3 The Concepts of Risk and Safety

Niklas Möller
University of Cambridge, Cambridge, UK

S. Roeser, R. Hillerbrand, P. Sandin, M. Peterson (eds.), *Handbook of Risk Theory*,
DOI 10.1007/978-94-007-1433-5_3,

Abstract: The aim of this chapter is to analyze the concepts of risk and safety in the context of societal decision-making. Risk and safety research is a heterogeneous field, and different areas have conceived of the nature of risk in different ways. In the chapter, I categorize risk perspectives in three broad groups: the scientist approach, the psychological approach, and the cultural approach to risk. Between these groups, the nature and status of risk and safety have been the debated subjects. I will attempt to bring some light onto complicated and controversial philosophical topics such as whether risk and safety are natural or normative notions, whether they are social constructions, objective, or even real. This investigation will focus on a range of different questions. I will distinguish between five common definitions of the term "risk," as well as contrast the notion of *risk* with both the notion of *safety* and the notion of *acceptable risk*. The main part of the chapter will focus on a quantitative or comparative concept of risk, that is, a notion that is in play in statements such as "the risk of flying is lower than the risk of traveling by car" and "the risk of nuclear power is 10^{-4} deaths per reactor year." The central aspects of such a notion of risk and safety will be discussed, in particular the notions of probability and harm. I will also discuss the common claim that it is the expectation value of the severity of harm that is the correct measure of risk. Furthermore, I investigate additional aspects such as epistemic uncertainty and other, more controversial aspects that have been proposed.

Introduction

Risk research is a discipline in rapid development with contributors from many areas of the natural and social sciences. This reflects a growing concern about risks in society. Both professional and non-professional awareness of risks are increasing, and much effort is put into risk assessment, risk management, and risk communication. As a consequence, there are now well-developed societal practices in place involving risk and safety. Still, the central concepts of risk and safety remain somewhat unclear. When characterized, risk and safety are often treated either as relatively straightforward natural science concepts, or, to the contrary, as fundamentally subjective notions ill fitting for scientific study. However, without an in-depth understanding of its central concepts, the subject matter of risk and safety research remains fuzzy and it is unclear what the objective of reducing risk and achieving safety really amounts to.

The aim of this chapter is to analyze the concepts of risk and safety in the context of societal decision-making. This investigation will focus on a range of different questions. In the first section, I will categorize risk perspectives in three groups: the scientist approach, the psychological approach, and the cultural approach to risk. It aims to give an initial background and show the different points of departure that are characteristic for the heterogeneous field of risk research, and to supply an initial context for us to orient from and relate to in what follows.

In the second section, I distinguish between several meanings of the term "risk," and discuss the notion of acceptable risk as well as the relation between the notions of risk and safety.

In the third section, I will discuss aspects of the quantitative concept of risk that will be in focus for the remainder of the chapter. I will investigate the fundamental aspects of probability and severity of harm, which are part of most quantitative conceptions of risk and safety. Furthermore, I will investigate additional aspects such as epistemic uncertainty and other, more controversial aspects that have been proposed.

The topic of the fourth section is the nature of risk and safety, by which is meant the status of the concepts: are risk and safety or normative notions, are they objective, or are they social constructions? Are they even real? The aim of the fifth section is to shed some light onto these complicated and controversial philosophical topics.

I suggest some topics for further research in the sixth section, before I end with a short conclusion in the final section.

Risk Perspectives

Several different fields of investigation have been interested in the concept of risk: engineering, economics, political science, sociology, psychology, and philosophy, to name just a few. This heterogeneous research arena has resulted in many different perspectives on risk. Some theorists have grouped these perspectives rather finely. Ortwin Renn, for example, has divided the risk approaches into seven categories: (1) the actuarial approach, (2) the toxicological and epidemiological approach, (3) the engineering approach, (4) the economical approach, (5) the psychological approach, (6) social theories of risk, and (7) cultural theory of risk (Renn 1992, p. 56). (Even a fine-grained categorization such as this is incomplete. As indicated by the present volume, *philosophy of risk* is a growing area of research; cf. e.g., Shrader-Frechette 1993; Lewens 2007; Asveld and Roeser 2009; Hansson 1998, 2009). Often, however, theorists are satisfied with a less fine-grained grouping, depending on the task at hand.

For the purpose of this chapter, let us distinguish between three broad approaches to risk. The first perspective can be called the *scientist approach* to risk. The basic idea is that risk is a phenomenon that may be investigated like most phenomena in science, that is, by employing the scientific method. Risk is something that can, at least in principle, be measured in a systematic way, and the main task of the researcher is to find a sufficiently precise measure of the phenomena and to find ways of reducing the risk as much as possible. On most interpretations, the first four of Renn's perspectives would belong to this category. Statistical and probabilistic tools are important in this perspective as ways of measuring and describing the risks, in addition to investigations into the causal mechanisms of various risk-related phenomena that are at the core of both the toxicological-epidemiological and the engineering approaches (Renn 1992).

The second perspective may be labeled the *psychological approach* to risk. The basic interest in this perspective is to study people's *perceptions* of risk, that is, people's beliefs about risks and their way of relating to them. A dominant psychological method is psychometrical research in which the researcher tries to establish reliable measures of risk perceptions. The aim of the approach is to get a clear and distinct picture of how people estimate risks and how they make choices in relation to them – in particular, what influences whether they deem a risk acceptable or not. Various attitudes toward risk in general – especially risk-aversive and risk-seeking behavior – as well as what types of risk we deem more important than others, are typical topics of interest on this approach (Slovic 2000; cf. also, Hansson 2010 and Summerton and Berner 2003 for this category).

The third approach may be called the *cultural approach* to risk (roughly corresponding to the last two of Renn's categories). Whereas the psychological approach mainly focused on the *individual* and her ways of conceiving of risk-related affairs, the cultural approach takes a broader perspective. On this approach, the main interest is to establish how our conceptions of risk are culturally mediated, that is, how they are formed by social contexts in our societies (e.g., identity and power). A particular risk statement is always articulated in a cultural context, singled out among many other possible formulations. Moreover, for every risk event we pick out as the interesting one, there are other potentially hazardous events we could have chosen. The cultural approach to risk is interested in, as Clarke and Short put it, "how social agents create and use boundaries to demarcate that which is dangerous" (Clarke and Short 1993, p. 79).

In this chapter, my main aim is analytical in the sense that I will analyze the conceptual aspects or constituent dimensions of risk and safety. What do terms such as "risk" and "safety" mean when we use them in statements such as "It is safe to fly," "Nuclear power is safer than other means of energy production," and "The risk of nuclear power is 10^{-4} deaths per reactor year?" In particular, I will investigate the comparable or quantitative notion of risk and safety exemplified in the latter two statements. Here, the scientist conception of risk and safety will form an interesting starting point.

Thus, I will focus on the meaning of the relevant terms, rather than psychological and social aspects such as what individuals think about risks (a topic for risk perception studies) or how these beliefs concerning risk and what we choose as the area of study concerning risk is dependent on our social context (a topic for cultural studies of risk). That does not mean that results from psychological or sociological studies of risk are irrelevant. Since there is a close relationship between our linguistic behavior and the meaning of the term that is used, actual practice matters. In particular, I will in section ➲ Further Aspects of Risk and Safety investigate some suggestions of risk aspects that have been put forward in the literature. Furthermore, I will return to the different perspectives on risk in section ➲ The Nature of Risk and Safety, where I will turn to the question about the *nature* of risk and safety, notably whether they are to be seen as objective, scientific concepts or whether they have some other status. Here, researchers belonging to different traditions have tended to take rather different stands.

Notions of Risk

Five Definitions of "Risk"

The terms "risk" and "safety" have been used with many related but distinct meanings. It is helpful to distinguish between at least five different, but clearly related meanings that have been used in the literature (Hansson 2004a; Möller et al. 2006):

1. Risk = an *unwanted event* which may or may not occur (Rosa 1998)
2. Risk = the *cause* of an unwanted event which may or may not occur
3. Risk = the *probability* of an unwanted event which may or may not occur (Graham and Weiner 1995)
4. Risk = the fact that a decision is made under conditions of *known probabilities* (Knight 1921; Douglas 1983)
5. Risk = the statistical *expectation value* of unwanted events which may or may not occur (Willis 2007; Campbell 2005)

The first meaning is displayed in a statement such as "Wildfires constitute the most serious environmental risk in Russia today" or "There is always the risk of an accident when driving in traffic." "Drunk driving constitutes a major traffic risk" and "Coronary heart disease is the number one risk of death in America" are two examples of the second meaning of "risk." A risk in the second sense is sometimes referred to as a *hazard* in an engineering context.

The third meaning is perhaps the most common, for example, in statements of the form "the risk of heavy rainfall this week is more than fifty percent." Note here that "risk" in this sense is not merely a synonym of "probability," but is used to mark the undesirability of the outcome, that is, rainfall. Hence a farmer would perhaps say that it is the *chance* of rainfall that is 50% rather than the *risk* of it.

The fourth sense of risk is a technical notion that is often used in decision theory. Here, one typically distinguishes between *decisions under certainty*, which are decisions where all the consequences of the decision alternatives are known; *decisions under risk*, which are decisions where the probabilities of the outcomes are known; and *decisions under uncertainty*, where the probabilities are unknown. In this context, claiming, say, that whether to install a certain warning system is a decision under risk is to claim that the situation can be treated as a decision where the probability of failure of the system is known.

The fifth sense is another technical sense in which the notion of risk is used. While the notion of expected value as such dates from the early development of probability theory in the seventeenth and eighteenth centuries, its application in the risk context is fairly new. It became common after the influential Rasmussen Report of 1975 and is now the standard definition of "risk" in risk analysis (Rechard 1999, p. 776).

The expectation value is the probability-weighted sum of the severity of harm. It measures the *magnitude* of the risk as the combination of two factors, the probability of an unwanted event, and its severity. It supplies an overtly *quantitative* sense of risk that is used both to compare risks, and to give a single magnitude of risk.

In risk analysis, the expectation value is often used for a single quantitative statement of a risk, or as the basis of comparative statements when claims are made that some technology is safer than another. Often in statements such as "the risk of flying is less than the risk of driving a car," it is the expectation values of both activities that are compared. (Note, however, the potential need for increased precision of such statements, for example, whether the basis for comparison is per traveled kilometer or per traveled hour.)

While all of these five meanings of "risk" are legitimate on their own terms, my main interest in this chapter is the application of the term for comparative or quantitative purposes. The last sense of the term, the notion of expected value, fits this aim. It gives a value of risk for an event that may be compared with other events, and it also gives a unique, freestanding measure of risk. In the next section, I will take a closer look at the quantitative notion, including the expected value conception.

A terminological note before I continue. In the remaining chapter, I will use "harm" instead of "unwanted event," for several reasons. If a potential unwanted event is small, it may be incorrect to talk about risk or safety. For example, drawing a blank ticket in a lottery would be an unwanted event, but the avoidance of this would not be described as a matter of safety (unless, of course, the lottery was about something severe, such as when a person participates in a game of Russian roulette). Thus, the nature of the unwanted event is relevant here: if its severity is below a certain level it does not count as a risk and safety issue. (The term "unwanted event" also refers to the subject's desires in an unfortunate way. If you, for some reason, had a desire to hurt yourself, and therefore engaged in a stunt act in which the probability of a severe accident is very high, we might say that you chose a *safe* way to kill yourself, meaning *certain*, but we would never say that you were safe.) I will therefore use the term "harm" that – contrary to "unwanted event" – implies a non-trivial level of damage.

Acceptable Risk

The quantitative notion of risk that is our main object of study should be distinguished from the notion of *acceptable risk*. The quantitative notion is in place when we compare risks or ascribe their magnitude. Even if we assume an answer to the magnitude question, whether we should accept a risk of a certain magnitude is yet a further question. In other words, we should distinguish between the *magnitude* of the risk on the one hand, and whether or not a risk of that magnitude should be *accepted*. They correspond to two different notions that must be kept separated in order to avoid conceptual fallacies about risk.

Historically, early studies in risk analysis were aimed at finding a level of risk – interpreted as the expectation value of harm – that should be accepted (Otway 1987). An example of such "low levels of risk" concerns the harmful effects of background radiation, or other risks that we accept in ordinary life. While helpful as comparability tools, the idea of any such fixed levels of acceptable risk have been questioned (Peterson 2002). The general objection to the idea is that if a risk is *additive*, that is, adds to existing risks, there will be an addition to the overall risk even if the risk is small, and hence we need a further justification for adding the risk.

A major reason for accepting or not accepting a risk is naturally due to the benefit of the risk in relation to the harm of imposing it. Cost–benefit analysis of risk is a method of risk analysis that aims to judge the acceptability of a risk by comparing the benefits to the "cost" that the risk corresponds to. While the viability of cost–benefit analysis of risk is heavily questioned (Le Grand 1991; Sen 1987; Hansson 2004b; Fischhoff et al., Kirmsky and Golding, Shrader-Frechette, contributions to Asveld and Roeser 2009), among other things since it predominately relies on the controversial presumption that we may compare risk and benefits on a single (monetary) measure, it is hard to avoid the general idea that the acceptability of a risk has *some* relation to its benefits.

The acceptability of a risk depends also on many moral aspects involving questions such as agency, rights, and volition. Arguably, if a risk is voluntary, such as smoking, it may be acceptable even if the same risk, were it involuntary, would not be. Similarly, a certain risk level may be acceptable if the persons taking the risk are the ones benefiting from them, but not otherwise (Hansson 2004b).

Is Safety the Antonym of Risk?

Risk and safety are closely related concepts. Until now, I have followed the tradition and mainly discussed the notion of risk, assuming that the relation between the two is uncontroversial. This, however, cannot always be assumed.

While risk is what we typically quantify and compare, safety is what we want to achieve. In the literature, the notion of safety is predominately used as the sought state-of-affairs. (This is the common usage in regulation and procedural documents, where the focus is on describing rules and procedures to enhance safety, but a direct and precise characterization of the concept of safety is missing; cf. e.g., IAEA 2000). We want the nuclear power plant to be safe, as we do the person walking alone in a city park and the cough medicine we use. Still, both risk and safety are used, as I have done in this chapter, to make comparative claims. The common picture of the relation between the two concepts is that they are antonyms: when the risk is low, safety is high, and conversely, when safety is low, risk is high. We could phrase a statement "it is safer to be at sea when you are sober than when you have been drinking," just as well as we could phrase it "it is less risky to be at sea when you are sober than when you are drunk."

There are, however, at least three potential complications for the antonym picture of risk and safety. The first comes from the fact that the two terms have different connotations. Safety is a positive property, while risk is generally something negative. Therefore, the level of risk may perhaps not be too high for us to claim that X is safer than Y. If there is a high risk for both the plague and cholera, but slightly less for cholera, it may be strange to say that we are safer from the one than the other, even if that would be correct on the antonym view.

Secondly, one should note that the term "risk" is often – explicitly or implicitly – given a technical definition. On such a usage, it is an open question whether safety is its antonym. It has been argued, for example, that, if the expected value of harm is used as a definition of risk, safety is not to be understood as the antonym of risk, since other aspects are relevant for safety (Möller et al. 2006). On a broader notion of risk, including the aspects that will be addressed in the next section, the antonym view is more plausible.

The third complication comes from the monadic use of the safety predicate, most clearly expressible with the term "safe" – such as in "the bridge is safe" rather than the dyadic (or comparative) "the new bridge is safer than the old one used to be." The application of the safety concept may be given an absolute and a relative interpretation. Consider the question "is my car safe?" One way of answering it is to reply: "No, since there is always a risk of being in a traffic accident – if you want to play it safe, stay home!" That way of interpreting the question would be the absolute sense of safety. According to the absolute interpretation, safety against a particular harm implies that the risk of that harm has been eliminated. Some authors take the absolute sense of safety for granted. For example, in the context of aviation safety it has been claimed that "[s]afety means no harm" (Miller 1988) and that "[s]afety is by definition the absence of accidents" (Tench 1985).

Another way of answering the question would be to reply: "Yes, the car is safe, since the risk of an accident in this car is low, and the latest safety features it comes equipped with also minimize the risk of a severe damage in case of an accident." That reply would be in line with a relative interpretation of safety, where safety means that the risk has been reduced or controlled to a certain acceptable level. A typical example taken from a safety application states: "[R]isks are defined as the combination of the probability of occurrence of hazardous event and the severity of the consequence. Safety is achieved by reducing a risk to a tolerable level" (Misumi and Sato 1999). The US Supreme Court is explicit about its use of a relative safety concept when they claim that "safe is not the equivalent of 'risk free'" (Miller 1988, p. 54).

The relative safety concept must not only be distinguished from the absolute interpretation, but also from the above-discussed notion of *acceptable risk*. As we have seen, whether a risk is acceptable may depend on the benefits, and on moral or social aspects. While there is a social element in expressions such as "tolerable level" of risk as well, we do not allow the same leeway to the notion of safety, and it is often reasonable to claim that while something is not safe, the risk is acceptable. Therefore, it seems as if it is actually the notion of acceptable risk that the American Department of Defense is referring to when it states that safety is "the conservation of human life and its effectiveness, and the prevention of damage to items, consistent with mission requirements" (Miller 1988, p. 54).

Both the absolute and the relative concepts of safety are legitimate, but they must be carefully separated in order to avoid misunderstandings. For most practical purposes, the absolute concept corresponds to an ideal that cannot be realized, and "safe" is used in the relative sense. Hardly anything for which it would be interesting to apply the concepts of risk and safety contains no risk at all, strictly speaking – not even, as the first answer suggested, staying at home.

Aspects of Risk

I will now look at the quantitative notions of risk and safety, that is, the concepts that are in place in statements such as "Flying is safer than going by car" or "Nuclear power has the lowest risk of all energy production methods."

In the last definition of risk in section ➋ Five Definitions of "Risk", "risk" was defined as the expected value of harm. As noted there, this is the most common definition in risk analysis, and some authors have even taken it as the only rational way of measuring risk. Bernard Cohen, for example, starts off a recent paper with the claim (Cohen 2003, p. 909):

- ▶ The only meaningful way to evaluate the riskiness of a technology is through probabilistic risk analysis (PRA). A PRA gives an estimate of the number of expected health impacts – e.g., the number of induced deaths – of the technology, which then allows comparisons to be made with the health impacts of competing technologies so a rational judgment can be made of their relative acceptability.

In this section, however, I will investigate whether the notion of expected value of harm gives a complete understanding of risk and safety, and I will discuss other aspects that have been suggested in the literature.

Before this, however, let us start with a methodological note. It may very well be held that as long as a definition of a term, for example, "risk," is internally coherent, it makes no sense to ask whether it is complete. We are *told* what the meaning of the term is, and if this meaning is intelligible, then no complain may be launched.

Indeed, this objection is correct as it stands. Here, however, our aim is to grasp the meaning of "risk" and "safety" in the context relevant to societal decision-making. Hence, whether the technical notion of expected value of unwanted events capture this meaning is another question. Although the expected value may supply a clear and distinct quantitative concept, the important question is whether it fits with the phenomena we are trying to capture.

In other words, our aim is to characterize the notion of risk involved in our natural language usage, not in any internal-to-science-only way. In the end, what decision-makers as well as laypeople want to know when they ask how the *risks* are distributed – or which of the alternatives is the *safest*, etc. – is the answer that is relevant for the intended meaning of the words. If we answer using the same terms but with another meaning, we have not given the adequate answer. Hence, it is important not only that we are clear about how our terms should be interpreted, but also that we are using the *right* interpretation. Risk research is principally an empirical field of study, but if we are dealing with inadequate conceptualizations, our analyses may be inadequate, too. (The approach used here has much in common with *ordinary language philosophy*, in which a core assumption is that knowledge of the meaning of terms is reached not primarily through theories in abstraction, but through a close attention to the details of ordinary language in which it is used (Soames 2005). Cf. also Hare (1952, p. 92) for a similar point against the claim that we may define moral terms as we please.)

Harm and Probability

The notions of harm and probability are central for any quantitative understanding of risk and safety. As soon as we move beyond an absolute conception as discussed in section Is Safety the Antonym of Risk?, where safe means no risk at all, we are dealing with events that *may* take place with different degrees of likelihood, and causing different levels of harm.

Starting with harms, we should note that it is far from trivial to compare their severity. On a commonsense notion of harm, it seems clear that there are many cases in which we may reasonably hold that one harm is more severe than another. In a traffic accident, for example, a death is more severe than a broken leg, and a broken leg is more severe than a bruise. But the relative severities between them are not so certain. How many broken legs, one may reasonably ask, are there on a death? It is not clear what the answer should be.

This is problematic if we think that we can always, in principle, compare risks. In particular, for the expected value of harm to be well defined, the severity of harm must be possible to measure rather extensively. More precisely, we must be able to decide not only that harm A is more severe than B, and that B is more severe than C, but also the relative severity between them. (In technical terms, we must be able to compare harms on an interval scale; cf. Resnik (1987) for a textbook presentation.)

In practice, when arguing for (and applying) the expected value notion of risk one typically limits the measure of the severity of harm to one significant value, typically casualties. So severity of harm in traffic is in the first instance measured in traffic deaths. A perhaps sometimes reasonable assumption behind such a measure is that the less severe harms "follow along," but that is sometimes a far-fetched idealization.

Even if we assume that death is the primary harm to take into account, it is unclear if the severity of harm is the same in all cases. Is, in all circumstances, the possibility of a 95-year-old person dying from a medical procedure of the same severity as if the patient were 25 years old? This is an area of severe controversy, but we should note that there are ambitious systems of measurement designed to take account of both the quantity and the quality of life generated by healthcare interventions, such as the QALY measure (quality-adjusted life-year).

The second aspect of risk, probability, is a much-investigated concept with a well-established mathematical content, and it is paramount in many important applications. The interpretation in cases of risk and safety is not evident, however. In probability theory there is a well-established distinction between subjective and objective interpretations of the concept. According to the objective interpretation, probability is a property of the external world, for example, the propensity of a coin to land heads up. According to the subjective interpretation, to say that the probability of a certain event is high means that the speaker's degree of belief that the event in question will occur is strong (Ramsey 1931; Savage 1972). When we are dealing with the repetition of technological procedures with historically known failure frequencies, it may be possible to determine probabilities that can be called objective. However, in most cases such frequency data are not available, unless perhaps for certain parts of the system under investigation. Therefore, frequency data will have to be supplemented or perhaps even replaced by expert judgment. Expert judgments of this nature are not, and should not be confused with, objective fact. Neither are they *subjective* probabilities in the classical sense, since by this is meant a measure of a person's degree of belief that satisfies the probability axioms but does not have to correlate with objective frequencies or propensities (see section Uncertainty). The judgments are better described as subjective estimates of objective probabilities. Furthermore, the probability estimates used in risk and safety analysis are not purely personal judgments even in this sense. Rather, they are based on the best possible judgments that can be obtained. Typically, a probability estimate is interpreted as the best possible judgment from the community of experts.

The theoretical issues of comparing harms and assigning probabilities do not only constitute a problem for proponents of the expected value notion of risk, but for any account that, while objecting to the expected value interpretation *as such*, still takes severity of harm and probability as basic aspects of risk and safety. To the extent that there is a problem in assigning a reasonable probability or severity of harm value for an event, there is a potential problem in assigning a measure of risk. As will be seen in the next two subsections, this problem is present even if we manage to agree about an interpretation of the most reasonable ascription of the probability and severity of harm that can be obtained.

Risk as the Expected Value of Harm

That the severity of harm and its probability are important aspects of risk is well in line with basic intuitions about the notions. Assuming that the measurement problems in the last subsection are solved, it seems reasonable to assume, *all else equal*, that if the likelihood of an accident increases, so does the risk; likewise, if the severity of an outcome increases, so does the magnitude of the risk. So far, it is hard to deny that these two aspects should be part of the quantitative concepts of risk and safety. That, however, falls short of accepting that the expected value of harm supplies the correct measure of risk.

Indeed, some authors, while endorsing the two aspects, have preferred more cautious formulations. Slovic, for example, has defined risk as “a blend of the probability and the severity of the consequences” (Slovic 2000, p. 365). Other authors have likewise expressed themselves in terms of “combinations” rather than explicitly stating how that combination is brought into a measure (Lowrance 1976; ISO 2002; Aven 2007). Still, as previously noted, the expected value is the dominant notion of risk in risk analysis (Hansson 2005).

There is, however, a strong reason for the claim that the expected value of harm is the proper measure of risk. It is, proponents argue, the only measure the use of which will minimize total harm *in the long run*. If we assume the *expected* number of people dying of heart attack in a country is 400,000 per year, that typically means that about 400,000 *will* actually die from heart attack on a given year. Imagine that we had a new revolutionary treatment for coronary heart disease that would decrease the expected number by 100,000. Switching to the first treatment would typically mean that around 300,000 people would die the following year.

This strong correlation between the *expected* value of an outcome and the *actual* value follows from what is called the Law of Large Numbers, which is the theorem in probability theory that states that if independent trials with the same probability of outcomes are repeated, the average value of the trials converges to the expected value. In other words, in the long run, given the above assumptions, actual value converges to the expected value. The classic illustration of this is the tossing of a (non-biased) coin: whereas there is nothing strange in tossing three heads in a row, we would expect the relative frequency after, say, 100 or 1,000 tosses to be very near to 0.5.

There are, however, several objections to defining the risk as the expectation value of harm. In the previous section I discussed some problems with comparing harms in a sufficiently extensive way as to make the notion of expected value well defined. Even if we assume that such a measure is available, however, there are problems with defining risk as expected value.

One main objection takes on the background assumption that is explicit in the very name of "Law of Large numbers." It is based on noting that as soon as the numbers of events in question are less than "large," there is no guarantee that the actual outcome is close to the expected one. (Indeed, *given* any number of events, there is no guarantee that the expected value is close to the actual outcome.) Let us imagine a future where we have a much smaller number of people dying from heart attack, say typically 20. In this future society, we are faced with a clear choice (don't ask me about the mechanisms for this, it is a thought experiment) between medicine A or B. With medicine A there is a 0.00001 probability that 1,000,000 will die, and 0.99999 that no one will die. With medicine B there is a probability of 0.5 that no one will die, and 0.5 that 30 will die. With medicine A, the expected value is 10, whereas the excepted value with medicine B is 15. The proponent of the expected value notion of risk would thus claim that using medicine A would mean the least risk, but this, the objectors argue, is not at all clear. The *actual* outcome could in fact be 100,000 deaths in heart attacks the following year, if treatment A is chosen, whereas in case B, it would be, at maximum, 30 persons. Hence, they may conclude, the safest option is B.

The basic idea in this objection is thus that there may be a discrepancy between the expected value and the actual value, depending on the actual distribution of outcomes and their matching probabilities, and that it is thus not irrational to oppose the general identification of risk with the expected value of harm. If we are dealing with "one-shot" events rather than frequent events of the same type and the same independent probabilities, the expected value may be nothing more than *one* aspect of the risk, not the full determinant of its magnitude. In particular, when the potential harmful outcome is extreme, such as the extinction of mankind, the expected value notion is deficient as a measure of the risk.

Uncertainty

While probability and harm are often alluded to as the basic aspects of risk and safety, some authors explicate the terms by focusing on the aspect of *uncertainty*. Hence, risk is sometimes explained with reference to an uncertain consequence of an event (Renn 2005; Aven and Renn 2009). Under some interpretations, however, this notion of uncertainty reduces into probability. Aven (2003: xii), for example, uses probability and probability calculus as "the sole means for expressing uncertainty." This probabilistic understanding of risk and safety has a substantial justification in the subjectivist *Bayesian framework* (Howson and Urbach 2006). As mentioned above, on a subjectivist interpretation, probability is conceived of as representing all aspects of a decision-maker's lack of knowledge. On a Bayesian construal, *all* rational decisions are fully representable with precise probabilities, since the rational decision-maker always, at least implicitly, assigns a probability value to each potential outcome. Faced with new information, the agent may change her probability assessment (in accordance with Bayes' theorem), but she always assigns determinable probabilities to all states of affairs. Thus, in the Bayesian view all uncertainty about what will happen is codified in the probability assessment for the outcome at hand (Ramsey 1931; de Finetti 1937; von Neumann and Morgenstern 1944; Savage 1972).

The reduction of uncertainty to probability is however not the only alternative. On the contrary, as was argued above the probabilities behind risk and safety ascriptions are not typically given a subjective Bayesian interpretation. Rather, it is more in line with what is called the classical view in decision theory, in which a basic distinction is made between situations with known probabilities and situations where probabilities are unknown or only partially known (Knight 1921). Proponents of the classical view have pointed out that there is a large difference between situations with well-determined probabilities, such as coin-tossing, and less well-determined situations such as assigning probabilities to whether a major accident will happen in a complex plant. In the latter case there is *epistemic uncertainty* that, according to these authors, may not be reducible to a unique probability value in a rational way. This has led many contemporary theorists also of Bayesian bent to argue for the inclusion of (non-probabilistic) epistemic uncertainty into the analysis (Ellsberg 1961; Kyburg 1968; Levi 1974; Gärdenfors and Sahlin 1982).

One way of illustrating the idea of epistemic uncertainty is to imagine two situations in which you are walking alone in the jungle, and are about to cross an old wooden bridge. The bridge looks unsafe, but you have been told by people in the local village that the probability of a breakdown of this type of bridge is less one in ten thousand. Contrast this with a case where you in your jungle walk are accompanied by a team of scientists, come across a bridge, and the scientists carefully investigate the bridge and conclude that the probability that it will break is rather one in five thousand. Even though the probability is now judged as higher, the epistemic uncertainty is far smaller and it is not unreasonable to regard this situation as preferable to the first one in terms of safety.

This kind of example suggests a picture in which the risk increases with the probability of harm, with the severity of harm, and with uncertainty. So two events with the same severity of outcome that are given the same probability for occurring in accordance with the best estimations possible may still warrant different ascriptions of risk, if the uncertainty is considered different in the two cases. Assume, for example, that we are to compare the safety of two different means of energy production that are both given the same probability of severe failure, one utilizing an old and well-tested method, and another utilizing a new method. Naturally, the new method has been well tested as far as possible, but we do not have the same source of performance data for the entire system as in the case of the old method. Here, it may be reasonable to claim that the risk of the new system is higher, due to the greater epistemic uncertainty in question.

Many methods and techniques in safety engineering may, in the first instance, reasonably be interpreted as methods of reducing the uncertainty. Take the method of *inherently safe design*. Here the idea is to switch from controlling potentially dangerous materials and subsystems to using materials and subsystems that are less dangerous in themselves, that is, that do not need to be controlled in the same degree. Hence, a possible hazard is *removed* rather than *contained*. For example, fireproof materials are used instead of inflammable ones, and this is considered superior to using inflammable materials but keeping temperatures low (Möller and Hansson 2008).

While the general notion may be reasonably clear, a fundamental question is how epistemic uncertainty should be characterized in more detail. This is a controversial area in which no consensus has been reached. The most extensive discussions have been in decision theory regarding how to express the uncertainty of probability assessments. Here, two major types of measures have been suggested, binary and multi-valued measures. A binary measure divides the probability values into two groups, possible and impossible values. In typical cases, the set of possible probability values will form an interval, such as: "The probability of an epidemic outbreak in this area within the next 10 years is between 5% and 20%." Binary measures have been used, for example, by Ellsberg, Kaplan and Levi (Ellsberg 1961; Kaplan 1983; Levi 1986).

Multivalued measures generally take the form of a function that assigns a numerical value to each probability value between 0 and 1. This value represents the degree of reliability or plausibility of each particular probability value. Several interpretations of the measure have been used in the literature, for example, second-order probability (Baron 1987; Skyrms 1980), fuzzy set membership (Unwin 1986; Dubois and Prade 1988), and epistemic reliability (Gärdenfors and Sahlin 1982). See Möller et al. (2006) for an overview.

In summary, while many authors agree that epistemic uncertainty is an important aspect for risk and safety *in addition* to probability, how to include it in more detail is strongly controversial. Perhaps this should not surprise us, however. A general measure of what we do not know may sometimes be too much to ask for, even in theory.

Further Aspects of Risk and Safety

For the quantitative notion of risk, several candidate aspects beyond probability, severity of harm and epistemic uncertainty have been suggested in the literature. In particular, empirical studies of risk – so-called *risk perception studies* – have established many other aspects that are relevant for how people judge risks. For example, in a seminal study Fischhoff et al. (1978/2000) presented eight aspects that determined people's preparedness to accept a risk. In addition to severity of consequence – which I have already discussed – the aspects were: Voluntariness of risk, Immediacy of effect, Knowledge about risk, Control over risk, Newness, Chronic-catastrophic, Common-dread. Results like these have helped shape a demand, directed at risk assessors, to take other aspects than the traditional under consideration.

Now, as I noted in section ➲ Acceptable Risk, there is an important distinction between the *magnitude* of a risk and its *acceptability*. Many of the aspects in the above study seem relevant to the *acceptability* of risk. Still, these aspects return as suggestions for aspects relevant to the very notion of risk as well. May there be any merit to such suggestions?

In evaluating such a claim, we should first distinguish individuals' *conceptualization* from the *concept* in question. While there typically is a deep connection between competence and actual use of a concept, many of our beliefs about a concept can be wrong even if we are competent with it. Moreover, we seldom think that we have full knowledge about how to exactly delimit a certain concept. I take myself as pretty competent with a rather wide range of concepts – such as *table*, *computer*, *book*, and *coffee mug*, just to mention some of the ones I just applied to objects in my proximity – but I am not sure that I would be able to delimit any of these concepts with any certainty. So even if we did show that there is an aspect that many people *take* to be part of the concept of risk, it does not mean that it actually *is*.

Such a skeptical stance has been a common attitude for many traditional risk analysts faced with suggestions beyond their preferred notion of risk, that is, typically the expected value notion (Hansson 2010). In their arsenal, they have used risk studies that seem to have shown the irrationality of people's risk perception. For example, it has been shown that people often take the risk of exposure to chemicals and radiation as something binary – either you are exposed or not – and neglect the dose factor (Kraus et al. 2000/2000). Since actual harm is predominately a factor of dose, it amounts to a misconception of the risk involved.

It certainly is a radical claim that such beliefs about chemical exposure are *irrational*. It is indeed an intelligible position – it just happens to be wrong. But false beliefs are not irrational beliefs. The point to make here, however, is that while there is a prima facie reason to take strongly held beliefs about a notion seriously, if we have sufficient evidence we may show that they are mistaken.

When it comes to many of the suggested candidate aspects of risk and safety, I believe indeed that they are mistaken, when thought of as *conceptual* aspects. Take the aspect of *control* (Möller et al 2006). To say that a certain risk is *more controllable* for an agent than another risk means, in the present context, that there is a more reliable causal relationship between the acts of the agent and the probability and/or the severity of the harm (this is an extension of the psychometrical concept, which only regards the subjective dimension). At first sight one might believe that, everything else being equal, more controllability implies less risk. Arguably, if we only have the option of flying on automatic pilot, we would be less safe than we would if we had the option of turning it off as well. However, this relationship does not seem to hold in general. We have, again, to distinguish between a psychological perception of safety and actual safety, since we may feel safe without actually being safe, and *vice versa*. Consider two nuclear power plants, one of which runs almost automatically without the intervention of humans, whereas the other is dependent on frequent decisions taken by humans. If the staff responsible for the non-automatic plant is poorly trained, it seems reasonable to maintain that the automatic plant is *safer* than the non-automatic one *because* the degree of controllability is lower in the automatic plant. Arguably, even if they are excellently trained, a high degree of control may lower the safety. Consider a third nuclear plant that is much less automated than the ones we have today. Due to cognitive and mechanical shortcomings, a human being could never make the fast adjustments that an automatic system does, but this plant has excellent staff that perform as well as any human being can be expected to do under the circumstances. In spite of this increased control this too would be a less safe plant. The reason is that the probability of accidents due to human mistakes would be much greater than in a properly designed, more automatic system.

The effects of control on safety in our nuclear plant example can be accounted for in terms of the effects of control on probability. If increased human control decreases the probability of an accident, then it leads to higher safety. If, on the other hand, it increases the probability of accidents, then it decreases safety.

Denying that control is part of the *concepts* of risk and safety is not denying that it may have other relations to the phenomena. For example, it may be a *heuristic device* in many situations: it might be that having the option to causally control our systems often would increase their safety. This correlation, however, is an *empirical* possibility rather than a conceptual one.

For other aspects, it seems rather clear that what is at stake is not the quantitative notion of risk and safety at all, but that they are reasons for the acceptability of a risk. Voluntariness of risk, for example, seems to be about the acceptability of imposing involuntary risk and not a claim that the degree of voluntariness changes the magnitude of risk in any sense.

There are, however, aspects of a potential risk event beyond probability, severity of harm and epistemic uncertainty that have been argued to belong to the concept of risk and safety (Möller 2009a, 2010). Although admittedly controversial, they are instructive in order to further demonstrate the complexity of inferring ascriptions of risk and safety.

Distributive aspects illustrate a further complication with harms and risk: even if we assume a measure, such as a reasonable way of comparing a broken leg with a casualty (as discussed above), how to infer the risk from different *distributions* of potential harms may not always be clear (Hansson 2005; Möller et al. 2006).

Imagine that a revolutionary method of building encapsulated nuclear power plants has made the population safe from even the unlikely event of a meltdown and thus decreased the total expected value of harm. The only drawback is that service staff must make some internal maintenance and this is very risky, having a high expected value of harm, several magnitudes above current levels for any staff. Let us further assume that the level of epistemic uncertainty is deemed to be negligibly low in both the current and the new and revolutionary method.

The question we can ask ourselves is whether, in cases like these, safety is merely a matter of receiving as low an expected value of harm as possible in a population, or if the distribution of potential harms also should count? Maybe we should not hold that a situation is safer than another if there are some people carrying significantly higher risks than others? In most societal activities carrying risks, there is an uneven distribution of potential harm. Persons living close to risky artifacts such as energy production plants or heavy chemical plants take a higher risk load than persons living further away. Likewise, for road traffic safety there is a debate about using cables for protecting the vehicle from driving off the road. The cables are successful in avoiding many potentially lethal car accidents; however, they may also be very dangerous for motorcyclists; much more so, in general, than if the motorcyclists merely went off the road. Since there are many more cars than motorcycles on the roads, the expectation value of harm is smaller when using the wires than not, yet for the motorcyclists the risk is much higher. It may be asked whether it is then reasonable to claim that the method is safe as long as the total expected value is kept low, regardless of the particular risks for motorcyclists.

Delimitation issues. The second aspect I will mention highlights the complicated nature of selecting the base events for ascriptions of risk and safety – in this case the harmful potential outcomes (Möller 2010). Let us imagine a situation where we are to judge whether automobile traffic between two particular cities is safer than airplane traffic. Extensive frequency data tells us that the probability of serious harm when traveling by airplane is lower than traveling by car. We have reason to believe, however, that to a significant extent, the one-car accidents are intentional, that is, acts of suicide. For air traffic, however, there are almost no accidents that may reasonably be labeled as suicide. If we were to exclude the suicide cases from the frequency data, the expected value of harm for car travel would be lower.

In this thought example, we face the question of how to delimit the risk and safety concepts: what events should count as risk events, and hence be included in frequency data and other means of evaluating the risk and safety at hand? If we were to include all harmful events, air travel would be safest. But it seems wrong to include also suicides in the statistics of road traffic safety. It is one thing to allow mistakes from the driver – and a main part of accidents certainly derive from the human factor – but in judging what means of transportation is the safest, it seems irrelevant that it is possible to use the car also as a tool of committing suicide. Note that this is not a question of epistemic uncertainty (although in practice there certainly is such a question as well). The question is rather that given that we know that an event is a traffic suicide event, should we include it in the basis for the frequency data? It is evident that only some harms count, not others, and we may reasonably doubt the answer, in the thought experiment, to which means of transportation is the safest.

Admittedly, in these two cases of distribution and delimitation, our intuitions may vary. But it seems at least plausible to doubt here that what has the lowest expectation value of harm in these cases are actually safest. Note that for this to be plausible, we do not need to claim that the actual risk of one event is greater or even equal to the other when the expected value of harm of it is less. The weaker claim that the risks are incommensurable, that we cannot reasonably compare them in such a case, is sufficient.

As mentioned above, there are many aspects of risk and safety proposed in the literature; distribution and delimitation aspects such as mentioned here are mere examples of the questions involved in inferring the risk from a given, complex situation. Still, they exemplify the possibility of further reasonable inferential aspects for the quantitative notion of risk, even beyond the general question of how the dimensions of probability and severity of harm should be combined into one measure. In the next section, we will investigate whether these aspects, among others, represent a problem for a scientific notion of risk.

The Nature of Risk and Safety

The nature of risk and safety has been understood in widely different ways in the literature. Some authors have taken risk to be a scientific, objective, or natural notion, picking out real properties in the world, while others have taken it to be something far more normative or subjective, or even denied its existence. Generally, these commitments can be mapped to a high degree with the three risk perspectives that were mentioned in section ➋ Risk Perspectives: the *scientist notion* of risk, the *psychological approach*, and the *cultural approach*. On the scientist approach, risk is typically seen as a scientific notion and, correspondingly, a phenomenon that may be investigated and measured in a systematic way, at least in principle. With that often comes the idea that risks are natural, objective, and real features of the world. On the *psychological approach*, the focus is on how individual people conceive of risks and their acceptability. But many psychological researchers are interested in how this relates to how the risks actually *are*, and often seem to assume that there is an objective fact of the matter to compare with (Summerton and Berner 2003, p. 6; Hansson 2010, p. 232).

The strongest opposition to conceiving risk and safety as objective concepts comes from proponents of the *cultural approach* to risk. As mentioned in section ➋ Risk Perspectives, on the cultural approach, risk is typically seen as a subjective and social phenomenon. Often this is expressed in terms of a social construction: risk is a social construction, since the content and delimitation of risks are socially articulated and treated. People from different cultures as well as within one culture have very different views on what constitutes a risk and how severe a risk may be. Sometimes proponents of this view express that there is no fact of the matter over and above these individual or cultural views, committing them to a denial of risk as objective and perhaps even real. (These are merely initial characterizations: as will become clear, all of the three risk perspectives are *compatible* with objectivist, subjectivist, or constructivist conceptions of risk as described below.)

Even if we assume that there is a lot to say in favor of the initial characterization just given, it should be unpacked a little more carefully. Taking risk to be *natural*, *objective*, or *real* are distinct thoughts that I will investigate further in the following section. First, in section ➋ Risk and Normativity, I will distinguish between taking risk and safety as *natural* concepts on the one hand, and as *normative* concepts on the other, and discuss some allegedly normative aspects of risk and safety. In section ➋ Thick Concepts and Reductionism, I will present an interpretation of risk and safety as what in philosophy has been called *thick concepts*, and discuss the idea that risk can be reduced to a natural concept. Lastly, in section ➋ Risk, Objectivity, and Social Constructions, I will discuss the idea that risk and safety are social constructions, and the related question whether that means that risks are not objective or real.

Risk and Normativity

In the traditional use, a natural concept is a concept that is invoked in scientific explanations or, more narrowly, used to express the laws of nature (Little 1994; Vallentyne 1998; the latter characterization is in line with traditional understanding of natural kinds: the close relation between natural kinds and law-like regularities is emphasized by most natural kind theorists, cf. Lange 2007; Boyd 1991; Hacking 1991; Wolf 2002). The philosopher G.E Moore has offered a number of influential characterizations to this effect, suggesting that a natural property is a property that is the subject matter of natural science and psychology, and that it is a property that can be known by means of empirical observation and induction (Moore 1903, pp. 25–27). Paradigmatic natural concepts are *water*, *gold*, *mass*, and *redness*.

Normative concepts are typically characterized as opposed to natural concepts. An often-used image is that the natural and the normative have opposite "directions of fit": to be true, natural claims should fit what the world is like, whereas it is what the world is like that should fit the true normative claim (Williams 1985, Blackburn 1998). Normative concepts such as *good*, *right*, and *fair* are thus action-guiding; that an action is good, right, or fair entails that we have a reason to perform it. (I am in this chapter using "normative" broadly in a sense that refers both to deontic concepts such as *right*, *ought*, and *permitted*, and to value concepts such as *good*, *bad*, and *better*. Sometimes "normative" is used to refer only to the former category, while "evaluative" is used only for the latter. While the relation between deontic and value concepts is an interesting topic in its own right, for the purpose of the present chapter, it is the distinction between the cluster of these broadly normative/evaluative concepts on the one hand, and the natural concepts on the other, that is of interest.) In contrast, natural concepts have no such *prima facie* action-guiding feature: that something contains water may be a reason to drink it if you are thirsty, but not otherwise.

Is risk a natural or a normative concept? As we have seen, many proponents of the scientist approach to risk view it as a perfectly natural, descriptive concept. This goes in particular for proponents of the expectation value of harm conception of risk, where risk is understood as a function of probability and severity of harm, where both harm and probability are taken as natural concepts.

However, the picture of risk and safety as natural, scientific concepts has been criticized. The criticism is grounded on either *external* or *internal normativity*.

External normativity is the normativity involved in answering questions about whether or not a risk should be imposed, is acceptable, or small enough to be safe. I label it "external" since it is external to the quantitative notions of risk and safety that have been the main subjects of investigation in the current chapter.

Many sources of external normativity have already been mentioned in section ➲ Acceptable Risk on acceptable risk. As soon as we leave the realm of absolute safety as discussed in section ➲ Is Safety the Antonym of Risk?, where safety is interpreted as a level of no risk, the question whether a certain level of risk should be accepted, or deemed safe, is relevant. In the early studies of acceptable risk, much work centered on what level of risk people thought was acceptable. Reasonably, what people *take* to be an acceptable level of risk is relevant for what level we *should* accept. Indeed, this is a basis for democratic decision-making. But generally they are different questions. The latter question about what level we should accept is a paradigmatic normative question. Hence, even if we take a certain level of safety as a given, whether that level is *sufficiently* small to be acceptable or safe is a further normative question about what we have reason to do.

This normative status is most obvious when the reasons for accepting or not accepting a risk do not have to do only with the levels of risk as such – e.g., whether they are larger than other risk levels that we have accepted (see section ➋ Acceptable Risk) – but with moral considerations, such as rights, agency, and autonomy. It is particularly evident when the acceptability of a risk depends on moral reasons that go in different directions. For example, that the direct risk of smoking is voluntary may be seen as a reason for allowing it in public areas. On the other hand, there may be a statistical connection between allowing people to smoke in public areas and the number of people that actually take up smoking. That could be considered as a reason against allowing it. The outcome of such a question of whether or not to allow public smoking, however based on scientific results on its harmful effects, is clearly a normative statement.

Internal normativity. That questions of acceptable risk and safety are normative, however, is hardly controversial today, especially after the evidently normative problems involved in trying to set a general limit for acceptable risk mentioned above. To the contrary, in modern risk analysis one is careful to distinguish between the quantitative assessment of risks with the assessment of the viability or acceptability of the risk. This corresponds to the typical division of risk analysis into two stages: *risk assessment* and *risk management* (NRC 1983; EC 2003; sometimes, as in the NRC model, risk assessment is divided into two stages as well, the research stage and the assessment stage). Risk assessment is traditionally considered the scientific stage where the estimations of the risks at hand are produced. Risk management then uses the output of the risk assessment as input for making a decision about the risk, ultimately whether to accept or reject it.

Thus, theorists that claim that risk is a natural concept – who we will refer to as *naturalists* – should typically be seen as referring to the *quantitative notion* of risk. Compare with paradigmatic measures such as *length*. Here, we may claim that someone's length may be a natural, scientific property and still agree that whether that length is acceptable for a certain task, such as being a basketball player, is a normative affair.

Proponents of a naturalist view on risk have several reasons in their arsenal. While the nature of probability is arguably a debated subject, as previously mentioned what is normally of interest in risk and safety contexts is not a subjective Bayesian notion but rather an objective notion such as relative frequencies (Resnik 1987). In addition, it is often claimed that severity of consequences may in many cases be measured to a sufficient degree. A basic measure is the number of lives at stake, so that higher severity means more potential deaths.

Critics, on the other hand, point to the very complications concerning the aspects of risk that I have discussed in this chapter. Probability, for example, may seem as an untouchably scientific notion due to its profound mathematical basis, but while a relative frequency is indeed a statistical fact, its role in deciding the actual probability of an event may be questioned, since it records historical data and the question is always how that applies to the situation at hand. Moreover, for many complex systems, we have sufficient frequency data for only parts, and have to estimate the overall probability, and there may be no value-free way of making that estimate.

This dependence on values is even clearer when we move beyond probability to epistemic uncertainty. There is no established measure of uncertainty. The many alternative suggestions put different aspects of uncertainty to the forefront, and it is unclear that there can be a general measure of such an elusive entity.

Naturalists can reply that these objections merely point to the fact that there are *epistemic values* involved when making risk and safety ascriptions. However, they may continue, this is not anything specific to the risk and safety area – in fact, philosophers of science have argued that epistemic values such as reliability, testability, generality, simplicity, etc., are integral to the entire process of assessment in science (McMullin 1982; Kuhn 1962; Lakatos and Musgrave 1970). Hence, they are the kind of values that the scientific experts are competent to apply in their scientific enterprise. Naturalists conclude that showing that risk ascriptions are normative in this sense merely puts it on the same level as our most paradigmatic natural notions such as *water* and *mass*.

In other words, on the naturalist defense, we have to be skeptics about *all* scientific notions in order to be skeptical about the notion of risk. Some theorists bite the bullet and draw that skeptical conclusion (Mayo 1991, pp. 254–255; Wynne 1982, p. 139). But even if we accept that all scientific claims are conditioned by epistemic values, and that this does not make them normative notions in the sought sense (merely an internal-to-science-only sense), it may be doubted that the values involved in risk ascriptions are limited to these values. The values involved in criteria for accepting a theory or an empirical result into the scientific corpus are typically biased toward avoiding false-positives, that is, taking something as a scientific fact when it is not. For risk and safety analysis, however, we may have a broader scope in order to avoid underestimating the risks (Möller 2009b; Wandall 2004).

Probability and epistemic uncertainty are only two dimensions of risk and safety, however, and the problem for the risk naturalist may be even larger when we turn to harm, and to the further problem of inferring risk and safety from all of these aspects. Even if we grant that a certain harm may be a perfectly natural notion, it is less clear that severity of harm may be ascribed in any value-free way. As I mentioned in section ➋ Further Aspects of Risk and Safety, neither comparing different harms nor ascribing a level of severity to an event such as a death seems to be possible to do in a value-free way. Indeed, it seems normative if anything is.

Furthermore, there is the added complication of inferring risk and safety from its "base aspects." As we have seen, the traditional risk assessment measure of expected value of harm has been heavily criticized. If we do not accept the expected value as a measure of risk, how then are we to infer the risk? Is there a scientific way of making this inference, one using – at the most – only values internal to science? The skeptic to risk as a natural concept doubts this.

If we include also the more controversial inferential *distribution* and *delimitation* aspects discussed in section ➋ Further Aspects of Risk and Safety, the risk naturalist faces even further obstacles. To argue – as in the case of the nuclear power plant where staff were to be exposed to a significantly larger potential harm than the population at large – that there are cases where the uneven distribution of potential harm is relevant for the very level of risk (and not only for the acceptability of a certain level) is to argue that distributive aspects are *part* of the very concept of risk and safety. And it seems doubtful, even if we grant that natural science may give us a probability and a severity of harm for an event, that there is anything for natural science to say about such distributive questions as these.

The delimitation aspect also seems to be a normative aspect of risk and safety, that is, a normativity due to which events should count as safety events, relevant for assessing the question of the risk and safety in the case at hand. The traffic suicide case indicates that not *all* potential harms are relevant for risk and safety. What is qualitatively identical on the level of physical harm is not necessarily identical for evaluating the risk and safety because one event is a suicide and the other an unintentional accident. It has been argued (Möller 2009a, 2010) that there are several potential properties of a harmful event that may motivate an exclusion of it as a risk-relevant event in certain circumstances, e.g., that it constitutes harm-seeking behavior or harms that are too small in the relevant circumstance (even if they are frequent), etc.

Thick Concepts and Reductionism

As we could see in the subsection above there are several arguments for the normative nature not only of notions such as *acceptable risk* and *safe*, but also for the normativity of the quantitative notion of risk itself. Is risk as a normative notion a threatening idea for the proponent of the naturalist notion of risk? Perhaps the naturalist may acknowledge that risk and safety are normative notions and yet argue that they may be reduced to natural ones? After all, it should be clear that the notion of risk has a substantial amount of non-normative content. In this section, I will place this idea in a broader philosophical context, namely, the philosophical debate of thick evaluative concepts.

When I introduced the notion of normative concepts above, I mentioned such normative concepts as *good* and *right*. Even if the arguments for risk and safety as normative concepts are sound, they seem to differ from these paradigmatic ones. That an action is good or right tells you that it is the thing to do, at least prima facie, but it does not tell you anything about what type of action it is. It does not have any (or at least not much) descriptive content, only normative content. To be kind to a stranger may be morally right, but so can, many believe, starting a war (that is the assumption behind the "just war" concept). And these actions seem to have very little in common, except for the – assumed – normative status of being right.

Certainly, there is also a normative dimension of risk. That something constitutes a risk is typically a reason against (allowing, using, performing) it, and the larger the risk, the stronger is that reason. Risk has a negative evaluative "direction," just as safety has a positive one. But risk is also importantly different from the paradigmatic normative notions of right and good, in that it is a notion with substantial descriptive content. Indeed, the richness of descriptive content is the main argument for the scientist approach to risk and safety. There are many descriptive characterizations of risk and safety that render them quite different from the paradigmatic normative notions of good and right: the central role of severity of harm; that a harm is not certain but may obtain; that a risk is greater the more likely it is that the harm may occur and the greater the harm, etc. Perhaps, proponents of the scientist approach to risk may argue, there is a possibility to acknowledge that risk has a normative aspect to it, but that we may, for scientific purposes, *reduce* the notion to its natural part.

In moral philosophy, there is an analogous debate among the kind of concepts that are called *thick normative concepts* – or thick concepts for short. In philosophy, John McDowell and Bernard Williams introduced the notion of thick concepts in the 1970s and 1980s (McDowell 1978, 1979, 1981; Williams 1985). Thick concepts are concepts such as *cruel*, *brave*, and *selfish*, concepts that have *both* descriptive and normative content. As such, thick concepts seem to differ from both paradigmatic natural concepts such as *water* and *length*, which have no such normative quality, as well as from paradigmatic normative concepts such as *good* and *right*, which are said to be solely or primarily normative. On the traditional analysis, a thick concept fills a double function: it describes a feature (world-guid*ed* function), and it evaluates it (world-guid*ing* function).

On our characterization of risk and safety, it seems as if these notions fit right into this category: that something is safe is a positive feature of the entity, and that something carries a risk is a negative feature of it. But it is not simply positive or negative, it is positive or negative *in a certain way*; it has a certain descriptive "shape," as I noted above. Grasping these aspects of the concepts is part of what one has to do in order to understand their meaning. Hence, it has been argued, risk and safety are to be understood as thick evaluative concepts (Möller 2009a, 2010).

When theorists first started to investigate (what would later be called) thick concepts, they believed that these could be analyzed as a *conjunction* of a descriptive part and an evaluative part, by which is meant that the descriptive content and the evaluative content can be independently given (Stevenson 1944; Hare 1952). "X is courageous" could therefore be analyzed as something along the lines of "X intended to act in the face of danger to promote a valued end" and "this is (*prima facie*) good-making." Furthermore, it was commonly assumed that the descriptive part determined the reference of the thick term. A thick concept, on this analysis, is a concept with a descriptive shape that decides the extension, and an evaluation that commends or condemns an entity for having these descriptive features. If this were correct, a descriptive reduction of a thick concept would, at least in principle, be possible. While something may still be said to be lacking in the understanding of the concept of courage in a person who grasped the descriptive part but took a neutral evaluative stance toward it (did not believe that courage was either good or bad), this person would be fully able to identify courage. For the concepts of risk and safety, this would mean that there indeed was a descriptive part that one could – in principle – isolate and operationalize. Perhaps an agent that understood only the natural part of risk and safety would not have a full understanding of the concept and all of what it signifies, but she would still be able to get the extension of the concept right. If this is possible, risk and safety would be normative concepts, but they would still be able to be *reduced* to natural concepts in the following important sense: natural concepts would still suffice as the descriptive kernel of the concept that would allow us to compare risk in a scientific, non-normative way. The normativity of risk and safety, on this understanding, could then be acknowledged without any threat for the scientific status of risk ascriptions.

To make this reductive idea clearer, it may be helpful to compare with the notion of body mass index (BMI), which is an individual's body weight divided by the square of his or her height). Having a BMI higher than 25 (based on the standard unit values kilogram and meter) is typically evaluated negatively, as it is considered bad for the health. Conversely at the other end of the spectrum, a BMI below 18.5 is considered a too low relative weight. In this sense, BMI may be seen as a concept that is a compound of both descriptive and evaluative parts. Here, however, the descriptive part is sufficient for picking out the extension of the term.

BMI is an example of a term with fully technical primacy: it is developed and applied using only natural notions. Hence, it is not at all surprising that these natural aspects are sufficient – it is a definitional fact of the term. What may be – and is – questioned is when a certain BMI is indeed unhealthy. That is, the evaluative aspects seem to have a secondary status as projections on the fundamentally descriptive or natural concept of BMI. The descriptive part, however, is rock solid.

The normative aspects argued for in the previous subsection, however, are hard to picture as secondary projections on the descriptive aspects. The point of the critique of a natural, non-normal conception of risk has been that normative aspects are prevalent in the very *ascription* of the quantitative risk, and not only in the evaluation of a previously already given quantification. Hence, in order to *establish* the magnitude of the harm, or the way in which the probability and severity are to be combined into a risk measure, we make evaluations that are over-and-above the natural basis of these notions.

For such notions where there is no clear primacy of a technical definition, philosophers in the thick concept debate have been increasingly skeptical of the reductive idea. Philosophers such as John McDowell, Bernard Williams, and Jonathan Dancy have argued that no such separation of a descriptive, autonomous part on the one side and an evaluation on the other can be given (McDowell 1978, 1979, 1981; Williams 1985; Dancy 1995, 2004). The only way to understand a thick concept is to understand the descriptive and evaluative aspects *as a whole*. The idea is that for a thick concept, the evaluative aspect is profoundly involved in the practice of using it, and therefore one cannot understand a thick concept without understanding also its evaluative point. (I here use the epistemological framing of the claim used by McDowell. Dancy (1995) goes even further, making the metaphysical claim that the descriptive and evaluative content is an indissoluble amalgam.)

On the analysis of risk and safety given in this chapter, it is hard to see the normativity of the aspects of risk and safety as a secondary, projective aspect of a notion that is otherwise essentially descriptive. The main reason for this is that the normative aspects of risk figured not only as an evaluation of the "finished product," as in the BMI example and in the question of the acceptability of the risk, but on the very ascription of the core aspects of risk, such as harm and epistemic uncertainty, and the delimitation and distributive aspects.

For risk and safety, thus, there is an essential interdependence between the natural-descriptive aspects and the normative-evaluative aspects. That an event has a higher expected value of death than another event – treated as a natural input – is typically a reason to ascribe the former a higher risk than the latter. But this is not the case if there are other harms than death that should count as well, or a questionable distribution of the harmful outcomes for the latter case (as in the staff in the novel nuclear plant), or if the epistemic security is much higher in the latter case, etc. – all arguably normative aspects. In this respect, risk and safety are unlike paradigmatic natural concepts such as gold. Even if people treated gold as something positive, as many do, this does not change the fact that only substances of a certain atomic number counts as gold: descriptive notions are all that matters. And it is unlike concepts such as BMI, or arguably even concepts as *fat* – on its quantitative notion at least – where descriptive aspects are all that counts for identifying something as having a higher BMI, or as being fatter.

On this reading, risk and safety seem to be of the same essential thick kind that philosophers such as John McDowell, Bernard Williams, and Jonathan Dancy argue cannot be separated into a natural part and a normative part, but must be treated as concepts that are *both* natural and normative, and for which no natural reduction can be given that suffices to pick out the extension of the terms. Currently, this anti-reductive camp is the dominant one in the debate about thick concepts, and also theorists such as Simon Blackburn who initially defended the reductive view (Blackburn 1984) have later abandoned it (Blackburn 1998). Even if Hare defended a reductivist strategy as late as in his 1997: 61, contemporary separatists argue instead for a weaker claim, namely, that a thick concept can be separated into a thin normative concept and a description, but that for the extension of a thick concept, the normative part is needed as well (Elstein and Hurka 2009). In other words, the output of recent moral philosophy is skepticism of the reductive claim for thick concepts such as risk and safety.

Risk, Objectivity, and Social Constructions

Let us now assume, in line with the arguments in the previous subsection, that risk and safety are thick normative concepts that cannot be reduced to natural ones. Is this bad news for proponents of the scientist or psychological approach to risk? Many theorists in the risk debate seem to make a close connection, if not equivalence, between a notion's status as natural or descriptive, and its being an objective concept. If a concept is not a natural concept, they seem to assume it is subjective rather than objective. It is important to note, however, that this fails to appreciate the large debate in philosophy concerning the status of the field of the normative, in particular the moral domain. In fact, many of the main proponents of the irreducible normativity of thick concepts mentioned above, for example, McDowell and Dancy, are moral realists who believe that there are objective truths in moral discourse. Hence, it is important to notice that claiming that risk is normative is not the same thing as claiming that it is non-objective.

In this subsection, I will investigate a common critique of the notions of risk and safety that seem to claim something stronger than mere normativity, namely, that these notions are essentially subjective or socially constructed. Objectivity and reality are philosophical questions that run deep very fast, and it would take us too far to argue for any particular position in this chapter. The current aim is rather to clarify the extent to which claims of subjectivity of risk, or even the non-reality of risk, are strong and radical claims.

Many theorists from the social sciences that approach risk on what I have called the *cultural perspective* argue that risk is a social construct, which is often labeled as a subjective, social feature rather than a feature of objective reality. People from different cultures as well as within one culture have very different views on what constitutes a risk and how severe a risk may be, they point out, and they go on to claim that there is no fact of the matter over and above these individual or cultural views (Douglas and Wildavsky 1982; Wynne 2000; Slovic 2000). Judith Bradbury, for example, describes the cultural view as the view that risk is conceived as "a socially constructed attribute, rather than as a physical entity that exists independently of the humans who assess and experience its effects" (Bradbury 1989, p. 381).

It should be noted that subjectivity in Bradbury's sense is a rather radical view. It means that the risk of climbing the ascent of K2, or driving 150 km/h in New York City, depends on humans assessing these activities rather than on physical facts that exist independently of such activities. In order to see the radicalism in such a position, and question whether the basic insights of the social constructivist idea motivate such radicalism, let us distinguish between three senses in which something can be socially constructed.

Let us call the first sense of social construct (essential) *reference-dependence*. A social system is needed in order for such an entity to exist. The paradigmatic example here is money (Searle 1995): in order for a piece of paper, metal object, or other physical entity to be money, it has to play a certain role in a social system. Before humans invented the concept of money and established the system in which it had a place, there was no such thing as money. If we take away the system, if we moved away from that sort of economy, or if intelligent life ceased to exist, there would be no money anymore.

The second sense of social construct is weaker: *interactive-dependence* (Hacking 1999). An entity that is interactive-dependent is not essentially dependent on any social system in which it is conceptualized. Still, its conceptualization in a social context affects the entity thus categorized. *Woman*, for example, is such an interactive category. By effectively categorizing someone as a woman – or stop doing so – one is affecting the way in which the person is viewed by others and herself, which behavior is typical and expected, right and wrong. Before we had a *concept* of woman, there existed women, in the sense that the kind of entities existed that (sufficiently) fit the concept. Its introduction, however, has changed the very category: a basic insight of feminism is that categorizing someone as a woman affects how she conceives of herself, which is one of the important problems for societal change. It is why merely *formally* allowing women entrance to societal positions and roles that used to be restricted to men does not by itself establish a substantial shift in the social structure, even if there are no essential biological restrictions that prohibit such a shift.

The third sense of social construct is the weakest: *sense-dependence*. (The distinction between reference-dependence and sense-dependence is taken from Brandom (2002, pp. 194–195), although it is here used in a slightly different way.) The idea here is that the entity in question is neither interactive-dependent nor essentially reference-dependent on its conceptualization and the social system in which it figures. However, our being able to *grasp* the concept and thus classify objects and properties in the world in accordance with it is dependent on a social context in which it has evolved. An example is the concept of a quark. The concept of a quark is dependent on the modern development of physics. While the concept of a quark was developed independently by Murray Gell-Mann and George Zweig in 1964, their being able to develop it was of course the fruit of social arrangements – ideas and institutions – throughout centuries. Indeed, all complex concepts such as those of theoretical entities in natural science theories are the fruit of social construction in this sense.

This third sense of the socially constructed is by far the weakest. Most philosophers of language would admit that *any* concept is socially constructed in the sense of having been developed in a social environment, for the simple reason that language *in total* is a social construct. Now, the extent of that sociality, and how necessary – as opposed to merely historically contingent – it is for language to be thus developed, may be disputed, but it seems to be an undeniable fact that the ability to express claimable content to other beings, to categorize objects and events in order to refer to them by language, is a social affair. How can it not be? Still, there is certainly a difference between the level of social exchange that is needed for different concepts. The concept of quarks need a bit more social lever – the development of advanced physics – in order to be constructed than, for example, the concept of food, which is closely connected to basic human needs.

Equipped with these three senses of social constructs, the one that most clearly fits Bradbury's description of risk as a "socially constructed attribute, rather than as a physical entity that exists independently of the humans who assess and experience its effects" is the first one, the essential reference-dependence. In the two other senses it seems (and I have assumed in the explanation of them) that while the *concept* is a social construct, the *entity* may exist independent of the humans that assess and experience its effects. It does make sense to state that women – and men, and children, and humans in general – existed prior to the construction of a concept to sort them out as such. Or, at the very least, further argument is needed in order to show why this is not so. For reference-dependent concepts such as money, it is the other way around: it makes little sense to claim that money existed before the conceptual practice involving it was in place.

In the debate, the range of claims sorted under the banner of social constructivism is very broad. Often in practice the core claim seems to be something closest to the third sense of social construct, which primarily focuses not on the entity as such but on the *concept* we have developed. For example, take Summerton and Berner's description of the main premise of the cultural approach:

- What is identified as a risk is not based on an a priori "fact" as identified by technical or medical experts; neither are risks merely rule-of-thumb assessments made by isolated individuals. The ways in which individuals – including experts – interpret risk can instead be seen as an expression of socially located beliefs and world views that to a large extent stems from the individual's situated position and experiences within social hierarchies, institutions and groups (Summerton and Berner 2003, pp. 6–7).

Their claim is much less radical than Bradbury's, since they explicitly limit themselves to what is *identified* as a risk by individuals, claiming first that it is not the same as what experts take it to be, and second that it is, for all individuals including experts, to a large extent a consequence of her social environment. None of this is a denial of the possibility of risks *existing* independently of individuals experiencing them.

Indeed, some theorists make clear that they are not arguing against objectivity per se. Clarke and Short point out that "objects might be dangerous in some objective sense" (Clarke and Short 1993, p. 79). The basic task on the cultural view, however, "is to explain how social agents create and use boundaries to demarcate that which is dangerous" (Clarke and Short 1993). In other words, the motive for making a claim of social construction in a domain can be to underline an explanatory focus: that of investigating how social circumstances, such as organizational environments, social hierarchies or other aspects of our socio-cultural context, play a part in how we conceptualize our world as well as what we take to be the extension of those concepts.

This explanatory task may of course be of paramount importance for how we should arrange our investigations into risks, and how to make decisions about them. Furthermore, it may convincingly show how our conceptions feed back into the phenomenon itself, that is, give detailed insights into the interactive-dependence of risk and safety. (Note that risk and safety are only indirectly interactive-dependent, since *we* are the ones affected by something's being categorized as risky – thus possibly changing the risk at hand – not the risk-entity itself.)

Importantly, however, neither sense-dependence nor interactive-dependence entails subjectivity. There may still perfectly well be an objective fact of the matter whether it is safe to climb K2, even if the actual safety depends on how we have perceived of it, as mediated, for example, by its manifestations in our training and preparation before attempting it, or in how we react to things that happen while climbing. Hence, even if risk and safety are both sense-dependent and interactive-dependent concepts, they can still be objective concepts corresponding to real properties in the world.

There are, however, claims that are made under the heading of social constructivism that deny the reality of risk (e.g., Ewald 1991, p. 199; Dean 1999, p. 177). Also a risk psychologist such as Paul Slovic makes this inference from constructivism to non-realism in a rather recent paper when he contrasts danger with risk, claiming that "danger is real, but risk is socially constructed." (Slovic 1999, p. 689) It is hard, though, to see the motivation for this denial of reality status for risk. Slovic's justification is representative: "Risk assessment is inherently subjective and represents a blending of science and judgment with important psychological, social, cultural, and political factors" (Slovic 1999). But even if we grant that risk *assessment* is subjective – for example, that it has uncorrected biases from cultural and political factors that result in faulty norms – this does not mean that risks cannot be real. Note that even if the *assessment* of our oil reserves are biased by wishful thinking as well as political and economic agendas, in addition to pure scientific uncertainty in the current models of estimation, it does not follow that oil is not real (or that Eve can have less oil in the tank of her car than Jones has in his).

If no further arguments are presented, it seems that the claims of subjectivity and anti-realism are unjustified inferences from sound insights into the societal nature of concept grasping and assessments of phenomena. In Slovic's claim in particular, this impression is enhanced from the seemingly arbitrary demarcation line between risk and danger. What constitutes a danger (or hazard) seems to be just as connected to human affairs as risk. So if danger is real, why not risk?

Lastly, we should note that even the strongest version of dependence I have mentioned, essential reference-dependence, in itself entails neither subjectivity nor anti-realism. Take the concept of money. The common wisdom is that money is indeed real, and claims concerning them may be objectively true or false, even if it is also true that the existence of potential money facts such as "My mobile phone is now worth roughly $200 less than when I bought it" or "I sold my car for $2,000" is dependent on a complex mix of social arrangements.

Naturally, none of the distinctions made in this subsection are by themselves an *argument* for the objectivity or reality of risk. Rather, the aim has been to distinguish between some of the different types of claims that the overall label of social construction may amount to, and to make plausible that denial of objectivity of risk and – even more – the reality of risk lies at the far end of this spectrum.

Further Research

As emphasized at the outset, risk and safety research is a heterogeneous field, and there are many interesting areas for research.

Analysis of Key Concepts

Mapping different risk and safety conceptions. In this chapter, several common interpretations of risk and safety have been discussed. Although it is well known that different people use the key concepts of risk research differently, a better mapping of the differences in usage of "risk" and "safety" in various fields is much needed. In particular, this is important for the psychological research where people's attitudes toward risks are the target, and where the terms are generally less clearly defined than in technical areas.

Risk versus safety. As indicated in this chapter, risk and safety are frequently treated as antonyms. While this may often be harmless, this treatment is not always viable. For example, it hides a difference in focus between risk research and safety research that should be more clearly explored.

Uncertainty. Uncertainty is an aspect of (lack of) safety that is sometimes hidden by the focus on handling the risk. Epistemic uncertainty is an important aspect of safety ascriptions, but, as emphasized in this chapter, there is large controversy as to how this notion is to be measured. Moreover, in many areas of risk and safety, the traditional method of coping with uncertainty is by an increased emphasis on probabilistic safety analysis (PSA). Sometimes, as in for example the nuclear industry, this has amounted to a very high level of descriptive sophistication. Even on this sophisticated level of PSA description, however, there are residual uncertainties, uncertainties that cannot be captured by PSA alone. Finding ways of handling these remaining uncertainties is an important task.

Safety versus security. While the concept of security is strongly interrelated with the concept of safety, there are important differences in the actual *practices* of the two fields. For example, security measures are more focused on protection against intentional threats, whereas safety measures have been directed against non-intentional harmful events as well as utilizing, to a greater extent, probabilistic methods. Investigating the differences and similarities in how safety and security are understood can be an important way to develop our understanding of both notions, and, consequently, develop our methods of ensuring safety and security.

Safety culture. The notion of a safety culture is studied by the social and behavioral sciences and employed by safety professionals. But it is rather unclear what a safety culture consists in. Is it an evaluative concept, and what counts as a good or bad safety culture? Can an organization lack a safety culture, or is that equal to a bad one? And what does "culture" signify in the context? Since it is considered an important "tacit" aspect in reaching safety, more conceptual clarity about the notion is called for.

The status of risk and safety. In this chapter, we touched on the complicated issues about the status of risk and safety. In addition to the overarching questions of naturalness, normativity, and constructivism, there are also interesting questions of subjectivity and objectivity when it comes to certain aspects of risk and safety. For harm and probability in particular, there are both subjective and objective interpretations. How these interpretations are best understood, and how this matters for risk and safety research, are pressing interdisciplinary concerns.

Further Issues

Related to these conceptual investigations, there are methodological, epistemic, and ethical issues for future research. Examples of such topics include:

Knowledge: implicit and explicit. Although effort is made in articulating explicit, propositional accounts of risk and safety, a significant amount of knowledge of risk and safety remains tacit in an organization. In connection with concepts such as uncertainty and safety culture, the relation between tacit and explicit knowledge of risk and safety should be investigated.

The sequential process approach. The viability of the common sequential process in safety analysis of starting with a scientific assessment of the risk, followed by an evaluative step in which the overall safety decision is made, depends on the viability of suitably isolating the descriptive, scientific aspects of the risk at hand. Any limitations to this approach have methodological consequences, not only for how we should go about gaining acceptance for a certain risk, or how we should compare risks to the advantages of the technology causing the potential risk, but for the actual enterprise of gaining *knowledge* of the risks involved. A common assumption in risk analysis is that ethical aspects come as merely an *addition* to the risk assessment, something to be considered in risk management only – an assumption that may be questioned, as in the arguments in this chapter that evaluative aspects surfaces already "inside" the very assessment of the risk. An important area of future research involves investigating the consequences of the internal as well as external presence of ethical aspects in risk assessment and management.

New versus old. Trying out new methods and techniques is paramount for technological progress. Yet, this means by definition less tested in actual circumstances, which often implies a large epistemic uncertainty. The circumstances when it is ethically correct to use new designs in potentially dangerous facilities is an interesting question in need of further research, both in general and in the many applied areas of risk and safety research.

Control versus integrity. Another trade-off of interest is the one between on the one hand surveillance and detailed supervision or control, and on the other, integrity and autonomy at the workplace. Here, what is deemed most safe and what is deemed as respecting the right to integrity and autonomy may not go hand in hand, and what to do with cases when these ideals clash is a central moral question in risk and safety ethics.

Disagreement. As indicated throughout this chapter, there are many different conceptions of safety, as well as many different ways of assessing the very same conception (such as the expected value conception). How should we best treat disagreement about risk and safety? Some disagreements are about scientific issues proper, and these seem reasonable to solve within the scientific community. Also scientists disagree, however, and sometimes on a level that is significantly relevant for the risk assessment. What should we do in such circumstances? But even when there is no disagreement of the underlying natural properties, we may disagree about the normative aspects of risk and safety, as well as about further ethical considerations. How we best treat such different cases of disagreement is a pressing topic for future study.

Conclusion

In this chapter, I have analyzed several aspects of the concepts of risk and safety. As I started by pointing out, there are many different – if overlapping – perspectives from which the questions of risk and safety are treated, as well as several interpretations of the terms "risk" and "safety" that are relevant in the context of societal decision-making. An evident, but still often neglected consequence of this heterogeneous set of interpretations is that we should always keep in mind the possibility of talking past each other in matters involving risk and safety. There are many viable ways of using the terms, and ascertaining that we are on the same track is a minimal condition for risk and safety communication.

The main focus in the chapter has then been on the quantitative notions of risk and safety. Central aspects such as probability, harm, and uncertainty have been treated. In particular, the controversial question of how risk and safety should be inferred from these aspects – and possibly even additional ones – should be kept in mind. Moreover, whether risk and safety are natural concepts or contain normative aspects that cannot be reduced to natural ones are important questions for how we should understand these concepts. Finally, I pointed out that even if risk and safety are normative concepts – as I indeed believe they are – this is not equal to the claim that risk and safety cannot be objective or that they are constructed in any deep sense. Still, risk and safety are central concepts in societal decision-making, and both scientific and extra-scientific concerns are vital if we are to make the best decisions when risks are involved.

Acknowledgments

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 237590.

References

Asveld L, Roeser S (eds) (2009) The ethics of technological risk. Earthscan, London

Aven T (2003) Foundations of risk analysis: a knowledge and decision-oriented perspective. Wiley, Chichester

Aven T (2007) A unified framework for risk and vulnerability analysis and management covering both safety and security. Reliab Eng Syst Safe 92:745–754

Aven T, Renn O (2009) On risk defined as an event where the outcome is uncertain. J Risk Res 12:1–11

Baron J (1987) Second-order probabilities and belief functions. Theory Decision 23:25–36

Blackburn S (1984) Spreading the word. Clarendon, Oxford

Blackburn S (1998) Ruling passions: a theory of practical reasoning. Clarendon, Oxford

Boyd R (1991) Realism, anti-foundationalism and the enthusiasm for natural kinds. Phil Stud 61: 127–148

Bradbury J (1989) The policy implications of differing concepts of risk. Sci Technol Hum Values 14: 380–399

Brandom R (2002) Tales of the mighty dead: historical essays in the metaphysics of intentionality. Harvard University Press, Cambridge, MA

Campbell S (2005) Determining overall risk. J Risk Res 8:569–581

Clarke L, Short J (1993) Social organization and risk: some current controversies. Annu Rev Sociol 19: 375–399

Cohen BL (2003) Probabilistic risk analysis for a high-level radioactive waste repository. Risk Anal 23: 909–915

Dancy J (1995) In: Defence of thick concepts, in French, Uehling, Wettstein (eds), Midwest studies in philosophy 20. University of Notre Dame Press, Notre Dame, Ind

Dancy J (2004) Ethics without principles. Clarendon, Oxford

de Finetti B (1937) La prevision: see lois logiques, ses sources subjectives, Annales de l'Institut Henri Poincaré 7

Dean M (1999) Governmentality: power and rule in modern society. Sage, London

Douglas EJ (1983) Managerial economics: theory, practice and problems, 2nd edn. Prentice Hall, Englewood Cliffs

Douglas M, Wildavsky A (1982) Risk and culture: an essay on the selection of technological and environmental dangers. University of California Press, Berkeley

Dubois D, Prade H (1988) Decision evaluation methods under uncertainty and imprecision. In: Kacprzyk J, Fedrizzi M (eds) Combining fuzzy impression with probabilistic uncertainty in decision making. Springer, Berlin, pp 48–65

EC (2003) Technical guidance document in support of commission directive 93/67/EEC on Risk Assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Joint Research Centre, EUR 20418 EN, Office for Official Publications of the EC, Luxembourg

Ellsberg D (1961) Risk, ambiguity, and the savage axioms. Quart J Econ 75:643–669

Elstein D, Hurka T (2009) From thick to thin: two moral reduction plans. Canad J Philos 39:515–536

Ewald F (1991) Insurance and risk. In: Burchell G, Gordon C, Miller P (eds) The Foucault effect: studies in governmental rationality. Harvester Wheatsheaf, Hemel Hempstead

Fischhoff B, Slovic P, Lichtenstein S, Read S, Combs B (1978/2000) How safe is safe enough? A psychometric study of attitudes toward technological risk and benefits. In: Slovic P (ed) The perception of risk. Earthscan, London, pp 80–103

Gärdenfors P, Sahlin N-E (1982/1988) Unreliable probabilities, risk taking, and decision making. In: Gärdenfors P, Sahlin N-E (eds) Decision, probability, and utility. Cambridge University Press, Cambridge, pp 313–334

Graham JD, Weiner JB (eds) (1995) Risk versus risk: tradeoffs in protecting health and the environment. Harvard University Press, Cambridge, MA

Hacking I (1991) A tradition of natural kinds. Phil Stud 61:109–126

Hacking I (1999) The social construction of what? Harvard University Press, Cambridge, MA

Hansson SO (1998) Setting the limit. Occupational health standards and the limits of science. Oxford University Press, Oxford

Hansson SO (2004a) Philosophical perspectives on risk. Techne 8:1

Hansson SO (2004b) Weighing risks and benefits. Topoi 23:145–152

Hansson SO (2005) Seven myths of risk. Risk Manage 7:7–17

Hansson SO (2009) Risk and safety in technology. In: Meijers A (ed) Handbook of the philosophy of science, vol 9: philosophy of technology and

engineering sciences. Elsevier, Amsterdam, pp 1069–1102
Hansson SO (2010) Risk: objective or subjective, facts or values. J Risk Res 13:231–238
Hare RM (1952) The language of morals. Clarendon, Oxford
Hare RM (1997) Sorting out ethics. Clarendon, Oxford
Howson C, Urbach P (2006) Scientific reasoning: the Bayesian approach, 3rd edn. Open Court Publishing Company, Chicago
IAEA (2000) Safety of nuclear power plants: design. International Atomic Energy Agency, Vienna
ISO (2002) Risk management – vocabulary – guidelines for use in standards. ISO/IEC Guide 73:2002
Kaplan M (1983) Decision theory as philosophy. Phil Sci 50:549–577
Knight FH (1921) Risk, uncertainty and profit. BoardBooks, Washington, DC (Reprinted 2002)
Kraus N, Malmfors T, Slovic P (1992/2000) Intuitive toxicology: experts and lay judgements of chemical risks. In: Slovic P (ed) The perception of risk. Earthscan, London, pp 285–315
Kuhn T (1962) The structure of scientific revolutions. University of Chicago Press, Chicago
Kyburg HE (1968) Bets and beliefs. Am Phil Q 5:54–63
Lakatos I, Musgrave A (eds) (1970) Criticism and the growth of knowledge. Cambridge University Press, London
Lange M (ed) (2007) Philosophy of science: an anthology. Blackwell, Malden
Le Grand J (1991) Equity and choice: an essay in economics and applied philosophy. Harper Collins Academic, London
Levi I (1974) On indeterminate probabilities. J Phil 71:391–418
Levi I (1986) Hard choices: decision making under unresolved conflict. Cambridge University Press, Cambridge
Lewens T (2007) Risk: philosophical perspectives. Routledge, New York
Little M (1994) Moral realism II: non-naturalism. Phil Books 35:225–232
Lowrance W (1976) Of acceptable risk – science and the determination of safety. William Kaufmann, Los Altos
Mayo D (1991) Sociological versus metascientific views of risk assessment. In: Mayo D, Hollander R (eds) Acceptable evidence, science and values in risk management. Oxford University Press, Oxford, pp 249–280
McDowell J (1978) Are moral requirements hypothetical imperatives? In: Proceedings of the Aristotelian society, Supplementary volume 52, pp 13–29
McDowell J (1979) Virtue and reason. Monist 62:331–350
McDowell J (1981) Non-cognitivism and rule-following. In: Holtzman S, Leich C (eds) Wittgenstein: to follow a rule. Routledge/Kegan Paul, London/Boston, pp 141–162
McMullin E (1982) Values in science. In: PSA: proceedings of the biennial meeting of the philosophy of science association, 2:3–28
Miller CO (1988) System safety. In: Wiener EL, Nagel DC (eds) Human factors in aviation. Academic, San Diego, pp 53–80
Misumi Y, Sato Y (1999) Estimation of average hazardous-event-frequency for allocation of safety-integrity levels. Reliab Eng Syst Safe 66:135–144
Möller N (2009a) Thick concepts in practice: normative aspects of risk and safety. Royal Institute of Technology, Stockholm
Möller N (2009b) Should we follow the experts' advice? Epistemic uncertainty, consequence dominance and the knowledge asymmetry of safety. Int J Risk Assess Manage 11:219–236
Möller N (2010) The non-reductivity of normativity in risks. In: de Vries MJ, Hansson SO, Meijers AWM (eds) Norms and the artificial: moral and non-moral norms in technology (Forthcoming)
Möller N, Hansson SO (2008) Principles of engineering safety: risk and uncertainty reduction. Reliab Eng Syst Safe 93:776–783
Möller N, Hansson SO, Peterson M (2006) Safety is more than the antonym of risk. J Appl Philos 23:419–432
Moore GE (1903) Principia ethica. Cambridge University Press, Cambridge
National Research Council (1983) Risk assessment in the federal government managing the process. National Academy Press, Washington, DC
Otway H (1987) Experts, risk communication, and democracy. Risk Anal 7:125–129
Peterson M (2002) What is a de minimis risk? Risk Manage 4:47–55
Ramsey F (1931) Truth and probability. In: Braithwaite RB (ed) The foundations of mathematics and other logical essays. Routledge & Kegan Paul, London, pp 156–198. (Reprinted in: Gärdenfors P, Sahlin N-E (eds). Decision, probability, and utility. Cambridge University Press, Cambridge, pp 19–47)
Rechard RP (1999) Historical relationship between performance assessment for radioactive waste disposal and other types of risk assessment. Risk Anal 19:763–807
Renn O (1992) Social theories of risk. In: Krimsky S, Golding D (eds) Social theories of risk. Praeger, Westport
Renn O (2005) Risk governance: towards an integrative approach. White paper no. 1, written by Ortwin Renn with an Annex by Peter Graham. International Risk Governance Council, Geneva

Resnik M (1987) Choices: an introduction to decision theory. University of Minnesota Press, Minneapolis

Rosa EA (1998) Metatheoretical foundations for post-normal risk. J Risk Res 1:15–44

Savage L (1972/1954) The foundations of statistics, 2nd edn. Dover, New York

Searle J (1995) The construction of social reality. The Free Press, New York

Sen A (1987) On ethics and economics. Blackwell, Oxford

Shrader-Frechette K (1993) Burying uncertainty: risk and the case against geological disposal of nuclear waste. University of California Press, Berkeley

Skyrms B (1980) Higher order degrees of belief. In: Mellor DH (ed) Prospects for pragmatism. Cambridge University Press, Cambridge, pp 109–137

Slovic P (1999) Trust, emotion, sex, politics, and science: surveying the risk-assessment battlefield. Risk Anal 19:689–701

Slovic P (2000) The perception of risk. Earthscan, London

Soames S (2005) Philosophical analysis in the twentieth century: the age of meaning. Princeton University Press, Princeton

Stevenson C (1944) Ethics and language. Yale University Press, New Haven

Summerton J, Berner B (2003) Constructing risk and safety in technological practice: an introduction. In: Summerton J, Berner B (eds) Constructing risk and safety in technological practice. Routledge, London, pp 1–19

Tench W (1985) Safety is no accident. Collins, London

Unwin S (1986) A fuzzy set theoretic foundation for vagueness in uncertainty analysis. Risk Anal 6:27–34

Vallentyne P (1998) The nomic role account of carving reality at the joints. Synthese 115:171–198

von Neumann J, Morgenstern O (1944) Theory of games and economic behavior. Princeton University Press, Princeton

Wandall B (2004) Values in science and risk assessment. Toxicol Lett 152:265–272

Williams B (1985) Ethics and the limits of philosophy. Harvard University Press, Cambridge, MA

Willis HH (2007) Guiding resource allocations based on terrorism risk. Risk Anal 27:597–606

Wolf M (2002) The curious role of natural kind terms. Pacific Phil Q 83:81–101

Wynne B (1982) Institutional mythologies and dual societies in the management of risk. In: Kunreuther H, Ley E (eds) The risk analysis controversy. Springer, Berlin, pp 127–143

Wynne B (1992) Carving out science (and politics) in the regulatory jungle. Soc Stud Sci 22:745–758