ON HOW TO CONCEPTUALISE AND DESCRIBE RISK

T. Aven

University of Stavanger, Stavanger, Norway
terje.aven@uis.no

ABSTRACT

A number of definitions and interpretations of the risk concept exist. Many of these are probability-based. In this paper we present and discuss a structure for characterising the definitions, which is founded on a clear distinction between (a) risk as a concept based on events, consequences and uncertainties; (b) risk as a modelled, quantitative concept; and (c) risk descriptions. The discussion leads to recommended perspective for conceptualising and assessing risk, which is based on risk defined by (a), and the probability-based definitions of risk can be viewed as related model parameters and/or risk descriptions. Two ways of detailing the framework are outlined: the relative frequency-based approach and the Bayesian approach.

1 INTRODUCTION

Risk is a fundamental concept for most scientific disciplines, but no consensus exists on how to define and interpret risk. Some definitions are based on probabilities, some on expected values, and others on uncertainty. Some consider risk as subjective and epistemic, dependent on the available knowledge, whereas others grant risk an ontological status independent of the assessors. The situation is chaotic and leads to poor communication. We are also afraid that it hampers effective risk management as well as the development of the risk field, as many of these definitions and interpretations lack proper scientific support and justification.

Of course, business needs a different set of risk methods, procedures and models than, for example, medicine and engineering. But there is no reason why these areas should have completely different perspectives on how to think when approaching risk and uncertainty, when the basic challenge is the same---to conceptualise that the future performance of a system or an activity could lead to outcomes different from those desired and planned, or not in line with stated objectives.

Think of an activity in the future, say the operation of an offshore installation for oil and gas processing. We all agree that there is some risk associated with this operation. For example, fire and explosions could occur leading to fatalities, oil spills, economic loss, etc. But it is not straightforward to explain what we mean by this risk if we require a precise definition and would like to use the concept in scientific studies. Risk analysts would introduce a set-up which directly or indirectly defines how risk is understood and assessed. The set-up would typically be probability-based, with probabilities interpreted either as relative frequencies or as subjective probabilities. An example would be the traditional statistical approach which considers risk as a relative frequency-interpreted probability or probability distribution, and the aim of the risk assessment is to accurately estimate this risk using models and hard data. All such set-ups can be challenged, as not being able to reflect risk in a proper way. Important risk aspects could be camouflaged or hidden by the set-up. Discussions of the set-up are therefore important, not only from a theoretical point of view but also from a practical risk management perspective. Many researchers have contributed to this discussion, e.g. Reid (1992) and Stirling (2007). Reid argues
that there is a common tendency to underestimate the uncertainties in risk assessments. According to Stirling (2007), using risk assessment when strong knowledge about the probabilities and outcomes does not exist, is irrational, unscientific and potentially misleading. Many other critical comments could have been added, but for the purpose of the present paper it is sufficient to conclude that there is a discussion in the scientific literature about the ability of the set-up of risk assessments to adequately reflect risk.

To be able to make judgments about this issue we need to clarify what risk is and how risk can and should be described. This is the topic of the present paper. A main purpose of the paper is to present a structure for characterising the various definitions of risk in a scientific context. This structure is based on a clear distinction between (a) risk as a concept based on events, consequences and uncertainties; (b) risk as a modelled, quantitative concept; and (c) risk descriptions. Examples of these categories are:

- Uncertainty about the occurrence of future events and their consequences (a)
- Frequentist-interpreted probability $P_f$ of an event (b)
- Estimates of $P_f$ (c)
- A subjective probability $P_s$ (c).

From this structure we establish a framework that integrates the (a), (b) and (c) definitions to obtain a hierarchy with the a) definitions as the overall risk concept. To further specify the framework we need to distinguish between the relative frequency-based approach and the Bayesian approach. This framework is the main contribution of the present paper.

In the paper we identify several definitions of risk that can be used as an overall, common definition. They all belong to the category (a). Many attempts have been made to establish a unified risk perspective, but none of these have obtained broad acceptance in practice. There could be many reasons for this. Firstly, the scientific work on risk may not have reached a sufficiently mature level for establishing such a definition. The exploring phase is not completed. Secondly, the scientific literature has a focus on the generation of new ideas and suggestions, and on a critique of other contributions. By its nature, it is hard to obtain broad consensus on scientific issues in general and risk definitions in particular. And thirdly, the standardisation organisations have not been able to produce sufficient broad and well-defined definitions which could be accepted by the scientific expertise on risk.

Consider for example the latest proposal from the International Standardisation Organisation (ISO 2009) for defining risk: Risk is the effect of uncertainty on objectives. What does this mean? Risk has to do with uncertainty, but is it the effect of uncertainty? And risk is related to objectives, but what if objectives are not defined? Then we have no risk? Asking experts on risk, there is no doubt that this definition would lead to numerous different interpretations. The definition is not sufficiently precise, and one may certainly also question its rationale as indicated.

The present paper is partly based on Aven (2010a). For some reflections on the ontological status of the concept of risk, see Aven et al. (2010).

2 A CLASSIFICATION OF RISK DEFINITIONS BASED ON THE PROPOSED STRUCTURE

As stressed above, there exist a number of definitions of risk. Here are some typical examples (list based on Aven and Renn 2009a):

1) Risk equals the expected loss (Verma and Verter 2007, Willis 2007).
2) Risk equals the expected disutility (Campbell 2005).
3) Risk is a measure of the probability and severity of adverse effects (Lowrance 1976).
4) Risk is the combination of probability and extent of consequences (Ale 2002).
5) Risk is equal to the triplet ($s_i$, $p_i$, $c_i$), where $s_i$ is the $i$th scenario, $p_i$ is the probability of that scenario, and $c_i$ is the consequence of the $i$th scenario, $i = 1, 2, \ldots, N$ (Kaplan and Garrick 1981).
6) Risk refers to uncertainty of outcome, of actions and events (Cabinet Office 2002).
7) Risk is a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain (Rosa 1998, 2003).
8) Risk is an uncertain consequence of an event or an activity with respect to something that humans value (Renn 2005).
9) Risk is the effect of uncertainty on objectives (ISO 2009).
10) Risk is equal to the two-dimensional combination of events/consequences and associated uncertainties (Aven 2007, Aven 2008).
11) Risk is uncertainty about and severity of the consequences (or outcomes) of an activity with respect to something that humans value (Aven and Renn 2009a).

For the measures that are based on probabilities and expected values, we may generate two versions, one where the probabilities are interpreted as relative frequencies (and the expected values as averages), and one where the probabilities are subjective probabilities (and the expected value is interpreted as the centre of gravity of the probability distribution). We write definitions $x_f$ and $x_s$, respectively, to separate the two categories, $x = 1, 2, \ldots, 5$. Consider as an example category 1, risk defined as the expected loss. According to definition 1$_f$, risk is understood as the average loss when considering an infinite number of similar situations, whereas 1$_s$ means that risk is the centre of gravity of the subjective probability distribution of the loss. Following the suggested structure for characterising the various risk definitions we have to place these definitions in one of the categories (a), (b) (c), defined in the previous section.

The result is that definition 1$_f$ is in category (b) and 1$_s$ is in category (c), as risk in the former case is based on the model of an infinite number of similar situations and risk in the latter case is a way for the assessor to describe or characterise risk. The expected loss $E_x$ when using subjective probabilities is a risk index based on the background knowledge (K) of the assessor.

This is in line with the rejection of risk as being defined by the expected value, as argued in, for example, Haimes (2004) and Aven (2010b). The expected value does not adequately capture events with low probabilities and high consequences. Take as examples nuclear accidents and terrorism risk, where the possible consequences could be extreme and the probabilities are relatively low. The expected value can be small, say 0.01 fatalities, but extreme events with millions of fatalities may occur, and this needs special attention.

A similar analysis is carried out for the other ten definitions. The result is shown in Table 1.

Table 1. Categorisation of the 11 risk definitions according to the structure (a) – (c). For definition 2 it is assumed that the expectation is taken with respect to a subjective probability distribution.

<table>
<thead>
<tr>
<th>Risk definition</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>1$_f$</td>
<td>b</td>
</tr>
<tr>
<td>1$_s$</td>
<td>c</td>
</tr>
<tr>
<td>2</td>
<td>c</td>
</tr>
<tr>
<td>3$_f$</td>
<td>b</td>
</tr>
<tr>
<td>3$_s$</td>
<td>c</td>
</tr>
<tr>
<td>4$_f$</td>
<td>b</td>
</tr>
<tr>
<td>4$_s$</td>
<td>c</td>
</tr>
</tbody>
</table>
Some comments are in place for the various definitions (2-11).

The second definition considers risk as the expected disutility, i.e. \( \text{Eu}(C) \), where \( C \) is the outcomes (consequences) and \( u(C) \) the utility function (Campbell 2005). The expectation is based on subjective probabilities. According to this definition, the preferences of the decision-maker are a part of the risk concept. In our view, and this view is shared by many risk experts, the preferences and values should not be a part of the risk concept and the risk assessments (Paté-Cornell 1996). There will be a strong degree of arbitrariness in the choice of the utility function, and some decision-makers would also be reluctant to specify the utility function as it reduces their flexibility to weight different concerns in specific cases. Risk should be possible to describe also in cases where the decision-maker is not able or willing to define his/her utility function.

Definitions 3-5 are all probability-based. The concept of risk comprises events (initiating events, scenarios), consequences (outcomes) and probabilities. Severity is a way of characterising the consequences, and refers to intensity, size, extension, scope and other potential measures of magnitude, and affects something that humans value (lives, the environment, money, etc.). Losses and gains, for example expressed by money or the number of fatalities, are ways of defining the severity of the consequences (Aven and Renn 2009a).

If relative frequency-interpreted probabilities \( P_f \) constitute the basis (definitions 3f, 4f and 5f), risk is a modelled, quantitative concept (category (b)) and we may formalise the definitions by writing

\[
\text{Risk} = (A,C,P_f),
\]

where \( A \) represents the events (initiating events, scenarios) and \( C \) the consequences of \( A \). Examples of events \( A \) are: gas leakage occurring in a process plant, and the occurrence of a terrorist attack. Examples of \( C \) are the number of casualties due to leakages, terrorist attacks, etc.

If on the other hand subjective probabilities constitute the basis (definitions 3s, 4s and 5s, the definitions must be viewed as risk descriptions as they express the analysts’ degree of belief concerning \( A \) and \( C \). Also the background knowledge \( K \) that the probabilities are based on, should be considered a part of the risk description.

A quick look at definitions 6-11 may give the impression that they are not that different from 3-5. However, there are important principle differences, as will be clear from the coming analysis. Probability is just a tool used to represent or express the uncertainties. The thesis of all the perspectives and definitions 6-11 is that risk should not be limited to \( (A,C,P) \). The uncertainties should be highlighted. Consider first definition 10, which we simply refer to as the \( (A,C,U) \) definition. Definition 11 may be viewed as a reformulation of this definition, based on the same ideas.

We consider an activity in the future, and something that humans value is at stake (lives, the environment, etc.). Undesirable (and desirable) events and consequences could occur. There are uncertainties about the occurrence of the events, and what will be the consequences (outcome) of these events if they should occur. How many will be killed? What will the value of the stock be?
Risk has two main components: i) the events and their consequences, and ii) uncertainty about these - will the events occur and what will the consequences be? These two components define risk.

According to definition 6, risk refers to uncertainty of outcome, of actions and events (Cabinet Office 2002). Hence, strictly speaking risk is not (A,C,U) but only U. As an example, consider the number of fatalities in traffic next year in a specific country. Then the uncertainty is rather small, as the number of fatalities shows rather small variations from year to year. Thus following this definition of risk, we must conclude that the risk is small, even though the number of fatalities is many thousands each year. Clearly, this definition of risk fails to capture an essential aspect, the consequence dimension. Uncertainty cannot be isolated from the intensity, size, extension etc. of the consequences.

According to definition 7, risk is a situation or event where something of human value (including humans themselves) is at stake and where the outcome is uncertain (Rosa 1998, 2003). Hence, strictly speaking risk is A, and not (A,C,U). However, Rosa expresses risk using the description (A,C,U), and refers to probability as a tool to describe the uncertainties. The Rosa (1998, 2003) definition is thoroughly discussed by Aven and Renn (2009a). The conclusion is that compared to common terminology, the Rosa definition leads to conceptual difficulties that are incompatible with the everyday use of risk in most applications. By considering risk as an event (A), we cannot conclude, for example, about the risk being high or low, or compare different options with respect to risk. The same conclusion is made for definition 8, which says that risk is an uncertain consequence of an event or an activity with respect to something that humans value (Renn 2005). This definition is similar to Rosa’s definition but the event A is replaced by the consequence C.

We have already commented on definition 9 that risk is the effect of uncertainty on objectives (ISO 2009). This definition seems to be in line with the two previous definitions. Alternatively, we may interpret the suggested ISO definition as (A,C,U), where the consequences (the effect) are seen in relation to the objectives.

Based on the perspectives and definitions 6-11, various types of risk descriptions (category (c)) can be specified, for example by using subjective probabilities. But we can also introduce modelled, quantitative risk concepts (category (b)). The result is a hierarchy of concepts which together provide a holistic system for conceptualising and describing risk. The next section will present the details of this system.

3  RECOMMENDATIONS FOR HOW TO CONCEPTUALISE AND DESCRIBE RISK

If we search for a widespread agreement on one definition we have to look among the categories (a). The others have to be excluded as they are based on either a model or an assignment of uncertainty using the tool, subjective probability. Risk should also exist as a concept without modelling and subjective probability assignments. We face risk when we drive a car or run a business, also when probabilities are not introduced. For risk assessment we need the probabilities, but not as a general concept of risk. In this way we obtain a sharp distinction between risk as a concept and risk descriptions (assessments).

The discussion in the previous section led to two candidates among the a-definitions; the (A,C,U) definitions (10-11) and the (A,C) definitions (7-8). The latter group means that the common risk terminology has to be revamped and we therefore prefer to use the (A,C,U) definition.

We will use this risk concept as a pillar for a recommended framework for conceptualising and describing risk. The next stage would then be to specify how to describe risk in this framework. To be able to do this we need to distinguish between a relative frequency-based approach and a Bayesian perspective as will be demonstrated by the following analysis. The main elements of the
frameworks for these two approaches are shown in Figures 1 and 2. We first look at the relative frequency case.

In this case we introduce relative frequency-interpreted probabilities $P_f$. These are in general unknown. Risk assessment is introduced to describe the risk, to estimate $P_f$. The description covers an estimate $P_f^*$ of $P_f$, as well as assessments of uncertainties about $P_f^*$ and $P_f$. Thus, if the relative frequency perspective to risk is the starting point, we are led to a risk description:

i) Risk description in the relative frequency case = $(A,C,P_f^*,U(P_f^*),U,K)$,

where $U(P_f^*)$ refers to an uncertainty description of $P_f^*$ relative to the true value $P_f$, $U$ refers to uncertainty factors not covered by $U(P_f^*)$, and $K$ is the background knowledge that the estimate and uncertainty description is based on. We may refer to $U(P_f^*)$ as a second-order uncertainty description.

One way of reflecting $U(P_f^*)$ is to use confidence intervals. These intervals describe the variation in the data available, but do not reflect other types of uncertainties, in particular uncertainties as a result of more or less relevant data.

If we use subjective probabilities $P_s$ to express our uncertainties about $P_f$, the risk description takes the form:

i') Risk description = $(A,C,U, P_s(P_f^*),P_s,K)$,

where $K$ now is the background knowledge that the estimate $P_f^*$ and the probability distribution $P_s$ is based on. Kaplan & Garrick (1981), see also Kaplan (1997), refer to this distribution as the second level definition of risk – it is combined with the first level $(A,C,P)$ definition. When including the second level definition the perspective is referred to as the probability of frequency approach. The risk description i)' can be viewed as an extended probability of frequency approach, as it covers all the elements of the probability of frequency approach and in addition address uncertainties $U$ not reflected by the $P_s$.

The $U$ covers in general factors not included in $U(P_f^*)$ or $P_s(P_f)$. Examples include the relevancy of the data when using confidence intervals and the fact that the subjective probabilities could produce poor predictions. The background knowledge $K$ could be poor. Probability assignments are conditioned on a number of assumptions and suppositions, and these could turn out to be wrong. One may assign a low probability of health problems occurring as a result of some new chemicals, but these probabilities could produce poor predictions of the actual number of people that experience such problems. Or one may assign a probability of fatalities occurring on an offshore installation based on the assumption that the installation structure will withstand a certain accidental load; in real-life the structure could however fail at a lower load level: the assigned probability did not reflect this uncertainty.

The analysts need to clarify what is uncertain and subject to the uncertainty assessment and what constitutes the background knowledge. From a theoretical point of view, one may think that it is possible (and desirable) to remove all such uncertainties from the background knowledge, but in a practical risk assessment context that is impossible. The assessment of the uncertainty factors would normally be qualitative, see approach indicated below.

Next we consider the Bayesian case. A risk description based on this definition would cover the following components:

ii') Risk description = $(A,C,U,P_s,K)$,

where $P_s$ is a subjective probability expressing $U$ based on the background knowledge $K$. This description covers probability distributions of $A$ and $C$, as well as predictions of $A$ and $C$, for
example a predictor $C^*$ given by the expected value of $C$, unconditionally or conditional on the occurrence of $A$, i.e. $C^* = EC$ or $C^* = E[C|A]$.

Using the description ii)’ there are no second-order probabilities, as talking about uncertainties of a subjective probability has no meaning. A subjective probability $P(A) = P(A|K)$ is interpreted as a knowledge-based probability with reference to a standard expressing the analysts’ uncertainty about the occurrence of the event $A$ given the background knowledge $K$. Following this interpretation the assessor compares his/her uncertainty (degree of belief) about the occurrence of the event $A$ with the standard of drawing at random a favourable ball from an urn that contains $P(A) \cdot 100\%$ favourable balls (Lindley 2006). The traditional betting interpretation of a subjective probability (Singpurwalla 2006) can also be used, but we prefer the reference to a standard definition as it does not mix uncertainty assessments with our attitude to money (Aven 2003, Aven 2010c). According to the betting interpretation, the probability of an event is the price at which the person assigning the probability is neutral between buying and selling a ticket that is worth one unit of payment if the event occurs, and worthless if not (Singpurwalla 2006).

Also in the Bayesian context we establish relative frequencies, but they are referred to as chances and not probabilities. A chance is the limit of a frequency of similar (formally exchangeable) random events. More generally we introduce probability models with unknown parameters. A chance is an example of such a parameter. By the Bayesian updating machinery, knowledge about the parameters is described first by the prior distribution, then updated to produce the posterior distribution to reflect observations. Finally, this distribution is used to generate the predictive distribution of the events $A$ and consequences $C$. These predictive distributions then incorporate the variation reflected by the probability model (and the chances) and the epistemic uncertainties about the true value of the parameters.

If probability models and chances are introduced, the Bayesian approach looks similar to the extended probability of frequency approach. However, there is a difference. In the relative frequency case, probabilities $P_f$ always need to be defined. They constitute the foundation of the approach. In the Bayesian case, chances are only defined when exchangeable sequences can be justified. Chances need some sort of model stability (Bergman 2009): populations of similar units need to be constructed (formally an infinite set of exchangeable random variables). We will for example not define a chance $p$ of an attack (Aven & Renn 2009b). It has no meaning. Subjective probabilities can however be used.

![Figure 1. The main elements of the recommended risk framework when it is based on relative frequencies](image-url)
Models, including probability models, are used in both cases. Instead of estimating $P_r$ or chances we estimate $g(q)$, where $g$ is the model and $q$ is a vector of parameters of the model. An event tree and a fault tree are two simple examples of such models. We may also use models to simplify and/or give rigour to the specification of the subjective probabilities $P_s$.

The U in the risk description covers an assessment of uncertainties that are not captured by the probabilistic analysis. A simple approach for how to do this assessment is outlined in Flage and Aven (2009). Assume that the events $A$ are described by probabilities $P(A|K)$ and conditional expected impact values given the occurrence of $A$, $E[C|A,K]$, are determined. For each $A$, a list of uncertainty factors is identified. These factors are assessed with respect to uncertainty and classified as “high” if one or more of the following conditions are met:

- The assumptions made represent strong simplifications.
- Data are not available, or are unreliable.
- There is lack of agreement/consensus among experts
- The phenomena involved are not well understood

In addition, the expected impact values must be sensitive to changes of the factor. The point is that a factor could be subject to large uncertainties but it is not considered important unless changes in the factor affect the impact.

4 CONCLUSIONS AND FINAL REMARKS

In this paper we have presented a new framework for conceptualising and assessing risk. Compared to earlier analyses of the risk concept (Aven 2009a,b, Aven and Renn 2009a), the framework provides a structure for both the relative frequency approach and the Bayesian approach within the same overall risk concept $(A,C,U)$. In this way two approaches are developed and specified. There could be different opinions on which approach should be preferred, but only the Bayesian approach would work in cases where relative frequency-interpreted probabilities (chances) cannot be meaningfully defined. In this sense the Bayesian approach is more general than the relative frequency approach.
By this framework it is acknowledged that risk is more than probabilities, probability distributions and expected values. The uncertainty dimension of risk extends beyond the probabilities. In this way the framework provides important input for making judgments about the quality of risk assessments. If a risk assessment is restricted to probabilities, important aspects of risk may be overlooked.

It may be a challenge to reveal and describe all the uncertainties. Qualitative approaches can be used, and further research is required to develop methods for proper identification and analysis of the uncertainties. But this is not the issue here. In this paper we address the overall conceptual structure of risk and risk assessment, not the analysis methods. Before such conceptual structures can be established it is difficult to develop suitable methods, as the methods would depend on the aim of the analyses. Risk analysis is a young discipline and has been characterised by many weakly justified risk perspectives and also by lack of consistency in approaches. The aim of the present paper has been to contribute to rectifying these problems by suggesting an overall holistic framework for conceptualising and assessing risk that could provide improved structure and guidance on how to think in a risk analysis context.

Acknowledgments

The work has been funded by The Research Council of Norway through the SAMRISK and PETROMAKS research programmes. The financial support is gratefully acknowledged.

References

24. William Kaufmann Inc.