Philos. Technol. (2011) 24:151–168
DOI 10.1007/s13347-010-0003-6

RESEARCH ARTICLE

# Should Probabilistic Design Replace Safety Factors?

**Neelke Doorn · Sven Ove Hansson**

Received: 8 June 2010 / Accepted: 7 September 2010 / Published online: 28 September 2010


**Abstract** Safety is a concern in almost all branches of engineering. Whereas safety was traditionally introduced by applying safety factors or margins to the calculated maximum load, this approach is increasingly replaced with probabilistic risk assessment (PRA) as a tool for dimensioning safety measures. In this paper, the two approaches are compared in terms of what they aim at and what they can, in fact, achieve. The outcome of this comparison suggests that the two approaches should be seen as complementary rather than mutually exclusive. PRA is particularly useful for priority setting and for the effect evaluation of safety measures; however, in most applications, uncertainties prevent PRA from providing an objective probability of failure or value of damage. Safety factors are indispensible for dealing with dangers that cannot be assigned meaningful probabilities.

**Keywords** Safety factor · Risk · Uncertainty · Vulnerability · Probabilistic design · PRA

## 1 Introduction

Safety is a concern in almost all branches of engineering. In structural engineering, for example, builders add extra strength in order to ensure the safety of their structures. This safety margin serves to protect individuals and society from the consequences of failure. The practice of adding extra strength to a construction can be traced back at least to antiquity; however, it is only in the nineteenth century that

N. Doorn (✉)
Department of Technology, Policy and Management,
Delft University of Technology, PO Box 5015, 2600 GA Delft, The Netherlands
e-mail: N.Doorn@tudelft.nl

S. O. Hansson
Department of Philosophy and the History of Technology, Royal Institute of Technology,
Teknikringen 78 B, 100 44 Stockholm, Sweden
e-mail: soh@kth.se

numerical safety factors were introduced to determine the sizes of these safety margins (Randall 1976). A safety factor is a ratio between the maximal load not leading to failure and the maximal load for which the construction is intended. In the last few decades, attempts have been made to replace safety factors by probabilistic calculations. Probabilistic risk assessment (PRA, also probabilistic safety assessment) is now increasingly used as a tool for dimensioning safety measures (Bedford and Cooke 2001; Vose 2000). With these methods, safety margins are calibrated to achieve a certain, sufficiently low, calculated probability of failure.

Probabilistic methodology has the obvious advantage of directly addressing the crucial goal of safety engineering, namely to reduce the probability of accidents and other adverse events. Safety factors address this goal only in a more indirect way. Proponents of probabilistic methods have been keen to point out that "the safety of the building constructions is a matter of calculating probabilities" (Mayer 1926; quoted in Elishakoff 2004: p. 1) and that "probability theory provides a more accurate engineering representation of reality" (Cornell 1969: p. 974); however, these general considerations are not sufficient to prove that probabilistic methodology is superior to the safety factor approach. The crucial issue is whether its use in practical engineering design will indeed promote safety more efficiently than the use of safety factors. Proponents of safety factors have argued that replacing them by probabilistic approaches would be "a dangerous lapse" (Knoll 1976), and it has been argued that probabilistic analysis tends to neglect some of the safety–critical issues that are traditionally covered by safety factors (Clausen et al. 2006; Möller and Hansson 2008).

Decision making regarding safety and risks is philosophically relevant for several reasons. First, different approaches to risk analysis represent different philosophical theories. Standard risk–benefit analysis, for example, is similar to classical utilitarianism in its disregard for persons. Drawbacks of classical utilitarianism pertain to risk–benefit as well (Hansson 2007). Secondly, although the establishment of risk exposure limits and other regulations is often presented as "scientific" and "value-free," risk-related decisions often contain value-based judgments on what risks to accept. It typically requires philosophical expertise to uncover these hidden value assumptions in decision making on technological risks (Hansson 2009c). Thirdly, risk-related decision making requires comparisons of values that are difficult, often seemingly impossible, to compare to each other: losses in human life, disabilities and diseases, the loss of animal species, etc. This issue of value incommensurability is a recurrent problem in philosophy in general and in the philosophy of engineering design in particular (Chang 1997; Van de Poel 2009).

In this paper, we will compare the two approaches in terms of what they aim at and what they can in fact achieve. In Section 2, some concepts that are crucial for the discussion will be clarified. We introduce safety factors in Section 3 and probabilistic analysis in Section 4. This is followed by overviews of the arguments for probabilistic design (Section 5) and safety factors (Section 6). In Section 7 we sum up and conclude.

## 2 Basic Terminology

The distinction between *risk* and *uncertainty* dates back to work in the early twentieth century by the economists JM Keynes and FH Knight (Arrow 1951;

Keynes 1921; Knight 1935[1921]). Knight pointed out that "[t]he term 'risk', as loosely used in everyday speech and in economic discussion, really covers two things which, functionally at least, in their causal relations to the phenomena of economic organization, are categorically different." In some cases, "risk" means "a quantity susceptible of measurement," in other cases "something distinctly not of this character." He proposed to reserve the term "uncertainty" for cases of the non-quantifiable type, and the term "risk" for the quantifiable cases (Knight 1935[1921]: pp. 19–20).

This terminological reform has spread to other disciplines, including engineering and it is now commonly assumed in most scientific and engineering contexts that "risk" refers to something that can be assigned a probability, whereas "uncertainty" may be difficult or impossible to quantify. It should be noted, however, that in everyday language, "risk" is often used without reference to probability. It should also be observed that although uncertainty and risk are commonly defined as two mutually exclusive concepts, it is in practice common to use "uncertainty" in lieu of "risk or uncertainty." Then "uncertainty" is used as a general term for lack of knowledge (whether probabilistic or not), and risk is a special form of uncertainty, characterized by the availability of a meaningful probability estimate. In what follows, we will adhere to this practice and use "uncertainty" in the broad sense that covers (probabilizable) risk.

Even in cases when the plausibility of a danger can be meaningfully summarized in a probability estimate, there may yet remain significant uncertainties about the accuracy of this estimate. In fact, only very rarely are probabilities known with certainty. Therefore, even if a decision problem is treated as a decision "under risk," then this does not mean that the decision in question is made under conditions of completely known probabilities. Rather, it means that a choice has been made to simplify the description of this decision problem by treating it as a case of known probabilities. This is practically important in safety engineering. Some of the probability estimates used in risk calculations are quite uncertain. Such *uncertainty about probabilities* should be taken into account when probabilistic analyses are used for decision-making purposes.

The unclear role of scientists taking part in risk policy decisions led in the 1970s to increasing awareness of the distinction between scientific assessments and policy decisions based on these assessments. This resulted in a standard view on the risk decision process that distinguishes strictly between its scientific and policy-making parts. This view was expressed in a 1983 report by the American National Academy of Sciences (National Research Council 1983). The decision procedure is divided into two distinct parts to be performed consecutively. The first of these, commonly called "risk assessment," is a scientific undertaking. It consists of collecting and assessing the relevant information and on this base characterizing the nature and magnitude of the risk. The second procedure is called "risk management." Contrary to risk assessment, this is not a scientific undertaking. Its starting point is the outcome of risk assessment, which it combines with economical and technological information pertaining to various ways of reducing or eliminating the risk in question, and also with political and social information. Based on this, a decision is made on what measures—if any—should be taken to reduce the risk. In order to protect risk assessments from being manipulated to meet predetermined policy

objectives, it was proposed to separate risk assessment organizationally from risk management; however, in spite of many official documents promoting this division between risk assessment and risk management, it is more often violated than obeyed.

In the public sector, it is often applied in areas where risk decisions are controversial and publicly debated, such as the control of chemicals. In areas where risk management is conceived as uncontroversial, such as the setting of technical standards for building constructions, risk management and risk assessment are usually performed by the same persons and in the same expert committees (Clausen and Hansson 2007). In the private sector, efficiency and coordination are prioritized, and therefore the management systems used in this sector tend not to be easily combined with a high degree of independence for a group of experts, such as risk assessors. We are not aware of any case of strict separation between risk assessment and risk management in the private sector.

Both the safety factors approach and the probabilistic approach to engineering design employ numerical limits to draw the line between sufficiently safe and too unsafe designs. These limits are different in nature, but in both cases they are, in practice, determined within the community of safety experts in a process that does not distinguish between risk assessment and risk management.

## 3 The Safety Factor Approach

The use of safety factors is a well-established method in the various branches of structural engineering. A safety factor is typically intended to protect against a particular integrity-threatening mechanism, and different safety factors can be used against different such mechanisms. Most commonly, a safety factor is defined as the ratio between a measure of the maximum load not leading to failure and a corresponding measure of the applied load. In some cases it may instead be defined as the ratio between the estimated design life and the actual service life. In addition to safety factors, the related concept of safety margin is used in several contexts. Safety margins are additive rather than multiplicative; typically a safety margin in structural engineering is then defined as capacity minus load.

It is generally agreed in the literature on structural engineering that safety factors are intended to compensate for five major types of sources of failure:

(1) Higher loads than those foreseen,
(2) Worse properties of the material than foreseen,
(3) Imperfect theory of the failure mechanism in question,
(4) Possibly unknown failure mechanisms, and
(5) Human error (e.g., in design) (Knoll 1976; Moses 1997).

The first two of these can in general be classified as variabilities, that is, they refer to the variability of empirical indicators of the propensity for failure. They are therefore accessible to probabilistic assessment (although these assessments may be more or less uncertain). In the technical terminology that distinguishes between risk and uncertainty they can be subsumed under the category of risk. The last three failure types refer to eventualities that are difficult or impossible to

represent in probabilistic terms, and therefore belong to the category of (non-probabilizable) uncertainty.

In order to provide adequate protection, a system of safety factors will have to consider all the integrity-threatening mechanisms that can occur. For instance, one safety factor may be required for resistance to plastic deformation and another one for fatigue resistance. Also different loading situations may be taken into account, such as permanent load ("dead load"; that is, the weight of the building) and variable load ("live load"; that is, the loads produced by the use and occupancy of the building); the safety factor of the latter being higher because of higher variabilities. Similarly, components with widely varying material properties (e.g., brittle materials such as glass) are subject to higher safety factors than components of less variable materials (e.g., steel and metallic materials). Geographic properties may be taken into account by applying additional wind and earthquake factors. Design criteria employing safety factors can be found in numerous engineering standards and building codes.

## 4 Probabilistic Risk (Safety) Assessment

Modern probabilistic risk assessment has largely been developed in the nuclear industry. Although the engineers designing nuclear reactors in the 1950s and 1960s aimed at keeping the probability of accidents very low, they did not have any means to estimate these probabilities. In the late 1960s and early 1970s, methodology was developed to make such estimates. The first PRA of a nuclear reactor was the Rasmussen report (WASH-1400) that was published in 1975 (Michal 2000; Rasmussen 1975). The basic methodology used in this report is still used, with various improvements, both in the nuclear industry and in an increasing number of other industries as a means to calculate and efficiently reduce the probability of accidents.

A PRA usually begins with the identification of the undesirable events to be covered by the analysis. In a nuclear reactor, most of these will be various types of accidents leading to core damage ("meltdown") or to the release of radioactivity. The next step is to identify for each of these adverse events the accident sequences that may lead to its occurrence. Typically, several such sequences will be identified for each event. Each sequence is a chain containing events such as mechanical equipment failure, software failure, lacking or faulty maintenance, mistakes in the control room, etc. Next, the probability of each of these accident sequences is calculated, based on the probability of each event in the sequence. Some of these probabilities can be based on empirical evidence, but others have to be based on expert assessments. The final step in the PRA consists in combining all this information into an overall assessment. In the early days of PRA, the overall assessment often included a total probability of a major accident and/or a statistical expectation value for the number of deaths per year resulting from accidents in the plant. Today, most PRA specialists in the nuclear industry consider such overall calculations too uncertain. Instead, their focus is on using analysis of accident sequences to identify weaknesses in the safety system. According to one leading expert, the final step in a PRA

> ... is to rank the accident sequences according to their probability of occurrence. This is done because risk must be managed; knowing the major

> contributors to each undesirable event that was defined in the first step is a major element of risk management. Also ranked are the SSCs—systems, structures, and components—according to their contribution to the undesirable event (Michal 2000: pp. 27–28).

The same basic methodology can be used in structural engineering. In the early 2000s, the Joint Committee on Structural Safety (JCSS) developed a Probabilistic Model Code for full probabilistic design. The code was intended as the operational part of national and transnational building codes that allow for probabilistic design but do not give any detailed guidance (Vrouwenvelder 2002). Contrary to the nuclear industry, structural engineering uses PRA more to dimension individual components than to identify and analyze different accident sequences (JCSS 2001; Melchers 2002). This difference depends in part on the complicated redistribution of the load effects after each component failure, which makes it difficult to predict the behavior of the system itself (Ditlevsen and Madsen 2007[1996]); however, attempts are made to broaden the scope of PRA in structural engineering and to view building structures as parts of wider infrastructure systems (Blockley and Godfrey 2000; Melchers 2007).

Within structural engineering practice, most PRA specialists defend a Bayesian interpretation of failure probabilities, in which "probabilities are considered as the best possible expression of the degree of belief in the occurrence of a certain event" and not as "unbiased predictors of occurrence frequencies that can be observed in practice" (JCSS 2001: p. 60; see also Ditlevsen and Madsen 2007[1996]).

## 5 Arguments for Using Probabilistic Design Methods

In the literature, we have found four arguments that are used to support the view that probabilistic approaches to design are preferable to deterministic ones such as safety factors. These are the possibility of economic optimization, improved precision, integral approach, and fitness for policy making (risk management). In this section, we will discuss each of these arguments.

### 5.1 Economic Optimization

The first, and probably most important, argument used in favor of probabilistic methods is that their output can be used as an input into economic optimization. Some argue that economic optimization of risk management measures is in fact the main objective of probabilistic risk analysis (Guikema and Paté-Cornell 2002). Traditional approaches in safety engineering, such as safety factors, provide regulatory bounds that may sometimes be overly conservative (Chapman et al. 1998). There is, for instance, no way to translate the difference between using the safety factor 2.0 and the safety factor 3.0 in the design of a bridge into a quantifiable effect on safety. Without a quantifiable effect (such as reduced expected number of fatalities) it is not possible to calculate the marginal cost of risk reduction, and therefore economic optimization of design is not possible. In contrast, a PRA that provides accident probabilities as outcomes makes it possible to calculate the

expected gains from a safer design. This is what is needed for an optimization of the trade-off between risks and benefits (Moses 1997; Paté-Cornell 1996).

Such optimization may involve trade-offs against other factors than money. A risk can, for instance, be weighed against other risks that are brought about by countermeasures against the first risk (Graham and Wiener 1995). It is also common for overdesign to have a price in terms of excess usage of energy and other natural resources. Accident probabilities obtained in a PRA can be used as inputs into a risk–benefit analysis (RBA) or cost–benefit analysis (CBA) in which different types of advantages and disadvantages are taken into account (Rackwitz 2004). In such an analysis, the different types of outcomes, including bodily injuries and loss of lives, are all assigned monetary values in order to achieve comparability. Madsen et al. (1986) warn that in the case of human lives, the trade-off between costs and risks cannot be reduced to technical quantification but should be supplemented with a variety of approaches in the practical selection of socially accepted safety levels (for a similar plea, see Hampshire 1972). Although the CBA methodology has many problems (Hansson 2007) it is widely used as a basis for economic decision making, in particular, in the public sector (Nathwani et al. 1997; Rackwitz 2001).

The major problem with this argument for PRA is that it puts very high demands on the probabilities that are outputs of the analysis. As we saw above, PRA analysts in the nuclear industry have largely given up the original idea that the outputs of probabilistic analysis of event sequences in nuclear reactors could be interpreted as reasonably accurate probabilities of various types of accidents. Instead, these calculations are used primarily to compare different event sequences and to identify critical elements in these sequences. If the outcome of PRA is interpreted in this latter sense, then the use of these probabilities as inputs into RBA or CBA is not justified.

The question that must be asked, then, is whether the outputs of PRAs in non-nuclear contexts, such as civil engineering, are accurate enough to be used as inputs into economic analysis. The answer to this question seems to differ between different contexts. Some relatively small and standardized infrastructure projects have effects that can be described fairly accurately in probabilistic terms. This applies for instance to some safety measures in road traffic such as central barriers on highways (Mak et al. 1998) or pedestrian crosswalks at intersections (Zegeer et al. 2006), for which the expected number of saved lives can be estimated with reasonable accuracy and weighed against the economic costs. On the other hand, in case of larger and more complex infrastructure projects, the probabilistic quantification of the effects of safety measures is generally not considered accurate enough to be used as direct input into economic analysis. The safety of gravity dams, for example, is largely dependent on seismic activity and how the structure responds to it. Both can, at most, be quantified artificially and roughly, making it difficult to provide accurate accident probabilities (Abbas and Manohar 2002). In cases like this, it is therefore recommended to develop a robust structural design rather than an economically optimized one (Takewaki 2005). Similar problems are faced in the design of other large infrastructure projects, such as flood defense structures and offshore facilities.

In summary, the argument that PRA provides means for economic optimization is not valid for PRA in general but only for those PRAs that provide probability estimates that are well calibrated with actual frequencies.

### 5.2 Improved Precision

The second argument in favor of probabilistic approaches states that probabilistic analysis is able to provide a more precise description of the design parameters. This argument is based on the presumed nature of risks and uncertainties. Savchuck (1992), for example, argues that contrary to traditional design approaches, the variables in probabilistic design methods are assumed to be random, which corresponds to the nature of the real states. Similarly, Thoft-Christensen and Baker state in the preface of their classical textbook that "most loads and other structural design parameters are rarely known with certainty and should be regarded as random variables or stochastic processes" (Thoft-Christensen and Baker 1982); however, this argument in favor of probabilistic modeling is weak for three reasons. First, describing the full stochastic behavior of a particular variable is not necessarily the best way of depicting it for all purposes. For example, for most purposes it would be of little worth to describe the behavior of a gas in terms of the stochastic properties of individual gas molecules.

Secondly, due to lack of data, probabilistic models may not always have the qualities needed for the alleged increase in precision to take place. The empirical basis of probabilistic models has to rely on events that are common enough to have given rise to data about their occurrence; however, in risk analysis the probabilities of very uncommon events may be the most important ones. Often such probabilities are inferred from models that are based on more common events. Although the central parts of the statistical distributions used in these models are fairly well-known, their tails can only be inferred under assumptions about the mathematical structure of the distribution that lack direct empirical evidence. This is the so-called *distribution arbitrariness* (Ditlevsen 1994; Harris and Soms 1983). Extreme value analysis often involves extrapolation to values beyond the largest or smallest observed value in order to assign probabilities to extreme events. Expert judgments (Slijkhuis et al. 1999) and boot-strapping techniques (Caers and Maes 1998) are used to reduce the uncertainty of the tail-based estimates; however, boot-strapping techniques still require sufficiently long data records and a careful analysis of the influence of data sampling uncertainties (Van Noortwijk and Van Gelder 1998).

Thirdly, many PRAs refer to situations that are influenced by uncertainties that are difficult or impossible to quantify. This can be dealt with either by excluding such uncertainties from consideration or by assigning values to them that may not be much better than guesses. Even if such values are precise in the sense of being exact, they are seldom accurate.

Serious objections have been raised against the introduction of such unreliable probabilities into probabilistic models. When uncertainty is introduced quantitatively in the probabilistic models, human knowledge becomes part of the system. In practice, this can mean that the probability of failure can be "improved by [...] increasing our knowledge" (Vrijling 2001: p. 340). If we are interested in the probability that a technological system will function successfully for a specific period of time, then this is a rather awkward conclusion. When striving to reduce the probability of system failure we are focusing on a property of the system that does not depend on the available knowledge but on the physical properties of the system itself. This is not a mere theoretical objection. The quantitative inclusion of

epistemic uncertainty in a probabilistic analysis may lead to suboptimal risk mitigation—contrary to the common assumption that the use of probabilistic approaches promotes efficiency in risk reduction.

In sum, the presumed precision of probabilistic methods in assessing uncertainty is a rather weak argument. High precision can be misleading, and even dangerously so, if the accuracy is low.

### 5.3 Integrated System Approach

A third argument in favor of probabilistic approaches is its presumed ability to provide a more integrated assessment of the safety of the full system. This claim comes in three forms.

First, some argue that by providing an integrated account of the full system, probabilistic approaches are able to provide a higher safety level, for instance in cases when several elements are cross-correlated (Allen 1981; Kuijper and Vrijling 1998). Kuijper and Vrijling (1998) give the example of a sea dike that protects a polder. If the polder is adjoined by a river as well, the resulting safety level is lower than if there is no river (and hence, no risk of flooding from the river).

Although this argument is often made, its theoretical underpinning is rather weak. True, the probability of failure of some components of a system can often be accurately calculated in a PRA; however, in the calculation of the system's total probability of failure, assumptions have to be made about how the different components contribute to the system's failure, how they interact, and whether or not the list of failure mechanisms can reasonably be assumed to be complete. Difficulties in correctly assessing these issues will make the estimated failure probability of the whole system uncertain. It would also be wrong to say that safety factors differ from PRA in not taking the whole system into account. In a safety-based design of a bridge, a safety factor is applied to the loading of the bridge as a whole, and the parts are optimized to achieve the targeted capacity. This is a holistic approach in the same way as the calculation of the probability of a bridge collapse is holistic.

A second claim based on the presumed integrative approach is that probabilistic design allows for comparison of the strengths of several elements within a system and that they can accordingly indicate which element to improve. Probabilistic approaches seem therefore fit for identifying critical elements and setting up maintenance schemes (Čepin 2002; Kong and Frangopol 2005; Vesely et al. 1994; Wang et al. 1996). This argument has force. As noted in Section 4, experience from the nuclear industry indicates that a PRA can be used to identify weak components in a system even if an accurate total probability of system failure cannot be calculated. Especially fault tree analyses can be useful in this way (Khan and Haddara 2003; Lapp 2005); however, it should be noted that non-probabilizable uncertainties need to be taken into account in maintenance planning and that a one-sided focus on known failure mechanisms may be dangerous (Arunraj and Maiti 2007).

A third proclaimed advantage related to the integrative approach of PRA is its ability to provide a more *inclusive* analysis of failure, including failure due to human error. Analysis of disasters and serious incidents shows that human error is one of the major causes of structural failure (Nowak and Collins 2000). Attempts have been

made to quantify risks of human error so that it can be included in an aggregate calculation of probability of failure (Dougherty 1997; Reer 2008; Sun et al. 2009); however, critics of such quantitative human reliability analysis argue that human error depends on elements such as commitment, attitude, and experience that are difficult to quantify. It is particularly doubtful whether meaningful numerical probabilities can be assigned to human reactions to new and untested technologies. These critics argue that attention should be redirected towards qualitative or human-centered approaches that aim at identifying new types of human failure rather than at quantifying currently known failure types (Hollnagel 1991; Mosneron-Dupin et al. 1997).

In sum, the claim that PRA is able to provide a more integrative and inclusive account of a system's safety and provide higher safety levels is rather weak, with one exception: for *intra*-system comparison PRA approaches have capacities that do not seem to be shared by any non-probabilistic methods. Such comparisons are important for the prioritization of risk reduction measures and for the planning of maintenance programmes.

### 5.4 Compatibility with Risk Assessment—Management Paradigm

A fourth advantage of probabilistic approaches concerns the organizational distinction between risk assessment and risk management. As mentioned in Section 2, the standard ideal for the risk decision process emphasizes a division of this process into two distinct parts: risk assessment and risk management (National Research Council 1983).

Compared to the safety factor approach, PRA seems more compatible with this organizational division between risk assessment and risk management. The selection of safety margins is a value-dependent exercise that forms part of the basis for scientific and technological work. In contrast, a PRA can be performed on the basis of scientific information alone. It is then up to the regulatory decision makers to set the acceptable probability of failure; however, as we also saw in Section 2, in most fields of engineering, there is no separation in practice between risk assessment and risk management.

Technical standards are typically set by groups of experts who are entrusted both with collecting and interpreting the scientific data and with proposing regulation. In structural engineering, for example, the establishment of the new European construction standard (Eurocodes) was characterized by organizational confluence of risk assessment and risk management (Clausen and Hansson 2007). Similarly, in hydraulic engineering, Vrijling et al. (1998) developed a unified framework for the assessment of safety in terms of acceptable individual and societal risks levels, which they derived from accident statistics and an estimate of the value of human life. Although the authors admit that the final judgment is, in the end, a political one, the proposed approach tries to merge risk assessment and management into one decision procedure.

These examples illustrate how the notion of probability and probabilistic design enter the domain of risk management (i.e., the domain where decisions on the acceptance of risks are to be made). Hence, although PRA in principle facilitates a clear distinction between risk assessment and risk management, the acceptable risk levels in PRA are often decided in the community of safety

experts who make the assessment as well. Furthermore, decisions on risk assessment and risk management issues are often made by the same expert committees, such as committees for technical standard setting or for the setting of exposure limits (Hansson 1998). Hence, the organizational structure does not support or encourage a separation between risk assessment and risk management. This is a severe limitation on the practical applicability of the proclaimed advantage of PRA that it is well suited for making this separation.

In this section, we discussed four arguments that are used in favor of probabilistic approaches. We showed that some of these were problematic and that the advantages of probabilistic approaches over the safety factor approach should therefore be somewhat qualified. The argument based on optimization was convincing but only for those PRA outputs that can be well calibrated with actual frequencies. The argument regarding compatibility with the standard view on the risk decision process was found to be sound but seldom utilized in practice. The two remaining arguments, improved precision and capability of system level analysis, were found to be rather weak.

## 6 Arguments for Using the Safety Factor Approach

In this section, we discuss three arguments against replacing safety factors by probabilistic risk assessment. The arguments refer to computational costs and simplicity, residual uncertainties, and security.

### 6.1 Analysis Costs and Simplicity

Probabilistic models promise to provide accurate estimates of failure probabilities that depend on many different input variables. The costs for data acquisition and computation tend to increase rapidly with the number of input variables. In practice, this leads either to unworkably long time for the analysis or simplifications of the model (with an unavoidable decrease in accuracy). Especially when the additional time also involves delays in the design and engineering process itself, the simplicity of the safety factor approach may be an advantage, also from a cost–benefit point of view. In the building industry, the efficiency of the building process is often more important than the amount of material used. Hence, reducing the construction time may—also from a cost–benefit perspective—be preferable over saving construction material.

Moreover, the simplicity of the safety factor approach can also make mistakes less likely. The importance of simplicity in safety work is known from chemical plant design. Plants with inherently safer technologies tend to be simpler in design, easier to operate, and more tolerant of errors (Overton and King 2006). Similarly, simpler calculation or design methods may be preferable over complex ones since they reduce the likelihood of mistakes in the calculations and hence, the likelihood of mistakes in the construction itself.

### 6.2 Capturing Residual Uncertainties

One of the important characteristics of probabilistic methods is that they can take potential adverse effects into account only to the extent that their probabilities can be

quantified (Clausen et al. 2006; Hansson 2009a; Knoll 1976). Although attempts are made to quantify as many elements as possible, most notably human errors, this can at most be done approximately. In practice, these difficulties may lead to a one-sided focus on only those dangers that can be assigned meaningful probability estimates. Probabilistic approaches tend to neglect potential events for which probabilities cannot be obtained (Hansson 1989; Knoll 1976: p. 411). Safety factors, on the contrary, are intended to compensate also for in practice unquantifiable uncertainties such as the possibility that there may be unknown failure mechanisms or errors in one's own calculations. It is a rational and not uncommon practice to set a higher safety factor to compensate for uncertainty. This is done routinely in toxicology (Fairbrother 2002; Santillo et al. 1998) and it seems logical to do this in other fields as well.

Safety factors are not the only method in safety engineering that takes uncertainties into account. The same applies to safety principles such as inherent safety and multiple safety barriers. These principles have in common that they introduce some degree of redundancy in the system, which is often an efficient way to protect also against dangers for which meaningful probability estimates are not available. If one of the safety measures fails for some unknown reason or if an unforeseen failure mechanism is activated, then an additional defense can provide protection.

Such "extra" safety measures may not be defensible from the perspective of a cost–benefit perspective, but they may still be justified from the perspective of protection against uncertainties (e.g., uncertainties about the probabilities of known risks and about unknown failure modes). For an example of this, suppose that a ship builder comes up with a convincing plan for an unsinkable boat. A PRA shows that the probability of the ship sinking is incredibly low and that the expected cost per life saved by lifeboats would be exceptionally high. There are several reasons why the ship should still have lifeboats: The calculations may possibly be wrong, some failure mechanism may have been missed, or the ship may be exposed to some unknown danger. Although the PRA indicates that such measures are inefficient, we cannot trust the PRA to be certain enough to justify a decision to exclude lifeboats from the design. Similar arguments can be used, for instance, for introducing an extra safety barrier in a nuclear reactor although a PRA indicates that it is not necessary. This is, of course, not an argument against performing PRAs but an argument against treating their outcomes as the last word on what safety requires.

### 6.3 Security and Vulnerability

A third argument in favor of the safety factor approach is related to security threats. So far, we have focused on safety, that is, the protection against unintended harm; however, the attacks on the twin towers on September 11, 2001 showed that not only "acts of nature" threaten the integrity of an engineering structure. We also need protection against another type of threat, namely those following from intended harm. This distinction is often expressed with the terms safety (against unintended harm) and security (against intended harm). Golany et al. (2009) refers to the former as probabilistic risk and the latter as strategic risk (where "strategic" refers to environments in which intentional actions are taken; it should be noted that Golany

et al. do not discuss the epistemic uncertainties that may also be present in strategic situations). An important distinction is that in the latter case there is an adversary who is capable of intelligent behavior and adapting his strategy to achieve his objectives. This has several implications.

First, it is in practice seldom meaningful to try to capture the likelihood of intended harms in probabilistic terms. Instead of assigning probabilities to various acts by a terrorist, it is better to try to figure out what actions would best achieve the terrorist's objectives. In such an analysis, the terrorist's responses to one's own preparative defensive actions will have to be taken into account (Parnell et al. 2008). Game theory (that operates without probabilities) is better suited than traditional probability-based analyses to guide prevention aimed at reducing vulnerability to terrorist attacks and most other intentional threats.

Secondly, as noted by Golany et al. (2009), whereas the criterion of effectiveness is adequate in safety work, in security work it should be replaced by the criterion of vulnerability. Vulnerability can be understood as a weakness that can be exploited by an adversary. The adversary's aim is not to maximize the likelihood of this loss but rather to maximize the loss itself (e.g., by targeting critical infrastructures and facilities). The optimal protection against terrorist attacks thus involves strategies to reduce this potential for loss. Probabilities do not have a central role in deliberations on how best to achieve such a reduction.

Sarewitz et al. (2003) add force to this line of argument by pointing out that vulnerability reduction can be considered a human rights issue, which in some situations may give it priority over economic optimization. Since modern society has an obligation to ensure that all citizens are provided a basic level of protection and that some fundamental rights are respected, economic arguments should not always be decisive in resource allocation. The authors give the example of the Americans with Disabilities Act, which requires that all public buses be provided with wheelchair access devices. This requirement was first opposed on economic grounds. Cost–benefit analyses showed that providing the buses with wheelchair access devices would be more expensive than providing, at public expense, taxi service for people with disabilities. The measure was nevertheless introduced in order to realize the right of people with disabilities to be fully integrated into society. The right to protection against violence can be seen as a similar fundamental right to be enjoyed by all persons. Such a right can justify protection even when a PRA or a CBA indicates that the resources would be "better" used elsewhere.

## 7 Discussion

Now that we have discussed the arguments in defense of probabilistic approaches and of safety factors, can we decide which approach is preferable? The (obvious) answer is no; both approaches are of value and it is does not seem constructive to see them as competitors. It should be recognized that neither of these methods can, in practice, tell the full truth about risk and safety (Hansson 2009b). In order to see how we could combine the insights from both approaches, let us reconsider the objectives of the two approaches as explained in Sections 3 and 4.

As we saw in Section 4, there are two different interpretations of the failure probabilities calculated in a PRA. One of these treats the calculated probabilities as relative indices of probabilities of failure that can be compared against a target value or against corresponding values for alternative designs. This interpretation seems unproblematic. It should be realized that it refers to a relative safety level; not all elements are included so it does not correspond to frequencies of failure in the real world (Aven 2009). Instead, this interpretation provides "a language in which we express our state of knowledge or state of certainty" (Kaplan 1993). It can be used to compare different engineering components within some engineering system or as a tool for priority setting and for the effect evaluation of safety measures. It is in this context of local optimization that probabilistic analysis has its greatest value (Lee et al. 1985).

The other interpretation treats the outcomes of PRA as objective values of the probability of failure. According to this view, these probabilities are not to be seen as merely relative indicators but as (good estimates of) objective frequencies. In a world that has no uncertainties but only known, quantifiable risks, this would indeed be a valid assumption; however, we do not live in such a world. In practice, this means that failure probabilities often include experts' quantified estimates of certain probabilities, and these estimates are unavoidably subjective (Caruso et al. 1999). Often some phenomena are excluded from the analysis. In these cases, we cannot translate a probability of failure from one context to the other. To compare the safety of nuclear power plant with the safety of a flood defense system on the basis of PRAs of the two systems is an uncertain and arguably not even meaningful exercise since their uncertainties are different and difficult or perhaps even impossible to compare.

Let us now turn to the safety factor approach that was said to be intended for compensating for five major categories of sources of failure. Two of these, namely higher loads and worse material properties than those foreseen, are targeted both by safety factors and PRA. Taking seriously the higher precision of probabilistic approaches, quantitative analysis of these sources of failure should at least in many cases preferably be based on probabilistic information.

The main advantage of the safety factor approach over PRA concerns the other three sources of failure that are unquantifiable: imperfect theory of the failure mechanisms, possibly unknown failure mechanisms, and human error (e.g., in design). PRA is not capable of capturing those uncertainties. This is a major reason why PRA should be seen as one of several tools for risk assessment and not as the source of final answers on risk assessment; however, current safety factors and other acceptance criteria, as laid down in codes, standards and regulations, may lead to suboptimal allocation of scarce resources (Rackwitz 2004).

Given that probabilistic approaches are incapable of capturing the non-quantifiable uncertainties and that current safety factors can be overly conservative with respect to known risks, how could current design codes in engineering be improved? In the literature, at least three approaches are described. The first two approaches depart from PRA and try to include uncertainties in the calculation of the probabilities, either by introducing an "extra" variability among the statistical parameters to account for the lack of information (Slijkhuis et al. 1999) or by applying safety margins to the probability itself (Vrijling 2001); however, especially

the first approach leaves us with the problem of secondary uncertainties (i.e., the problem on how to estimate the uncertainty about the uncertainty). The third, more widely advocated approach is the inverse of the second. In this approach, the safety factor is defined in probabilistic terms ("partial safety factors") (Elishakoff 2004). The reason why this approach is preferable over the second (applying safety margins to the probability itself) is that it leaves more room for incorporating considerations of security. As shown in Section 6, protecting against security threats requires a fundamentally different policy, namely reducing vulnerability rather than estimating in numerical terms the probability that a particular plant, building, or technological system will be subject to an attack. Traditional safety engineering approaches are better equipped for addressing this.

To conclude, our comparison of safety factors and PRA suggests that the two should be seen as complementary rather than mutually exclusive. Using PRA may lead to a one-sided focus on those dangers that can be assigned meaningful probabilities. Since not all dangers can be quantified and since most decision making is done under conditions of uncertainty, PRA cannot provide the final answer to safety issues. This holds even more when security threats come into play. On the other hand, when optimization becomes important—be it in the prioritization of maintenance measures or in situations where we are faced with hazards that cannot be eliminated—PRA can be an indispensable tool for priority setting and for the effect evaluation of safety measures.

**Acknowledgements** This paper was written during the first author's stay as visiting doctoral researcher at the Philosophy Department of the Royal Institute of Technology, Stockholm, which was funded by the Netherlands Organisation for Scientific Research under grant number 360-20-160.



## References

Abbas, A. M., & Manohar, C. S. (2002). Investigations into critical earthquake load models within deterministic and probabilistic frameworks. *Earthquake Engineering and Structural Dynamics, 31*(4), 813–832.

Allen, D. E. (1981). Limit states design—what do we really want. *Canadian Journal of Civil Engineering, 8*(1), 44–50.

Arrow, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situations. *Econometrica, 19*(3), 404–437.

Arunraj, N. S., & Maiti, J. (2007). Risk-based maintenance—techniques and applications. *Journal of Hazardous Materials, 142*(3), 653–661.

Aven, T. (2009). Perspectives on risk in a decision-making context—review and discussion. *Safety Science, 47*(6), 798–806.

Bedford, T., & Cooke, R. M. (2001). *Probabilistic risk analysis: Foundations and methods*. Cambridge: Cambridge University Press.

Blockley, D. I., & Godfrey, P. S. (2000). *Doing it differently.* London: Thomas Telford.

Caers, J., & Maes, M. A. (1998). Identifying tails, bounds and end-points of random variables. *Structural Safety, 20*(1), 1–23.

Caruso, M. A., Cheok, M. C., Cunningham, M. A., Holahan, G. M., King, T. L., Parry, G. W., et al. (1999). An approach for using risk assessment in risk-informed decisions on plant-specific changes to the licensing basis. *Reliability Engineering & System Safety, 63*(3), 231–242.

Čepin, M. (2002). Optimization of safety equipment outages improves safety. *Reliability Engineering & System Safety, 77*(1), 71–80.
Chang, R. (Ed.) (1997). *Incommensurability, incomparability, and practical reason*. Cambridge: Harvard University Press.
Chapman, P. M., Fairbrother, A., & Brown, D. (1998). A critical evaluation of safety (uncertainty) factors for ecological risk assessment. *Environmental Toxicology and Chemistry, 17*(1), 99–108.
Clausen, J., & Hansson, S. O. (2007). Eurocodes and REACH: Differences and similarities. *Risk Management, 9*(1), 19–35.
Clausen, J., Hansson, S. O., & Nilsson, F. (2006). Generalizing the safety factor approach. *Reliability Engineering & System Safety, 91*(8), 964–973.
Cornell, C. A. (1969). Probability-based structural code. *ACI Journal, 66*, 974–985.
Ditlevsen, O. (1994). Distribution arbitrariness in structural reliability. In G. I. Schueller, M. Shinozuka, & J. T. P. Yao (Eds.), *Structural safety & reliability, vols. 1-3. ICOSSAR '93* (pp. 1241–1247). Rotterdam: Balkema.
Ditlevsen, O., & Madsen, H. O. (2007[1996]). *Structural reliability methods (internet edition 2.3.7)*. Chichester: John Wiley & Sons Ltd.
Dougherty, E. M. (1997). Is human failure a stochastic process? *Reliability Engineering & System Safety, 55*(3), 209–215.
Elishakoff, I. (2004). *Safety factors and reliability: friends or foes?* Dordrecht/Boston/London: Kluwer Academic Publishers.
Fairbrother, A. (2002). Risk assessment: lessons learned. *Environmental Toxicology and Chemistry, 21*(11), 2261–2263.
Golany, B., Kaplan, E. H., Marmur, A., & Rothblum, U. G. (2009). Nature plays with dice—terrorists do not: allocating resources to counter strategic versus probabilistic risks. *European Journal of Operational Research, 192*(1), 198–208.
Graham, J., & Wiener, J. (1995). *Risk versus risk*. Cambridge: Harvard University Press.
Guikema, S. D., & Paté-Cornell, M. E. (2002). Component choice for managing risk in engineered systems with generalized risk/cost functions. *Reliability Engineering & System Safety, 78*(3), 227–238.
Hampshire, S. (1972). *Morality and pessimism*. Cambridge: Cambridge University Press.
Hansson, S. O. (1989). Dimensions of risk. *Risk Analysis, 9*(1), 107–112.
Hansson, S. O. (1998). *Setting the limit. Occupational health standards and the limits of science*. New York: Oxford University Press.
Hansson, S. O. (2007). Philosophical problems in cost-benefit analysis. *Economics and Philosophy, 23*(2), 163–183.
Hansson, S. O. (2009a). From the casino to the jungle. *Synthese, 168*(3), 423–432.
Hansson, S. O. (2009b). Risk and safety in technology. In A. W. M. Meijers (Ed.), *Handbook of the philosophy of science, vol. 9. Philosophy of technology and engineering sciences* (pp. 1069–1102). Amsterdam: Elsevier/North-Holland).
Hansson, S. O. (2009c). An agenda for the ethics of risk. In L. Asveld & S. Roeser (Eds.), *The ethics of technological risk*. London: Earthscan.
Harris, B., & Soms, A. P. (1983). A note on a difficulty inherent in estimating reliability from stress strength relationships. *Naval Research Logistics, 30*(4), 659–663.
Hollnagel, E. (1991). What is a man that he can be expressed by a number? In G. Apostolakis (Ed.), *Probabilistic safety assessment and management*. Beverly Hills: Elsevier.
JCSS (2001). *Probabilistic model code. Part 1—basis of design*. ISBN: 978-3-909386-79-6, Joint Committee on Structural Safety.
Kaplan, S. (1993). Formalism for handling phenomenological uncertainties. The concepts of probability, frequency, variability, and probability of frequency. *Nuclear Technology, 102*(1), 137–142.
Keynes, J. M. (1921). *A treatise on probability*. London: Macmillan.
Khan, F. I., & Haddara, M. A. (2003). Risk-based maintenance (RBM): a quantitative approach for maintenance/inspection scheduling and planning. *Journal of Loss Prevention in the Process Industries, 16*(6), 561–573.
Knight, F. H. (1935[1921]). *Risk, uncertainty and profit*. Boston: Houghton Mifflin.
Knoll, F. (1976). Commentary on the basic philosophy and recent development of safety margins. *Canadian Journal of Civil Engineering, 3*(3), 409–416.
Kong, J. S., & Frangopol, D. M. (2005). Probabilistic optimization of aging structures considering maintenance and failure costs. *Journal of Structural Engineering, 131*(4), 600–616.
Kuijper, H. K. T., & Vrijling, J. K. (1998). Probabilistic approach and risk analysis. In K. Pilarczyk (Ed.), *Dikes and Revetments: Design, Maintenance and Safety Assessment* (pp. 443–62). Rotterdam: AA Balkema.

Lapp, S. A. (2005). Applications of fault tree analysis to maintenance interval extension and vulnerability assessment. *Process Safety Progress, 24*(2), 91–97.

Lee, W. S., Grosh, D. L., Tillman, F. A., & Lie, C. H. (1985). Fault tree analysis, methods, and applications—a review. *IEEE Transactions on Reliability, 34*(3), 194–203.

Madsen, H. O., Krenk, S., & Lind, N. C. (1986). *Methods of structural reliability.* Englewood Cliffs: Prentice-Hall.

Mak, K. K., Sicking, D. L., & Zimmerman, K. (1998). Roadside safety analysis program—a cost-effectiveness analysis procedure. *General Design and Roadside Safety Features, 1647*, 67–74.

Mayer, M. (1926). *Die Sicherheit der Bauwerke und ihre Berechnung nach Grenzkräften anstatt nach zulässigen Spannungen*. Berlin: Springer.

Melchers, R. E. (2002). Probabilistic risk assessment for structures. *Proceedings of the Institution of Civil Engineers: Structures and Buildings, 152*(4), 351–359.

Melchers, R. E. (2007). Structural reliability theory in the context of structural safety. *Civil Engineering and Environmental Systems, 24*(1), 55–69.

Michal, R. (2000). The nuclear news interview. Apostolakis: On PRA. *Nuclear News*, 27–31 (March).

Möller, N., & Hansson, S. O. (2008). Principles of engineering safety: risk and uncertainty reduction. *Reliability Engineering & System Safety, 93*(6), 798–805.

Moses, F. (1997). Problems and prospects of reliability-based optimization. *Engineering Structures, 19*(4), 293–301.

Mosneron-Dupin, F., Reer, B., Heslinga, G., Strater, O., Gerdes, V., Saliou, G., et al. (1997). Human-centered modeling in human reliability analysis: some trends based on case studies. *Reliability Engineering & System Safety, 58*(3), 249–274.

Nathwani, J. S., Lind, N. C., & Pandey, M. D. (1997). *Affordable safety by choice: the life quality method*. Waterloo: University of Waterloo Press.

National Research Council. (1983). *Risk assessment in the federal government: managing the process*. Washington: National Academy Press.

Nowak, A. S., & Collins, K. R. (2000). *Reliability of structures*. Boston: McGraw Hill.

Overton, T., & King, G. M. (2006). Inherently safer technology: an evolutionary approach. *Process Safety Progress, 25*(2), 116–119.

Parnell, G. S., Borio, L. L., Brown, G. G., Banks, D., & Wilson, A. G. (2008). Scientists urge DHS to improve bioterrorism risk assessment. *Biosecurity and Bioterrorism: Biodefense Strategy Practice and Science, 6*(4), 353–356.

Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: six levels of treatment. *Reliability Engineering & System Safety, 54/2–3*, 95–111.

Rackwitz, R. (2001). *A new approach for setting target reliabilities. In safety, risk and reliability. Trends in engineering* (pp. 531–536). Zürich: IABSE.

Rackwitz, R. (2004). Optimal and acceptable technical facilities involving risks. *Risk Analysis, 24*(3), 675–695.

Randall, F. A. (1976). The safety factor of structures in history. *Professional Safety*, 12-28 (January).

Rasmussen, N. C. (1975). *Reactor safety study. An assessment of accident risks in U.S. Commercial Nuclear Power Plants (WASH-1400, NUREG 75/014)*. U.S. Nuclear Regulatory Commission.

Reer, B. (2008). Review of advances in human reliability analysis of errors of commission—Part 2: EOC quantification. *Reliability Engineering & System Safety, 93*(8), 1105–1122.

Santillo, D., Stringer, R. L., Johnston, P. A., & Tickner, J. (1998). The precautionary principle: protecting against failures of scientific method and risk assessment. *Marine Pollution Bulletin, 36*(12), 939–950.

Sarewitz, D., Pielke, R., & Keykhah, M. (2003). Vulnerability and risk: some thoughts from a political and policy perspective. *Risk Analysis, 23*(4), 805–810.

Savchuk, V. P. (1992). Some applications of probabilistic methods in space structures design. *Reliability Engineering & System Safety, 37*(2), 129–138.

Slijkhuis, K. A. H., Van Gelder, P. H. A. J. M., Vrijling, J. K., & Vrouwenvelder, A. C. W. M. (1999). On the lack of information in hydraulic engineering models. In G. I. Schueller & P. Kafka (Eds.), *Safety and reliability. Proceedings of the ESREL '99* (pp. 713–718). Munchen: A.A. Balkema.

Sun, Z. Q., Xie, H. W., Shi, X. J., & Liu, F. Q. (2009). Engineering approach for human error probability quantification. *Journal of Systems Engineering and Electronics, 20*(5), 1144–1152.

Takewaki, I. (2005). A comprehensive review of seismic critical excitation methods for robust design. *Advances in Structural Engineering, 8*(4), 349–363.

Thoft-Christensen, P., & Baker, M. J. (1982). *Structural reliability theory and its applications*. Berlin: Springer-Verlag.

Van de Poel, I. R. (2009). Values in engineering design. In A. W. M. Meijers (Ed.), *Handbook of the philosophy of science, vol. 9. Philosophy of technology and engineering sciences* (pp. 973–1006). Amsterdam: Elsevier/North-Holland.

Van Noortwijk, J. M., & Van Gelder, P. H. A. J. M. (1998). Bayesian estimation of quantiles for the purpose of flood prevention. In B. L. Edge (Ed.), *26th International Conference on Coastal Engineering* (pp. 3529–3541). Copenhagen: ASCE.

Vesely, W. E., Belhadj, M., & Rezos, J. T. (1994). PRA importance measures for maintenance prioritization applications. *Reliability Engineering & System Safety, 43*(3), 307–318.

Vose, D. (2000). *Risk analysis* (2nd ed.). New York: Wiley.

Vrijling, J. K. (2001). Probabilistic design of water defense systems in The Netherlands. *Reliability Engineering & System Safety, 74*(3), 337–344.

Vrijling, J. K., van Hengel, W., & Houben, R. J. (1998). Acceptable risk as a basis for design. *Reliability Engineering & System Safety, 59*(1), 141–150.

Vrouwenvelder, A. C. W. M. (2002). Developments towards full probabilistic design codes. *Structural Safety, 24*(2–4), 417–432.

Wang, J., Yang, J. B., Sen, P., & Ruxton, T. (1996). Safety based design and maintenance optimisation of large marine engineering systems. *Applied Ocean Research, 18*(1), 13–27.

Zegeer, C. V., Carter, D. L., Hunter, W. W., Stewart, J. R., Huang, H., Do, A. et al. (2006). Index for assessing pedestrian safety at intersections. In *Pedestrians and bicycles* (pp. 76–83). Washington: Natl Acad Sci.