed and analysed trial and error. See [20] for an offshore application of adaptive risk management in a (C, U) risk perspective.

The risk-based approaches incorporate risk assessments but need to be extended and have a broader scope than the standard probabilistic analysis commonly seen in textbooks and practice today, as indicated in the previous section. A focus on knowledge building, transfer of experience and learning, represents an important means to reduce the risk related to surprises and black swans, by obtaining improved understanding of relevant systems and activities, models and the ability to predict what is coming. To provide a suitable foundation for such improvements, we need a platform that incorporates adequate concepts, assessment and management principles and methods. This is a research issue, and some ideas for how such a platform can be defined are presented by [12]. The idea is to integrate the conceptual framework of the (C, U) risk perspective as outlined in Sect. 3.2, with associated assessment and management principles and methods, and add theories and practical insights from other fields specifically addressing the knowledge dimension and the black swans. In [12], two areas are given main attention, first the collective mindfulness concept with its five principles: preoccupation with failure, reluctance to simplify, sensitivity to operations, commitment to resilience and deference to expertise. These elements represent criteria for checking the soundness of the understanding and treatment

of the risk at various stages of the accident scenario, including the planning phase
of the actual operation. There is a vast amount of literature (see e.g. [36, 45, 61,
62]) demonstrating that these five principles explain High Reliability Organizations
(HROs) well and that the collective mindfulness concept thus can be used as an
effective instrument for managing risks, the unforeseen and potential surprises.

In addition, the use of ideas from the quality discourse, with its link to the concepts
of "common-cause variation" and "special-cause variation", and the continuous focus
on learning and improvements, are added. Common causes capture "normal" system
variation, whereas the special causes are linked to the unusual variation and the
surprises (black swans) [18, 28, 55, 56].

There are a number of other risk management approaches, based on different
ideas and traditions, which could be added to those briefly mentioned above. See
for example [43, 46, 47]. However, for the purpose of the present chapter, the above
review suffices, as it provides a basis for discussing how the integrated framework
and thinking outlined above (summarised in [12]) can be used in a practical risk
management context. The main uses of the framework are:

1. As a general guideline for designing methods and tools for understanding, assess-
   ing, managing and communicating risk and safety
2. As a means for evaluating (also covering suggestions for improvement) the quality
   ("goodness") of various risk management activities, such as use of various types
   of risk assessment, barrier principles, risk management strategies (for example
   use of the ALARP[2] criterion)
3. As a supplement to standard accident investigation procedures, by drawing atten-
   tion to critical issues such as the violation of the five collective mindfulness prin-
   ciples, and issues related to knowledge, uncertainties, variation, the unforeseen,
   surprises and black swans, from different perspectives (individuals and groups)
   and points in time
4. As a means for evaluating the quality ("goodness") of a concrete risk assessment
   being conducted
5. As a means for evaluating the quality ("goodness") of the risk understanding in
   relation to critical operations, reflecting the ability to understand the total system,
   relevant knowledge, transfer of experience and learning.

To illustrate the use, let us look into some examples from the petroleum industry.

As a first example, let us consider the guidelines from the Norwegian Petroleum
Safety Authority (PSA-N) for barrier management principles [51]. With respect to
concepts, principles and main thinking for the understanding, assessment and man-
agement of risk and safety, it can be concluded that the document is very much in
line with the integrated framework presented here. This applies, for example, to the
meaning of the concept of risk. A further analysis of the five collective mindfulness
principles and aspects highlighted by the quality discourse reveals two main chal-
lenges. The first one relates to the reluctance to simplify criterion of the collective
mindfulness concept. Although the barrier principles highlight the need for a total

---

[2]ALARP: As Low As Reasonably Practicable.

system view with a focus on total barrier functions, the strong emphasis on the specification of detailed barrier element performance requirements may lead to difficulties in practice. Meeting the barrier element performance requirements may give the false perception that the risks are low and the barrier functions fulfilled. As we know, the connections between barrier element performance, risk and satisfying barrier functions are often unclear. Good performance numbers for the detailed barrier elements are no guarantee of safety. Holistic thinking is important, particularly for being able to identify black swans as well as for ensuring robustness and resilience. From a practical point of view, we may find that both the industry and the agency, through their auditing, are happy with a regime that highlights the barrier element as it is simple and easily followed up, compared to broader more judgemental assessments of barrier function performance and risks. The PSA-N is aware of this challenge, and it will be interesting to see how the implementation process will proceed.

The second challenge of these barrier management principles relates to management by objectives and the compliance focus, and is linked to the discussion above concerning the reluctance to simplify. The approach recommended by the PSA-N document has a strong emphasis on formulating, assigning and satisfying performance requirements, which can easily lead to an excessively strong focus on meeting requirements rather than on identifying the overall best solutions and measures. Experience has shown that such an approach represents a serious challenge – the compliance regime prevails, much more than the improvement processes which are always highlighted in theory but often fail to be given priority when in competition with the convenience and practical attractiveness of compliance procedures.

The second example of applications 1.–5. relates to the use of job safety analysis in an industrial setting. In a job safety analysis, hazards linked to the various subtasks are identified, and risk is described by a risk matrix covering consequence and probability. Based on this risk description, an overall judgement about safety and risk acceptability is made. An evaluation of this method (Use 2), quickly reveals that the way risk is described is based on a probability-based thinking on risk, and, as discussed in Sect. 3.2, this perspective can be criticised by not properly reflecting the strength of knowledge that the assessments are based on, uncertainties, surprises and black swans. It is beyond the scope of the present chapter to provide a detailed solution to this challenge, but some ideas have already been pointed to above, in the last paragraph of Sect. 3.4.1, by incorporating judgements of the strength of knowledge that the probability-based judgements are based on, and making assessments of the so-called assumption deviation risk. Specific related black swan type of analysis should also be considered, some type of red-team analysis as mentioned in Sect. 3.4.1, for example using the ideas outlined in [6]:

> First make a list of all types of risk events having low risk as scored by the three dimensions: assigned probability, consequences, and strength of knowledge. Then carry out a review of all possible arguments and evidence for the occurrence of these events, for example by identifying historical events and experts' judgements not in line with common beliefs. To conduct these assessments, experts are needed that are not members of the core group of analysts conducting the more standard parts of risk assessment. The idea is to allow for and stimulate different views and perspectives, in order to break free from the prevailing beliefs and obtain creative processes.

Further examples can be found in [8]. This leads us to the third example, linked to the fourth use of the integrated framework. A risk analysis team has conducted a risk assessment in relation to a critical maintenance operation, a job safety analysis. To evaluate the quality of this analysis, we apply the principles of the integrated framework, presuming that the intention now is to conduct it in line with the principles just outlined. A check is made that the analysis covers what can go wrong, causes, barriers, consequences and likelihood, as well as issues linked to uncertainties, knowledge, surprises and black swans. A key point is that the analysis is clear on what are the key assumptions and beliefs that the judgements are based on, and that these assumptions and beliefs are challenged as discussed in an appropriate way. Key control questions relate to the data and information supporting the assessment, and the ability of the analysts to provide a meaningful presentation of the results, which is clear on what the analysis provides and what the limitations are. In the concrete example studied, it turned out that the results relied on a key assumption which was not questioned in the uncertainty analysis, despite the fact that the review showed that some critical questions to it had been raised in some other contexts. These questions were, however, not known to the analysis team first performing the risk analysis.

This critical maintenance operation is also considered in the example adopted for the fifth use. We would like here to apply the framework to evaluate the risk understanding of personnel that are to carry out the work. The focus is on the potential hazards, their causes and consequences, including barriers, as well as the likelihood. Special attention is paid to signals and warnings that can give an indication of the occurrence of some more severe hazardous situations, and how to sense these signals and warnings and make adequate adjustments. The degree to which the insights provided by the relevant risk assessments are known and understood is questioned. This relates to all the issues mentioned above (what can go wrong, causes, barriers, consequences and likelihood), but also covers issues linked to uncertainties, knowledge, surprises and black swans (if such issues have not been addressed by the risk assessment, a separate analysis of such issues needs to be added to be in line with the recommended thinking). A key issue is to reveal what are the main assumptions and beliefs that the judgements are based on, and how to cope with surprises relative to these assumptions and beliefs. In the specific case considered here, the fourth collective mindfulness principle was pointed to as a weakness, as the issue of robustness and resilience had not been properly thought through. The current standard in the company for how to carry out such operations, fails to stress the importance of commitment to resilience. Moreover, it was identified that key personnel had a lack of understanding of the system, and that important experience from some recent operations had not been captured by the risk assessment and was not known to all the personnel involved in carrying out the critical operation.

For an example linked to accident investigation (Use 3.), see [59, p. 82]. See [6, 12] for some other examples of Use 2., linked to risk acceptance criteria and the implementation of the risk reduction principle ALARP.

3.5 Conclusions

We have reviewed some recent advances in the risk field, linked to the conceptualisation of risk and specifically addressing unforeseen events, surprises and so-called black swans. Some ideas for new types of risk analyses were briefly outlined before an integrated framework was presented which extends the traditional probability-based perspectives on risk to broader ways of thinking about risk, which give due attention to the uncertainties and also draw on ideas from the quality discourse and organisational learning (collective mindfulness and its five principles: preoccupation with failure, reluctance to simplify, sensitivity to operations, commitment to resilience and deference to expertise). Using some simple examples it is shown how this framework can be used in practice, as a tool to improve the understanding, assessment, management and communication of risk. Further research and testing are, however, required to be able to use this framework in practical settings. Such research and testing are currently being conducted in the Norwegian oil and gas industry.

References

1. Althaus CE (2005) A disciplinary perspective on the epistemological status of risk. Risk Anal 25(3):567–588
2. Aven T (2000) Risk analysis — a tool for expressing and communicating uncertainty. In: Proceedings of ESREL 2000, Balkema Publishers, Edinburgh, 21–28
3. Aven T (2009) Safety is the antonym of risk for some perspectives of risk. Saf Sci 47(7):925–930
4. Aven T (2011) On some recent definitions and analysis frameworks for risk, vulnerability, and resilience. Risk Anal 31(4):515–522
5. Aven T (2012) The risk concept — historical and recent development trends. Reliab Eng Syst Saf 99:33–44
6. Aven T (2013) Practical implications of the new risk perspectives. Reliab Eng Syst Saf 115:136–145
7. Aven T (2013) Probabilities and background knowledge as a tool to reflect uncertainties in relation to intentional acts. Reliab Eng Syst Saf 119:229–234
8. Aven T (2014) Risk, surprises and black swans. Routledge, Abingdon
9. Aven T (2015) The concept of antifragility and its implications for the practice of risk analysis. Risk Anal 35(3):476–483
10. Aven T, Baraldi P, Flage R, Zio E (2014) Uncertainty in risk assessment: the representation and treatment of uncertainties by probabilistic and non-probabilistic methods. Wiley, Chichester
11. Aven T, Kristensen V (2005) Perspectives on risk: review and discussion of the basis for establishing a unified and holistic approach. Reliab Eng Syst Saf 90(1):1–14
12. Aven T, Krohn BS (2014) A new perspective on how to understand, assess and manage risk and the unforeseen. Reliab Eng Syst Saf 121:1–10
13. Aven T, Renn O (2009) On risk defined as an event where the outcome is uncertain. J Risk Res 12(1):1–11
14. Aven T, Renn O (2010) Risk management and governance: concepts guidelines and applications. Springer, Heidelberg
15. Aven T, Zio E (2011) Some considerations on the treatment of uncertainties in risk assessment for practical decision making. Reliab Eng Syst Saf 96(1):64–74

16. Ambrose F, Ahern B (2008) Unconventional red teaming. In: Anticipating rare events: can acts of terror, use of weapons of mass destruction or other high profile acts be anticipated?
17. Beck U (1992) Risk society: towards a new modernity. Sage, London
18. Bergman B (2008) Conceptualistic pragmatism: a framework for Bayesian analysis? IIE Trans 41(1):86–93
19. Ben Ale JM (2002) Risk assessment practices in the Netherlands. Saf. Sci. 40:105–126
20. Bjerga T, Aven T (2015) Adaptive risk management using new risk perspectives – an example from the oil and gas industry. Reliab Eng Syst Saf 134:75–82
21. Brown GG, Cox LA Jr (2011) How probabilistic risk assessment can mislead terrorism risk analysts. Risk Anal 31(2):196–204
22. Campbell S, Currie G (2006) Against Beck: in defence of risk analysis. Philos Soc Sci 36(2):149–172
23. Chermack TJ (2011) Scenario planning in organizations: how to create, use, and assess scenarios. BK publishers, San Francisco
24. Cox LAT (2012) Confronting deep uncertainties in risk analysis. Risk Anal 32(10):1607–1629
25. de Moivre A (1711) De mensura sortis. Philos Trans 27:213–264
26. Dekker SW (2011) Drift into failure: from hunting broken components to understanding complex systems. Ashgate, Aldershot
27. Dekker SW, Nyce JM (2004) How can ergonomics influence design? Moving from research findings to future systems. Ergonomics 47(15):1624–1639
28. Deming WE (2000) The new economics for industry, government, education. MIT Press, Cambridge
29. Douglas M, Wildavsky A (1982) Risk and culture: an essay on the selection of technological and environmental dangers. University of California Press, Berkeley
30. Rosa, EA (1998) Metatheoretical foundations for post-normal risk. J Risk Res 1(1):15–44
31. Gross M (2010) Ignorance and surprise. MIT Press, London
32. Hardy CO (1923) Risk and risk bearing. University of Chicago Press, Chicago
33. Haynes J (1895) Risk as an economic factor. Quart J Econ 9(4)
34. Holling CS (ed) (1978) Adaptive environmental assessment and management. Wiley, London
35. Hollnagel E, Woods DD, Leveson N (2006) Resilience engineering: concepts and precepts. Ashgate Publishing, Aldershot
36. Hopkins A (2013) Issues in safety science. Saf Sci 67:6–14
37. Jasanoff S (1999) The songlines of risk. Environ Value 8(2):135–152
38. Kaplan S, Garrick BJ (1981) On the quantitative definition of risk. Risk Anal 1(1):11–27
39. Kaplan S, Visnepolschi S, Zlotin B, Zusman A (1999) New tools for failure and risk analysis: anticipatory failure determination (afd) and the theory of scenario structuring. Ideation International Inc., Southfield
40. Kay NM (1984) The emergent firm: knowledge ignorance and surprise. Macmillan, London
41. Khorsandi J, Aven T (2014) A risk perspective supporting organizational efforts for achieving high reliability. J Risk Res 17(7):871–884
42. Knight FH (1921) Risk, uncertainty and profit. Mifflin Company, Houghton
43. Kunreuther H, Useem M (2009) Learning from catastrophes: strategies for reaction and response. Wharton School Publishing, Upper Saddle River
44. Latour B (2005) Reassembling the social: an introduction to actor network theory. Oxford University Press, Oxford
45. Le Coze J-C (2013) Outlines of a sensitising model for industrial safety assessment. Saf Sci 51(1):187–201
46. Leveson N (2012) Engineering a safer world: systems thinking applied to safety. MIT Press, Cambridge
47. Loch CH, DeMeyer A, Pich M T (2006) Managing the unknown: a new approach to managing high uncertainty and risk in projects. Wiley, Hoboken
48. Masys AJ (2012) Black swans to grey swans: revealing the uncertainty. Disaster Prev Manag 21(3):320–335

49. Taleb NN(2007) The black swan: the impact of the highly improbable. Random House, New York
50. Taleb NN (2012) Antifragile: things that gain from disorder. Random House, New York
51. PSAN (2013) Principles for barrier management in the petroleum industry. Technical Report 29/1- 2013, Norwegian Petroleum Safety Authority (PSA)
52. Renn O (2008) Risk governance: coping with uncertainty in a complex world. Routledge, Abingdon
53. Rosa EA (2003) The social amplification of risk, chapter the logical structure of the social amplification of risk framework (sarf): metatheoretical foundation and policy implications. Cambridge University Press, Cambridge
54. Senge P, Roberts C, Smith BJ (1990) The fifth discipline fieldbook: strategies and tools for building a learning organization. Currency, New York
55. Shewhart WA (1931) Economic control of quality of manufactured product. Van Nostrand, New York
56. Shewhart WA (1939) Statistical method from the viewpoint of quality control. Dover Publications, Washington
57. Sternberg RJ (1999) Handbook of creativity. Cambridge University Press, Cambridge
58. Turner BA, Pidgeon NF (1997) Man-made disasters, 2nd edn. Butterworth-Heinemann, Amsterdam
59. UCB (2011) Final report on the investigation of the Macondo well blowout. Technical report, University of California at Berkeley, Deepwater Horizon Study Group
60. Vatn J (1998) A discussion of the acceptable risk problem. Reliab Eng Syst Saf 61:11–19
61. Weick K, Sutcliffe K (2007) Managing the unexpected: resilient performance in an age of uncertainty, 2nd edn. Jossey-Bass, San Francisco
62. Weick KE, Sutcliffe KM, Obstfeld D (1999) Organizing for high reliability: processes of collective mindfulness, Research in organizational behaviour. Elsevier, New York, pp 81–123
63. Weiss A, Wright S (2006) Dealing with the unknown: a holistic approach to marketing and competitive intelligence. Compet Intell 9(5):15–20
64. Wintle B, Lindenmayer D (2008) Adaptive risk management for certifiably sustainable forestry. For Ecol Manag 256(6):1311–1319
65. Wynne B (1992) Social theories of risk. Risk and social learning: reification to engagement, Praeger, Westport, p 275–297
66. Zio E (2007) An introduction to the basics of reliability and risk analysis. World Scientific, Singapore