Three influential risk foundation papers from the 80s and 90s: Are they still state-of-the-art?

Terje Aven

University of Stavanger, Norway

ARTICLE INFO

Keywords:
Risk science
Risk characterization
Uncertainties
Models

ABSTRACT

Three of the most influential scientific works in the risk field, at least in the engineering environment, are Stan Kaplan and John Garrick’s paper from 1981 on risk quantification, George Apostolakis’ paper on probability from 1990, and Elisabeth Paté-Cornell’s paper on uncertainty levels in risk assessments from 1996. The present article reviews and discusses these works, the aim being to acknowledge their important contributions to risk science and provide insights on how these works have influenced and relate to the state-of-the-art of the risk science of today. It is questioned to what extent these papers still represent state-of-the-art. Recent documents by the Society for Risk Analysis are used as a reference for comparison, in addition to related publications in scientific journals.

1. Introduction

The risk field and science is young. Considerable foundational work was conducted in the 80s and 90s, as reflected by the many papers on such issues in that period, for example in Risk Analysis and Reliability Engineering & System Safety; see, for example, overviews in SRA [82]. A science can be viewed as the most warranted statements that can be made at the time being on the subject matter covered by the relevant knowledge community [41,42]. Such statements are built on scientific work, and there will always be a discussion as to what these statements are. The present article is to be seen as a contribution to this end for the risk field and science, by looking into its foundational literature and a selection of papers that have been particularly influential. Certainly, many papers have made strong contributions to the risk science, and it is not straightforward to point to just a few papers. However, looking at citation numbers and highlighting the technical literature, the following three papers clearly stand out: Kaplan and Garrick [50], Apostolakis [3] and Paté-Cornell [67]. The present article will focus on these papers. Many other papers deserve attention for their contributions, but the aim and scope of the present article is not to provide an all-inclusive review and discussion of foundational papers for the risk field and science but to highlight some selected ones. If we were to discuss the most influential scientific works in the risk field in general, reference should clearly be given to the early work on risk perception [e.g. (79)] and the social amplification of risk [e.g. (51)]. See also discussion in Greenberg et al. [36]. Describing the historical development of the risk field and science as such, would be a different type of paper. A much broader reference basis would then have been presented, covering scientific papers, professional reports as well as governmental policy documents. Examples include Farmer [25], NRC [62,63], Fischhoff et al. [26,27], Kuhlman [55], Griffith [37], Health Council of the Netherlands [44], HSE [45,46], Rechard [73,74], Renn [75] and Ale [1,2], with their many references. Studying these source documents provides a background for understanding these three papers, what they aim at accomplishing and what is the regulation and policy context.

The present paper has, however, a different aim and scope. New knowledge is sought developed by studying these three highly recognized and influential scientific papers from the 80s and 90s, addressing fundamental issues of risk analysis, reviewing their main contributions and comparing them with the risk science of today. The review and comparison allow for reflections on some of the key foundational challenges of the risk science, with a historical perspective, but also pointing to development trends and current discussions. The paper is concerned with scientific, foundational issues of risk analysis, with a main focus on the engineering environment, risk assessments and related conceptualization, characterization and use. With these restrictions, there should be little discussion about the importance of these three papers, acknowledging that there could be different views among risk scientists and professionals what are the most influential scientific papers.

These three papers are reviewed in Section 2. The focus is on issues such as the meaning of the concepts of risk and probability, different types of uncertainty, model uncertainty, the difference between uncertainty and probability, and how to quantify and describe risk.
Section 3 follows up these topics and discusses them in view of developments in the risk science and the current state-of-the-art. It is shown that many of the main points made in these three papers are still considered up-to-date risk knowledge. For example, Kaplan and Garrick [50] refer to an interpretation of probability to express uncertainties based on comparisons with lottery drawings, which is today considered a main pillar for understanding probabilities.

A main reference for the discussion of the current state-of-the-art of the risk science is documents produced by the Society for Risk Analysis (SRA), including a glossary and key principles [80,81,83,84]. These documents have been developed by a broad group of senior risk scientists, with input from members of the society. In general, the reference for what is state-of-the-art is the scientific risk literature. There is clearly an element of subjectivity in interpreting what are the most warranted statements of this literature, and this is acknowledged. The key is the argumentation put forward, which is open for critique and analysis, which in its turn can lead to further discussion and new risk science insights.

2. Review of the three papers

In the following, a brief review of the three papers is given. Foundational issues, which are considered interesting from a current risk science perspective, are highlighted.

2.1. The Kaplan and Garrick paper

The Kaplan and Garrick paper from 1981 has strongly influenced the way risk is conceptualised, in particular in engineering settings. It has about 3000 citations in Google Scholar. The aim of the paper is “to provide some suggestions and contributions toward a uniform conception of risk” (50). As the title of the paper indicates, its main focus is the quantification of risk, and most professionals in the risk field are familiar with the risk triplet introduced in this paper:

Risk is quantitatively expressed by the set of triplets \( (s_i, p_i, c_i) \), 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, \( i = 1, 2, \ldots N \). In risk assessment, this risk description is derived by answering the following three questions:

i) What can happen? (i.e., what can go wrong?)
ii) How likely is it that will happen?
iii) If it does happen, what are the consequences?

To take into account potential scenarios not identified, it is suggested that a category of ‘others’ is included. These scenarios represent potential surprises and are today commonly referred to as ‘black swan’ types of events [12,69,85], refer to discussion in Section 3.3.

A second level of risk characterisation is also presented, the so-called ‘probability of frequency’ approach. Here, ‘frequency’ is used to denote a frequentist probability, interpreted as the fraction of times the event would occur if the experiment could be repeated infinitely under similar conditions, and ‘probability’ is used as subjective probability, reflecting a degree of belief or a measure of confidence. Following Lindley [57], Kaplan and Garrick explain that a probability is to be interpreted by reference to a comparison standard, for example a lottery basket containing coupons numbered from 1 to 1000. Then, if a probability of, say, 0.002 is assigned for an event \( A \), it means that the assessor has the same degree of belief in \( A \) occurring as randomly drawing a coupon numbered 2 or less. Kaplan and Garrick show that the probability of frequency approach is in line with the Bayesian approach, starting from a probability model on the basis of frequentist probabilities, and using subjective probabilities, together with Bayes’ formula, to update degrees of belief about the frequentist probabilities. Kaplan and Garrick reject the idea that risk is probability times consequence. In line with the triplet definition, risk is to be considered probability and consequence. For short, we will refer to this risk perspective as \((C,P)\). However, when discussing qualitative perspectives on risk, Kaplan and Garrick refer to risk as uncertainty plus damage (to be understood as uncertainty ‘and’ damage). The point being made is that to refer to risk there must be some type of damage or loss that might be received.

Kaplan and Garrick also make some remarks concerning risk acceptance. It is concluded that it does not make sense to talk about risk in isolation. We have to look at the options, the costs, benefits and risks of each and find the overall best option. “The risk associated with that option is acceptable. All others are unacceptable” [50].

2.2. The Apostolakis paper

The Apostolakis paper addresses many fundamental risk science topics, including how to conceptualise and treat uncertainties in risk assessments. Apostolakis distinguishes between uncertainties in physical models and state-of-knowledge uncertainties about the parameters and assumptions of these models. Using Apostolakis’ terminology, the model is a function \( G \) of some parameters \( \varphi \), which can be written \( G(\varphi | M,H) \), where \( M \) is the set of model assumptions that define the model, and \( H \) is the entire body of knowledge and beliefs of the modeller. Both \( \varphi \) and \( M \) are subject to uncertainties, and it is suggested that subjective probabilities are used to express these. In a risk context, the focus is on events, and we are led to probability statements of the form \( P(A|H) \), where \( A \) is the event of interest.

As an example, consider a case where \( A \) is defined by “no accident occurring in the period considered”, and a Poisson model with parameter \( \lambda \) is introduced to reflect variation in the number of events occurring. Then \( \varphi = \lambda \), and an example of \( M \) is the assumption that the occurrence rate is constant in time. By establishing a subjective distribution over \((\lambda, M)\), Apostolakis shows how to compute \( P(A|H) \) using the rule of total probability. He also shows how to update the probabilities in view of new information and evidence, using the Bayesian updating machinery. Apostolakis underlines that probabilities, regardless of where they appear, are always measures of degrees of belief reflecting lack of knowledge (but he does not provide an interpretation of these probabilities beyond that). Frequencies and model parameters are something else. They are uncertain quantities and are assessed using subjective probabilities. Frequencies reflect variation (aleatory uncertainties).

Apostolakis also highlights the need for expert opinions, because safety assessments must deal with rare events. The challenge is how to process the judgements and combine them with observations and evidence. The Bayesian subjectivist approach is considered the solution which ensures coherence in the uncertainty judgements.

2.3. The Paté-Cornell paper

Paté-Cornell’s risk perspective is similar to \((C,P)\) and the risk triplet of Kaplan and Garrick. The paper highlights the difference between aleatory and epistemic uncertainties. The former is said “to stem from variability in known (or observable) populations and, therefore, represent randomness in samples”, whereas the latter are “those that come from basic lack of knowledge about fundamental phenomena” [67]. It is underlined that Bayesian theory allows the systematic integration of these uncertainties.

The paper’s main contribution relates to uncertainties in risk analyses and the taxonomy introduced with six treatment levels:

- Level 0: Identification of hazards
- Level 1: Worst-case approach – how bad is the worst?
- Level 2: Quasi-worst cases and plausible upper bounds. Different types of reasoning are used to define what is plausible, including historical data and established practice, and in some cases also probability.
- Level 3: Best estimates and central values.
- Level 4: Probabilistic risk assessment, single risk curve. As an example of such a curve, think about risk described by a subjective probability distribution for the number of fatalities.
- Level 5: Probabilistic risk analysis, multiple risk curves. The probability of frequency approach represents one way of establishing such curves. Then, subjective probability distributions are established for frequentist probabilities and related curves. The multiple risk curves can also reflect disagreement among experts or different modelling assumptions.

These levels are thoroughly discussed by Paté-Cornell. Strong argumentation against the use of the conservatism adopted in Levels 1 and 2 is provided. The key point made is that this approach “does not allow meaningful risk ranking because the degree of conservatism varies from case to case, making the results incomparable” ([67], p. 96).

In the paper, Paté-Cornell shows how these six levels correspond to regulatory practices in the US. She also underscores that risk assessments are only an information tool. Value judgements reflecting the decision maker’s preferences are also needed to make the right decisions. The paper discusses the need for seeing beyond expected utility theory to guide the decision-making. “Decision makers may need and/or ask for a full display of the magnitudes and the sources of uncertainties before making an informed judgment” [67]. According to the paper, rationality is more complex than simply maximizing expected utility.

The author also discusses the concept of risk aversion and its link to randomness and the lack of fundamental knowledge to support the probabilities. She states that “Conventional risk aversion implies that policies tend to focus on large losses that might occur at once. The danger is to neglect (by comparison) cumulative small losses that are collectively more costly, and to fail to protect sufficiently the potential victims of isolated accidents. Similarly, under ambiguity aversion, the danger is to spend large amounts of resources to mitigate risks that are poorly known and to neglect those that are known. The societal benefits are likely, in the end, to be smaller than they would have been given an attitude of neutrality towards ambiguity, but the policies may have satisfied another criterion of prudence” [67]. Here, ’ambiguity’ is used with the same meaning as ’epistemic uncertainties’. She also stresses that “it is important to separate the facts and their analysis from the value judgments”.

The work by Paté-Cornell [67] may be viewed as a reaction to the idea that risk can be managed effectively through financial metrics - a key reference being Morrall [59] who ranked a number of regulatory interventions in US by cost per life saved. Paté-Cornell [67] points to a “trend in the US regulatory and legal climate towards a soft cost-benefit analysis in environmental and health regulations. Risk assessment methods that were designed to provide a regulatory number through plausible upper bounds are no longer sufficient since they do not permit support of notions of cost-effectiveness”. To show the potential risk reduction benefits, uncertainty quantification is considered needed. Similar discussions where conducted in Europe in the 80s and also earlier, see for example Farmer [25], Health Council of the Netherlands [44] and HSE [45].

3. Discussion of the papers in view of the state-of-the-art

Based on the review in the previous section, the following topics will be further discussed in this section:

- How risk is conceptualised
- How risk is described
- How rational decision-making can be conducted in the face of risks

3.1. How risk is conceptualised

The three papers have a focus on risk description and characterisation. However, Kaplan and Garrick also touch upon the more fundamental question about what risk is. According to these authors, risk is qualitatively to be understood as uncertainties plus damage. This is an interesting comment, as it points to a broader definition of the risk concept than the quantitative triplet suggestion. There is a difference between uncertainties and risk. The point being made by Kaplan and Garrick is that, when faced with risk, there must be a potential for undesirable or negative consequences or outcomes (loss, damage). This is in line with the broad definition of the risk concept by SRA [80], which states that there needs to be at least one outcome that is considered undesirable or negative, when referring to risk. The neutral term ‘consequences’ is preferred ([80], as, when considering a future activity (for example, an investment or the operation of a process plant), there is also a potential for positive or desirable outcomes – as they typically are the driver for the activity to be realised. Information about the whole spectrum of consequences and outcomes is useful for supporting the decision-making, for example in relation to an investment. At least there is a need to include neutral outcomes, such as ‘as planned’, ‘no fatalities’, etc. Referring to the ‘consequences’ of the activity, the concept is applicable to all types of settings and applications.

Thus, Kaplan and Garrick qualitatively refer to risk as (C,U), when using a similar terminology to (C,P) introduced in Section 2.1. Risk has two dimensions: consequences and uncertainties. The consequences can have many dimensions (associated with life, health, environmental issues, economic values, etc.). Following Kaplan and Garrick, before the activity, we are not sure – we do not know – what C will be; we are in a state of uncertainty. In SRA [80], several definitions of risk are included, but they all express ideas similar to (C,U). It is a matter of taste whether one writes (U,C) or (C,U), but in the present article the latter is preferred, as the uncertainties are associated with the consequences.

Seeing uncertainties as a main component of risk goes back a long way. For example, Hardy [43] refers to risk as uncertainty, in regard to cost, loss or damage; see historical overview in Aven [6]. In a business context, it is common to think about risk as uncertainties, expressed through, for example, the variance. However, large uncertainty in itself is not the same as high risk; we have to look at both dimensions: the consequences and the associated uncertainties. If the damage is large, the risk is high, even with small uncertainties.

Uncertainty is not the same as probability. We use probability to represent or express uncertainty – it is an instrument adopted for this purpose. Other instruments exist, including imprecise probabilities, possibilistic measures and qualitative methods [22,30]. The use of probability to describe uncertainty has been thoroughly discussed in the literature (e.g. [30,58,64]), and will be further discussed in the following sections.

Distinguishing between the overall qualitative definition of risk and the way risk is described is important, because it stimulates reflections about what are suitable risk measurements and characterisation, as well as work to further develop and improve current approaches and methods. For example, when using probability to express uncertainties, we need to question the degree to which this tool is suitable for this purpose. If the point of departure is that risk is probability, such fundamental issues are less likely to be addressed. Although Kaplan and Garrick do not focus much on the qualitative risk definition in their paper, they present a framework that is based on this dichotomy, with risk as a concept comprising both consequences and uncertainties. It has inspired much of the work on risk conceptualisation in recent years (e.g. [6,80]).

3.2. How risk is described

For all three papers, the set of triplets $(s_i, p_i, c_i)$ is the basis for the risk description or characterisation. As the probability $p$ relates to the
scenarios as including the outcomes – or at least being linked to a specific consequence – and the term ‘consequences’ is to be seen as synonymous with ‘outcome’. The scenarios could, for example, reflect events leading to specific numbers of fatalities, as in event trees. Note that the use of the term ‘consequences’ in Section 3.1 is different: it covers the total of scenarios, events and outcomes of the activity considered. An alternative formulation of the risk description, which captures the same ideas as the risk triplet and is common today, is to refer to \((A',C',P)\), where \(A'\) denotes some specified events, \(C'\) the consequences of these events, and \(P\) the related probabilities. The ‘\(A'\) sign is used to distinguish between the actual quantities occurring \((A\text{ and } C)\) and the quantities defined in the risk assessment \((A',C')\). An event \(A\) may occur which is not captured by the set \(A'\) specified in the assessment. In a process plant a component of \(A'\) can refer to a gas leakage and \(C'\) to the number of fatalities, given the leakage. Probabilities are related to both \(A'\) and \(C'\). This is in fact what Kaplan and Garrick do in their Section 5.3, when they discuss uncertainties related to \(C\). Hence, it would be more accurate to write \((s_p\text{ and } p)\), instead of \((s_p\text{ and } c)\). In several places in their paper, it seems that they think about scenarios more as events of the type \(A\) they refer, for example, to the scenario ‘pipe break’. They also emphasise that the scenarios \(s\) are really categories of scenarios, as in ‘pipe break’, which “actually includes a whole category of different kinds and sizes of breaks that might be envisioned, each resulting in a slightly different damage” \((50),\ p.\ 13)\).

It is essential, as mentioned by Kaplan and Garrick, that the scenarios/events are “chosen so that they are mutually exclusive and the same event does not show up in more than one category”. Thus scenarios can be defined by any number of events. Often a simple event is sufficient, especially in the case of ‘pinch points’ as discussed by Garrick \([34]\).

### The probability of frequency approach, and evidence-based probabilities

In the above formulation, \(p\) and \(P\) refer to a subjective probability. This probability expresses the degree of belief of the assessor. The characterisation of risk is labelled the ‘first level’. The ‘second level’ is based on the probability of frequency approach. It presumes the existence of frequentist probabilities, and the use of subjective probabilities to express epistemic uncertainties about the ‘true’ value of the frequentist probabilities. The approach can be seen as consistent with the Bayesian set-up for analysing unknown quantities and incorporating new observations, starting from a probability model representing a family of frequentist probabilities, with some unknown parameters. The approach can lead to comprehensive analysis and be resource demanding, but it benefits from a solid theoretical foundation for coherent uncertainty judgments \([38,58]\). The probability of frequency approach has been extended by allowing for imprecise probabilities and non-probabilistic representations of uncertainty, see for example Helton et al. \([39,40]\) and Flage et al. \([30]\).

All three papers discuss carefully the distinction between the frequentist probabilities and the subjective probabilities, but it is interesting to note that Kaplan and Garrick \((50),\ p.\ 22)\) in fact dispute the subjectivity, by indicating that the subjective probabilities and related risk descriptions are ‘objective’, given the evidence. These authors acknowledge that any risk description is subjective, just as is probability itself: it depends on the knowledge of the assessors. However, they add, that, given the same totality of information, they must come to the same probabilities and thus agree on the risk description: the evidence determines the probabilities; they can be said to be evidence-based (see also \([49]\)). According to Kaplan \([49]\), true Bayesian uses probability in sense of degree of credibility (confidence) dictated by the evidence, using Bayes’ theorem - there is no personality in it, no ‘opinion’. Kaplan refers to Jaynes \([48]\):

> Probability theory is an extension of logic, which describes the inductive reasoning of an idealized being who represents degrees of plausibility by real numbers. The numerical value of any probability \((A/B)\) will in general depend not only on \(A\) and \(B\), but also on the entire background of other propositions that this being is taking into account. A probability assignment is ‘subjective’ in the sense that it describes a state of knowledge rather than any property of the “real” world; but is completely ‘objective’ in the sense that it is independent of the personality of the user; two beings faced with the same total background of knowledge must assign the same probabilities. -E.T. Jaynes

The argumentation leads us to the concept of logical probability as was first proposed by Keynes \([52]\). The basic idea is that there exists a number in \([0,1]\), denoted \(P(H/K)\), which in an objective way measures the degree of logical support that the evidence \(K\) give to the hypothesis \(H\) \([51]\). However, several scholars have argued that this type of probabilities cannot be justified \([8,17,18]\). Using logical probabilities it is not clear how to interpret a number (say) 0.3 compared to 0.4. Lindley \([58]\) writes:

> Some people have put forward the argument that the only reason two persons differ in their beliefs about an event is that they have different knowledge bases, and that if these bases were shared, the two people would have the same beliefs, and therefore the same probability. This would remove the personal element from probability and it would logically follow that with knowledge base \(K\) for an uncertain event \(E\), all would have the same uncertainty, and therefore the same probability \(P(E/K)\), called a logical probability. We do not share this view, partly because it is very difficult to say what is meant by two knowledge bases being the same. In particular it has proved impossible to say what is meant by being ignorant of an event, or having an empty knowledge base, and although special cases can be covered, the general concept of ignorance has not yielded to analysis. \((58),\ p.\ 44)\)

The present author cannot see that convincing arguments have been provided for generally justifying logical or evidence-based probabilities as referred to above. As thoroughly discussed by Lindley \([58]\) there is a subjective leap from the knowledge base to the probability assignment, even if the Bayesian theory is used to update the probabilities when new information is available. As Lindley discusses in the above quote, there is no objective way of determining a prior distribution in case of an empty knowledge base. The problem has led to the development of imprecise probabilities, as will be discussed below.

Nonetheless, the term ‘subjective probability’ is often avoided by many analysts as the word ‘subjective’ is associated with arbitrariness and lack of scientific rigour \([61]\). Although the probabilities are to varying degree based on evidence, they are subject to an objective logic for how to make coherent judgments about uncertainties. The subjectivity term gives also the impression that the alternative, the frequentist probabilities in contrast, are objective. However, as discussed by Paté-Cornell \([67]\), there are also subjective elements in forming these probabilities. Models allowing for infinite repetition of the situation considered need then to be formulated, and this requires subjective judgments. The issue of objective probabilities has been thoroughly discussed in the literature, see e.g. Bernardo and Smith \([16]\), p. 101), Singpurwalla \((78),\ p.\ 17)\), Aven and Reniers \([13]\). For unique events, like the occurrence of a serious terrorist attack the next 10 years in a country, or the occurrence of a major nuclear accident in the next 30 years, there is no alternative to the use of subjective probabilities. Repeating the situations to produce a frequentist probability does not make sense. Also as noted by Apostolakis \([3]\), subjective probability is the only way of expressing uncertainties as to whether a statement is true or not. Frequencies has no meaning.

### Imprecise probabilities

Acknowledging the subjectivity of these probabilities, considerable research has been conducted to develop alternatives. The use of imprecise probabilities (also referred to as interval probabilities) is
considered the most important one [30]. The idea is that the imprecise probabilities should better reflect the knowledge available. Such imprecision intervals are used a lot in the risk field but are not addressed by the three papers reviewed in the present article. For example, in risk matrices, such intervals are commonly used, with categories expressing 19 intervals such as ‘at least a probability of 0.95’. In the context of subjective probabilities, the interpretation is this [13,58]: Consider an urn comprising 100 balls, at least 95 of which are red. The exact number is not specified. The probability is comparable to – the uncertainty and degree of belief are the same as – randomly drawing a red ball out of the urn. If the number of balls is specified to be 95, the probability is interpreted as 0.95. The concept of imprecision is commonly misinterpreted as a uniform distribution. Using such a distribution, there is however no imprecision, all probabilities are precise numbers.

It is observed that, when referring to a subjective probability, Kaplan and Garrick use this type of urn interpretation and not one of the many others, originating from, for example, de Finetti [20], Ramsey [72] or Savage [76], which mix uncertainty and value judgements; see discussion in Aven and Reniers [13]. Consider the following example. A subjective probability, for the event A that a specific hypothesis is true, is assigned to be 0.95. Then, this probability can be interpreted as expressing that 0.95 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” (see e.g. [13,77]). Strong arguments can be provided for not using such an interpretation in a risk analysis context: the uncertainty assessment should not reflect the assigner’s attitude to money; see discussion in Lindley [58] and Aven and Reniers [13]. Through their recommendations to use the urn type of probability interpretation, Kaplan and Garrick have taken a stand on this issue and shown the way for the risk science.

The knowledge and assumptions that the probabilities are founded on

The use of imprecise probabilities ensure that the process from knowledge to probabilities becomes more objective. The intervals are more evidence-based than the specific assigned probabilities. However, this does not mean that the interval probabilities or the supporting knowledge become more objective as such. Knowledge is basically justified beliefs [80] and these beliefs can be more or less strong and even erroneousness. The evidence could, for example, be based on expert opinion, which could be poor or wrong. This aspect of risk is addressed by Apostolakis [3] when examining the assumptions of the model used in the risk assessment. Apostolakis outlines an approach for dealing with the issue and, although questions can be raised about the suitability of this approach (see below), the work is important, as it points to the need to discuss this type of uncertainties and risk. As such, the paper by Apostolakis is one of the pioneering works for treating assumptions in risk assessments. The importance of assumptions in risk assessments has also been thoroughly discussed by Paté-Cornell (e.g. [66,67]). Today, this challenge is a current issue of risk science, particularly in relation to evaluating the quality of risk assessments [84].

The approach by Apostolakis is illustrated by the use of the Poisson process model to represent events occurring. By using the law of total probability, the approach recommended produces the output probability \( P(A|H) \), where \( A \) is the event of interest and \( H \) the total knowledge of the assessor. The parameters \( \varphi \) and the set \( M \) of assumptions that define the model have been integrated out. This requires that probability distributions are established for \( \varphi \) and \( M \). Let \( M_1, M_2, \ldots, M_n \) represent the different model assumptions. Then, according to the approach presented, the assessor needs to assign probabilities \( P(M_i | H) \), expressing the probability that assumption \( i \) is true. For \( i = 1 \) in this Poisson case, we can think about the assumption that the occurrence rate is constant in time. It is, however, not clear what the other model assumptions should be. What type of deviations from \( M_1 \) should we include? We may for example restrict attention to an inhomogeneous Poisson process, but that means that other models are not considered.

Thus new types of assumptions need to be made. It may be thought that it is possible to establish a model so broad that the whole space of probability models is covered. In practice, this is not possible. The models are introduced to simplify the analysis and obtain focus on the issues that matter. Thus, there will always be a question about the quality and performance of the model. We are led to considerations of the reasonability of the assumptions made. As the assumptions can be viewed as justified beliefs for the assessment conducted, the critical issue is the strength of these beliefs. Other beliefs of the knowledge base also need to be considered. Such beliefs may, for example, be linked to the relevance of the data used in the analysis or the competence of the experts involved in the study. Considerable work has been conducted to establish suitable approaches for how to do this. See, for example, Flage and Aven [29] and Aven [11]. To evaluate the strength of the knowledge, it is suggested that issues, such as the following, are addressed:

- The reasonability of the assumptions made
- The amount and relevancy of data/information
- The degree of agreement among experts
- The degree to which the phenomena involved are understood and accurate models exist
- The degree to which the knowledge \( K \) has been thoroughly examined (for example, with respect to signals and unknown known; i.e. others, but not the analysis group, have the knowledge).

Another related system for assessing the knowledge strength is the so-called NUSAP system (NUSAP: Numeral, Unit, Spread, Assessment, and Pedigree) [32,33,53,54,56,86,87].

These types of evaluations are qualitative, and the risk description then captures probability or imprecise probabilities related to the quantities of interest, as well as strength-of-knowledge judgements. In addition, the background knowledge of the assessment is seen as a part of the risk description. Formally, and generalising \((A', C', P)\), the risk description can be written \((A', C', Q, K)\), where \( Q \) is a description of uncertainty and \( K \) is the knowledge supporting \( Q \). Above, \( Q = (P, SoK) \), where the SoK is a judgement of the Strength of the Knowledge \( K \).

All three papers discussed in the present article advocate a probabilistic and quantitative approach. The knowledge \( K \) of the analysis is acknowledged as important, but without being evaluated. The above discussion has aimed at showing that SoK judgements are needed. Comprehensive probability curves can be established as argued for in, for example, the probability of frequency approach, but, for the evaluation of the numbers produced, whether they are based on a strong knowledge or not is critical.

The use of \( SoK \) type of judgements means the acknowledgement of qualitative assessments, in addition to the probabilities. The problems with qualitative analyses are well-known. Rankings and comparisons are more difficult to conduct, and the scores and classifications are often strongly dependent on the assessors. However, it can be argued that the alternative is worse: represent and express risk in a mathematical, quantitative way, which ignores important aspects of risk and introduces strong elements of arbitrariness. Attempts have been made to show that it is possible to integrate out the assumptions and key aspects of the background knowledge, as in relation to the model uncertainty case of Apostolakis, see also discussion by Mosleh and Bier [60], but, as indicated above, such integration leads to new types of assumptions and beliefs. In general, it is neither possible nor desirable to present the final result of a risk assessment as a \( P(A) \) unconditional on ‘everything’. Any model used means that there is some background knowledge, some beliefs that the assessment is based on. In practice it is usually preferable to highlight and discuss the conditional probabilities \( P(A|Z = z, H) \), where \( Z \) is an unknown quantity reflecting for example the different model assumptions. Different approaches exist for this purpose. One approach is to use sensitivity analysis and see how the results are affected by changes in \( z \). Another approach is to include the
assumptions in the background knowledge \( H (K) \) and include a strength-of-knowledge (SoK) assessment, as outlined above. We could also perform an ‘assumption deviation risk’ assessment ([7]), in which deviations from the assumption made are analysed. The study is qualitative, highlighting potential deviations from the assumptions made, the implications of the deviations, judgements of probability, and related strength of knowledge.

Finally in this section, a comment on the use of conservatism in risk assessment. Paté-Cornell [67], see also Paté-Cornell [66], provides clear guidance on the issue, as briefly discussed in Section 2.3. The argumentation and recommendations provided stand out equally strongly today as in 1996; see the recent discussion by Aven [9].

### 3.3. How rational decision-making can be conducted in the face of risks

All three papers studied in this article emphasise that risk assessment supports decision-making, it does not prescribe what is the best decision. There are concerns other than risk that need to be taken into account when making decisions in situations where risk is an issue. Some type of decision analysis is in place. A backbone of decision analysis is the expected subjective utility theory. It is discussed in all three papers. The theoretical basis and its usefulness for guiding risk decisions is acknowledged, but also its limitations. Paté-Cornell provides a thorough discussion of the topic. She points to the fact that, according to the expected utility theory, a rational decision-maker is assumed to be indifferent to the level of uncertainty (ambiguity) beyond its effect on the outcome subjective probability distribution. Whether the probability is founded on a strong or weak knowledge basis is not relevant, refer to the Ellsberg paradox [23,24]. She refers to “firm” and “soft” probabilities, respectively. However, ignoring this aspect of knowledge strength can be challenged from both an empirical and a normative perspective, as discussed by Paté-Cornell [67] and Fishburn [28], Gilboa and Marinacci [35], Aven [5], pp. 120–122), see also Paté-Cornell and Fischbeck [71], Davis and Paté-Cornell [19] and Paté-Cornell [68]. Paté-Cornell has made an important contribution on this issue, by clarifying the difference between risk analysis and decision analysis. Her analysis is to a large extent also state-of-the-art today.

Kaplan and Garrick stress that much information is lost in the expectation operation. Rather than computing an explicit utility function, they question whether it would not be better to use a broad risk characterisation, as outlined in their paper using risk curves, and ask: “Design B will cost, say, \( \Delta \) dollars more than A. Is it worth it to you?” ([50], p. 23). Kaplan and Garrick link this discussion to the question about acceptable risk. They state: “... Risk cannot be spoken of as acceptable or not in isolation, but only in combination with the costs and benefits that are attendant to that risk” ([50], p. 24). Risk science today expresses the same thing (see e.g. [14,81,84]), although risk management practice is not always in line with these ideas (see e.g. [10]).

Apostolakis focuses on the problem of the expected utility theory in relation to major societal decisions that involve many decision-makers (many stakeholders). In such settings, it breaks down. As we know, the theory works for a single decision-maker but not for two or more decision-makers.

Nonetheless, we find risk scholars that advocate a ‘rational’ approach based on utility theory, supported by mathematical and probabilistic modelling, subjective probabilities (often interpreted in accordance with, for example, Ramsey [72] and Savage [76]). This perspective, and the work performed within its scope, add knowledge to the risk science, but it has limitations, as discussed in the present article. It is argued that is necessary to see beyond the quantitative perspective, to adequately conduct risk analysis. A key problem is that the transformation from the knowledge and evidence to the subjective probabilities is often difficult to justify. Another challenge is the fact that this knowledge can be more or less strong, and even erroneous.

The three papers studied in this article to a large extent support a quantitative approach, although they point to the need to address assumptions and the background knowledge of the assessments. With reference to the criteria adopted in the present paper (refer to the SRA guidance documents), the knowledge aspects of risk can be viewed as stronger highlighted today compared to what were the standards in earlier works. For example, when subjective probabilities are used to measure or describe uncertainties and risk today, recommendations are also given for including judgements of the strength of the knowledge supporting these. This cannot be done mathematically and is not covered by the quantitative theory referred to. Knowledge were also highlighted in earlier works on risk, see for example discussions in Otway and Peltu [65], but the risk conceptualization and characterisation were to large extent based on probability judgments. Incorporating knowledge and its strength as basic elements of the risk characterisation, the awareness of these aspects of risk is increased and changes in knowledge are more easily reflected in the risk characterisations.

Also, potential for surprises is highly relevant for risk management, but it is not an aspect of the mathematical perspective referred to above. In risk science today, it is a major area for research and development. The topic of surprise and the unforeseen is not discussed in detail in the three papers, but Kaplan and Garrick make some interesting points in relation to the issue when arguing for including scenario categories of the type ‘others’ to ensure completeness and reflect potential surprises (black swans types of events). As discussed in the literature such events represent a challenge in risk management not only because of limits of imaginability, but also unwillingness to think out of the box and anticipate potential problems [70].

Through their works, the authors of these three papers have strongly highlighted the need for distinguishing between risk assessment (including risk conceptualisation and characterisation) on the one hand, and risk management on the other. Risk assessment informs decision makers [4]. In the 80s and 90s also other perspectives were promoted. For example scholars founded on cultural theory and constructivism, argued that risk is the same as risk perception [21,47,88]. According to Beck ([15], p. 55): “Because risks are risks in knowledge, perceptions of risks and risk are not different things, but one and the same.” Researchers as Kaplan, Garrick, Apostolakis and Paté-Cornell contributed to developing concepts, theories and frameworks making it possible to separate risk and risk assessment from risk perception and risk management.

### 4. Conclusions

The findings in this article are summarised in Table 1.

The overall conclusion is that these three papers have contributed greatly to the development of the risk science, and they still to a large extent represent state-of-the-art. These papers have addressed important foundational issues and provided guidance that have strongly influenced risk analysis and management. As highlighted in the introduction section, there is a huge literature on the foundations of risk analysis and management, and the three papers need to be seen as contributions along with many other papers and works. A considerable number of developments have also been made on the fundamentals of risk science the last 20–30 years - rethinking, refining and sharpening existing beliefs and argumentation.

As mentioned in the introduction section, documents from SRA and related works constitute the main study reference for the present analysis. As also highlighted in Table 1, common practice is to varying degree in line with the most recent theoretical advancements. For example, in industry traditional perspectives on risk based on consequences and probability are to large extent prevailing. Scientific developments require discussion and scrutiny of all aspects of the scientific process, and the author of the present paper hopes that the present paper can stimulate a discussion on what is in fact state-of-the-art of the risk field and science, also challenging the perspectives and argumentation provided in this article.
Table 1
Summary of main conclusion of the present article, based on the criteria and argumentation used in the paper.

<table>
<thead>
<tr>
<th>Message from the three papers</th>
<th>Risk science of today</th>
<th>Match between these two</th>
<th>Has this message influenced the development in risk science?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk is uncertainty + damage (KG)</td>
<td>Risk = Consequences + Uncertainty</td>
<td>Strong</td>
<td>Yes, but common risk analysis practice is still very much based on the risk concept being a function of probabilities</td>
</tr>
<tr>
<td>Risk described by the set of triplets, (s,p,c) (KG)</td>
<td>The set of triplet, (s,p,c), represents an example of a risk description. More generally, risk can be described by (C, Q, R) or (A, C, Q, R).</td>
<td>The current description extends the traditional triplet approach</td>
<td>Yes, the current one has developed from the traditional triplet approach</td>
</tr>
<tr>
<td>The probability of frequency approach (KG, A, PC)</td>
<td>This approach is still a cornerstone in quantitative risk analysis. The need for further knowledge considerations is acknowledged</td>
<td>Strong, but current risk science highlights the need for additional qualitative analysis</td>
<td>Yes, these works provide a foundation for the use of the probability of frequency approach</td>
</tr>
<tr>
<td>A subjective probability is to be interpreted by reference to an urn type of comparison (KG)</td>
<td>The same</td>
<td>An importance difference</td>
<td>Yes, but this interpretation is not broadly used in practice</td>
</tr>
<tr>
<td>A subjective probability can be viewed as an evidence-based probability (KG)</td>
<td>The same</td>
<td>Yes, the reference to aleatory (stochastic) uncertainties is confusing, as it represents variation and not uncertainties</td>
<td></td>
</tr>
<tr>
<td>There is only one type of uncertainty (epistemic uncertainties) (KG, A, PC)</td>
<td>The same</td>
<td>Yes, but alternative approaches have also been developed</td>
<td></td>
</tr>
<tr>
<td>Uncertainties about (model) assumption should be integrated out (A)</td>
<td>A more nuanced perspective is adopted, also including alternative approaches for assessing the risk or importance of the assumptions made</td>
<td>Some differences in the thinking</td>
<td>Yes, but the approach is still often used in practice</td>
</tr>
<tr>
<td>Conservatism should be avoided (PC)</td>
<td>The same</td>
<td>Yes, but current practice is still often risk-based (the analysis prescribes what to do)</td>
<td></td>
</tr>
<tr>
<td>Risk assessment informs decision-makers</td>
<td>The same, but the risk characterisations are broader than probability curves and numbers</td>
<td>The same, but the risk characterisations are broader than probability curves and numbers</td>
<td>Yes, but current practice is still often risk-based (the analysis prescribes what to do)</td>
</tr>
<tr>
<td>Decision analysis and expected utility theory in particular is a useful tool for risk decision-making in many cases, but have some limitations (A, PC)</td>
<td>The same</td>
<td>The same</td>
<td></td>
</tr>
</tbody>
</table>

KG: Kaplan and Garrick [50], A: Apostolakis [3] and PC: Paté-Cornell [67].

The three papers should all be honoured as major contributions within risk science, and their influence on both research and practice is notable. Yet, we need to acknowledge that risk analysis practice does not fully reflect the basic insights and recommendations made in these papers, for example on how to interpret a probability. Continuous and systematic work is therefore needed to bring out the risk science messages and improve the practice of risk analysis concepts, principles, approaches, methods and models.

A term like probability does not have the same meaning and interpretation everywhere and at different points in time, and differences can in some cases be seen more as a reflection of the context than a sign of fundamental deficiencies in the understanding of the users or the underlying science – some issues just do not get resolved. However, it is an aim of the scientific work to obtain clarity and precision on terminology. The present paper argues that the issue about probability is to large extent about science and precision, when this term is used in a professional risk assessment and management context. There are some basics. A frequentist probability has a definition and interpretation, so do subjective probabilities. Most issues can be resolved, but if they cannot, it is essential to clarify what are the problems and what are the different stands.

Acknowledgments

The author is grateful to four reviewers for their useful comments and suggestions to earlier versions of this paper.

References

[22] Dubois D. Representation, propagation and decision issues in risk analysis under


