

# The Epistemology of Technological Risk

Sven Ove Hansson

soh@kth.se

Department of Philosophy and the History of Technology,  
Royal Institute of Technology  
Stockholm, Sweden

## 1. Introduction

Beginning in the late 1960's, growing public concern with new technologies gave rise to a new field of applied science: the study of risk. Researchers from the technological, natural, behavioural, and social sciences joined to create a new interdisciplinary subject, risk analysis. (Otway 1987) The aim of this discipline is to produce as exact knowledge as possible about risks. Such knowledge is much in demand, not least among managers, regulators, and others who decide on the choice and use of technologies.

But to what extent can risks be known? Risks are always connected to lack of knowledge. If we know for certain that there will be an explosion in a factory, then there is no reason for us to talk about that explosion as a risk. Similarly, if we know that no explosion will take place, then there is no reason either to talk about risk. What we refer to as a risk of explosion is a situation in which it is not known whether or not an explosion will take place. In this sense, knowledge about risk is knowledge about the unknown. It is therefore a quite problematic type of knowledge.

## 2. Risk as a quantitative concept

Although many have tried to make the concept of risk as objective as possible, on a fundamental level it is an essentially value-laden concept. More precisely, it is negatively value-laden. Risks are unwanted phenomena. The tourist who hopes for a sunny week talks about the "risk" of rain, but the farmer whose crops are threatened by drought will refer to the possibility of rain as a "chance" rather than a "risk."

There are more distinctions to be made. The word “risk” is used in many senses, that are often not sufficiently distinguished between. Let us consider four of the most common ones.

(1) risk = an *unwanted event* which may or may not occur.

This is how we use the word “risk” when we say that lung cancer is one of the major risks that affect smokers, or that aeroembolism is a serious risk in deep-sea diving.

However, we may also describe smoking as a (health) risk or tell a diver that going below 30 meters is a risk not worth taking. We then use the word “risk” to denote the event that caused the unwanted event, rather than the unwanted event itself:

(2) risk = the *cause* of an unwanted event which may or may not occur.

In addition to identifying the risks, i.e. finding out what they are, we also want to determine how big they are. This is usually done by assigning to each risk a numerical value that indicates its size or seriousness. The numerical value most often used for this purpose is the probability of the event in question

Indeed, risks are so strongly associated with probabilities that we often use the word “risk” to denote the probability of an unwanted event rather than that event itself. This terminology is particularly common in engineering applications. When a bridge-builder discusses the risk that a bridge will collapse, or an electrical engineer investigates the risk of power failure, they are almost sure to use “risk” as a synonym of probability.

(3) risk = the *probability* of an unwanted event which may or may not occur.

As should be clear from the three definitions presented thus far, the word “risk” is a very versatile – and ambiguous – one. Consider a situation in which the probability is 10% that a certain person will contract lung cancer due to smoking. We can then use the word “risk” to denote either the cancer, the act of smoking, or the 10% probability.

It is important to note that probability, and hence risk in the third sense, always refers to a specified event or type of event. If you know the probability (risk) of

power failure, this does not mean that you have a total overview of the possible negative events (risks) associated with the electrical system. There may be other such events, such as fires, electrical accidents etc., each with their own probabilities. In risk analysis, a concept of risk has been developed that aims at quantifying the total amount of risk associated with a technological system:

- (4) risk = the statistical *expectation value* of unwanted events which may or may not occur.

Expectation value means probability-weighted value. Hence, if 200 deep-sea divers perform an operation in which the individual risk of death is 0.1% for each individual, then the expected number of fatalities from this operation is  $200 \times 0.1\% = 0.2$ .

Expectation values have the important property of being additive, i.e. they can be added. Suppose that a certain operation is associated with a 1% probability of an accident that will kill five persons, and also with a 2% probability of another type of accident that will kill one person. Then the total expectation value is  $0.01 \times 5 + 0.02 \times 1 = 0.07$  deaths. In similar fashion, the expected number of deaths from a nuclear power plant is equal to the sum of the expectation values for each of the various types of accidents that can occur in the plant. (Bondi 1985, p. 9) The following is a typical example of the jargon:

The worst reactor-meltdown accident normally considered, which causes 50 000 deaths and has a probability of  $10^{-8}$ /reactor-year, contributes only about two per cent of the average health effects of reactor accidents. (Cohen 1985, 1)

In risk-benefit analysis, i.e. the systematic comparison of risks with benefits, risks are measured in terms of their expectation values. This is also the standard meaning of risk in several branches of risk research. Hence, in studies of risk perception the standard approach is to compare the degree of severity that subjects assign to different risk factors (“subjective risk”) to the expectation values that have been calculated for the same risk factors (“objective risk”). The underlying assumption is that there is an objective, knowable risk level that can be calculated with the expectation value method.

However, this measure of risk is problematic for at least two fundamental reasons. First, probability-weighting is normatively controversial. A risk of 1 in

1000 that 1000 persons will die is very different from a risk of 1 in 10 that 10 persons will die. Although the expectation values are the same, moral reasons can be given to regard one of these two situations as more serious than the other. In particular, proponents of a precautionary approach maintain that prevention against large but improbable accidents should be given a higher priority than what would ensue from an expectation value analysis. (O’Riordan and Cameron 1994, O’Riordan et al 2001)

The other problem with the expectation value approach is that it assesses risks only according to their probabilities and the severity of their consequences. Most people’s appraisals of risks are influenced by factors other than these. In particular, the expectation value method treats risks as impersonal entities, and pays no attention to how risks and benefits are distributed or connected. In contrast, the relations between the persons affected by risks and benefits are important in most people’s appraisals of risks. It makes a difference if it is myself or someone else that I expose to a certain danger in order to earn myself a fortune. If the expectation value is the same in both cases, we can arguably say that the size of the risk is the same in both cases. It does not follow, however, that the risk is equally acceptable in the two cases. More generally speaking, if we use expectation values to measure the size of risks, this must be done with the reservation that the size of a risk is not all that we need to in order to judge whether or not it can be accepted. Additional information about its social context is also needed.

### **3. The tuxedo syndrome**

From a decision-maker’s point of view, it is useful to have risks quantified so that they can easily be compared and prioritized. But to what extent can quantification be achieved without distorting the true nature of the risks involved?

As we have just seen, probabilities are needed for both of the common types of quantification of risk (the third and fourth senses of risk). Without probabilities, established practices for quantifying risks cannot be used. Therefore, the crucial issue in the quantification of risk is the determination of probabilities.

Not all dangers come with probabilities assigned to them. In decision theory, the terms “risk” and “uncertainty” are used to distinguish between those that do and those that do not. By a decision “under risk” is meant a decision with known or

knowable probabilities. By a decision “under uncertainty” is meant one that has to be taken with unknown probabilities. In one of the most influential textbooks in decision theory, the terms are defined as follows:

We shall say that we are in the realm of decision-making under:

(a) *Certainty* if each action is known to lead invariably to a specific outcome (the words prospect, stimulus, alternative, etc., are also used).

(b) *Risk* if each action leads to one of a set of possible specific outcomes, each outcome occurring with a known probability. The probabilities are assumed to be known to the decision maker. For example, an action might lead to this risky outcome: a reward of \$10 if a ‘fair’ coin comes up heads, and a loss of \$5 if it comes up tails. Of course, certainty is a degenerate case of risk where the probabilities are 0 and 1.

(c) *Uncertainty* if either action or both has as its consequence a set of possible specific outcomes, but where the probabilities of these outcomes are completely unknown or are not even meaningful. (Luce and Raiffa 1957, p. 13)

This usage of the key terms “risk” and “uncertainty” differs distinctly from everyday usage. In everyday conversations, we would not hesitate to call a danger a risk although there are no means to determine its probability. By uncertainty we would mean a state of mind rather than the absence of information. However, these are well-established technical terms, and in what follows, they will be used in their standard technical meanings.

The gambler’s decisions at the roulette table are clear examples of decisions under risk, i.e. decisions with known probabilities. Given that the wheel is fair, the probabilities of various outcomes – gains and losses – are easily calculable, and thus knowable, although the gambler may not take them into account. For a clear example of a decision under uncertainty, consider instead the decision of an explorer to enter a distant part of the jungle, previously untrod by human foot. There are tigers and poisonous snakes in the jungle, but no estimates better than guesses can be given of the probability of being attacked by them. Such attacks are known dangers with unknown probabilities. In addition, it is reasonable to expect that the jungle contains a large number of other species – from microorganisms to large animals – some of which may be dangerous although they are completely unknown. Not only their probabilities but their very existence is unknown. (Unknown dangers are not included in Luce and Raiffa’s

definition of uncertainty, as quoted above, but in subsequent discussions the concept of uncertainty has been extended to include them.)

In real life we are seldom in a situation like that at the roulette table, when all probabilities are known with certainty (or at least beyond reasonable doubt). Most of the time we have to deal with dangers without knowing their probabilities, and often we do not even know what dangers we have ahead of us. This is true not least in the development of new technologies. The social and environmental effects of a new technology can seldom be fully grasped beforehand, and there is often considerable uncertainty with respect to the dangers that it may give rise to. (Porter 1991)

Risk analysis has, however, a strong tendency towards quantification. Risk analysts often exhibit the *tuxedo syndrome*: they proceed as if decisions on technologies were made under conditions analogous to gambling at the roulette table. In actual fact, however, these decisions have more in common with entering an unexplored jungle. The tuxedo syndrome is dangerous since it may lead to an illusion of control. Risk analysts and those whom they advise may believe that they know what the risks are and how big they are, when in fact they do not.

#### **4. Unknown probabilities**

When there is statistically sufficient experience of an event-type, such as a machine failure, then we can determine its probability by collecting and analysing that experience. Hence, if we want to know the probability that the airbag in a certain make of car fails to release in a collision, we should collect statistics from the accidents in which such cars were involved.

For new and untested technologies this method is not available. Due to the lack of accident statistics, it cannot be used to determine the probability of airbag failure in a new car model that is just to be introduced. If the construction is essentially unchanged since previous models, then we may rely on statistics from these earlier models, but this is not advisable if there have been significant changes.

Even after many years' experience of a technology there may be insufficient data to determine the frequencies of unusual types of accidents or failures. As one example of this, there have (fortunately) been too few severe accidents in nuclear

reactors to make it possible to estimate their probabilities. In particular, most of the reactor types in use have never been involved in any serious accident. It is therefore not possible to determine the risk (probability) of a severe accident in a specified type of reactor.

One common way to evade these difficulties is to calculate the probability of major failures by means of a careful investigation of the various chains of events that may lead to such failures. By combining the probabilities of various subevents in such a chain, a total probability of a serious accident can be calculated. Such calculations were in vogue in the 1970's and 1980's, but today there is a growing skepticism against them, due to several difficult problems with this methodology. One such problem is that accidents can happen in more ways than we can think of beforehand. There is no method by which we can identify all chains of events that may lead to a major accident in a nuclear reactor, or any other complex technological system.

Another problem with this methodology is that the probability of a chain of events can be very difficult to determine even if we know the probability of each individual event. Suppose for instance that an accident will happen if two safety valves both fail. Furthermore suppose that we have experience showing that the probability is 1 in 500 that a valve of this construction will fail during a period of one year. Can we then conclude that the probability that both will fail in that period is  $1/500 \times 1/500$ , i.e.  $1/250000$ ? Unfortunately not, since this calculation is based on the assumption that failures in the two valves are completely independent events. It is easy to think of ways in which they may be dependent: Faulty maintenance may affect them both. They may both fail at high temperatures or in other extreme conditions caused by a failure in some other component, etc. It is in practice impossible to identify all such dependencies and determine their effects on the combined event-chain.

In spite of these difficulties, the construction and analysis of such event chains (often called fault-trees) is not a useless exercise. To the contrary, this can be an efficient way to identify weaknesses in a complex technological system. It is important, though, to keep in mind that an exhaustive list of negative events cannot be obtained, and that therefore total risk levels cannot be determined in this way.

Technological risks depend not only on the behaviour of technological components, but also on human behaviour. Hence, the risks in a nuclear reactor

depend on how the personnel act both under normal and extreme conditions. Similarly, risks in road traffic depend to a large degree on the behaviour of motorists. It would indeed be difficult to find an example of a technological system in which failure rates depend exclusively on its technical components, with no influence from human agency. The risks associated with one and the same technology can differ drastically between organizations with different attitudes to risk and safety. This is one of the reasons why human behaviour is even more difficult to predict than the behaviour of technical components.

Humans come into probability estimates in one more way: It is humans who make these estimates. Unfortunately, psychological studies indicate that we have a tendency to be overly confident in our own probability estimates. In other words, we tend to neglect the possibility that our own estimates may be incorrect. It is not known what influence this bias may have on risk analysis, but it certainly has a potential to create errors in the analysis. (Lichtenstein 1982, 331) It is essential, for this and other reasons, to make a clear distinction between those probability values that originate in experts' estimates and those that are known through observed frequencies.

### **5. Unknown dangers**

There are occasions when decisions are influenced by worries about possible dangers although we have only a vague idea about what these dangers might look like. Recent debates on biotechnology are an example of this. Specific, identified risks have had only a limited role in these debates. Instead, the focus has been on vague or unknown dangers such as the creation of new life-forms with unforeseeable properties.

It is easy to find examples in which many of us would be swayed by considerations of unknown dangers. Suppose, for instance, that someone proposes the introduction of a genetically altered species of earthworm that will aerate the soil more efficiently. If introduced in nature, it will ultimately replace the common earthworm. For the sake of argument we may assume that all concrete worries have been neutralized. The new species can be shown not to induce more soil erosion, not to be more susceptible to diseases, etc. Still, it would not be irrational to say: "Yes, but there may be other negative effects that we have not been able to think of. Therefore, the new species should not be introduced." Similarly, if someone proposed to eject a chemical substance into the stratosphere for some good purpose or other, it would not be irrational to



oppose this proposal solely on the ground that it may have unforeseeable consequences, and this even if all specified worries can be neutralized.

However, it would not be feasible to take such possibilities into account in all decisions that we make. In a sense, *any* decision may have catastrophic unforeseen consequences. If far-reaching indirect effects are taken into account, then – given the unpredictable nature of actual causation – almost any decision may lead to a disaster. In order to be able to decide and act, we therefore have to disregard many of the more remote possibilities. Cases can also easily be found in which it was an advantage that far-fetched dangers were not taken seriously. One case in point is the false alarm on so-called polywater, an alleged polymeric form of water. In 1969, the prestigious scientific journal *Nature* printed a letter that warned against producing polywater. The substance might “grow at the expense of normal water under any conditions found in the environment,” thus replacing all natural water on earth and destroying all life on this planet. (Donahoe 1969) Soon afterwards, it was shown that polywater is a non-existent entity. If the warning had been heeded, then no attempts would have been made to replicate the polywater experiments, and we might still not have known that polywater does not exist. In cases like this, appeals to the possibility of unknown dangers may stop investigations and thus prevent scientific and technological progress.

We therefore need criteria to determine when the possibility of unknown dangers should be taken seriously and when it can be neglected. This problem cannot be solved with probability calculus or other exact mathematical methods. The best that we can hope for is a set of informal criteria that can be used to support intuitive judgement. The following list of four criteria has been proposed for this purpose. (Hansson 1996)

*1. Asymmetry of uncertainty:* Possibly, a decision to build a second bridge between Sweden and Denmark will lead through some unforeseeable causal chain to a nuclear war. Possibly, it is the other way around so that a decision not to build such a bridge will lead to a nuclear war. We have no reason why one or the other of these two causal chains should be more probable, or otherwise more worthy of our attention, than the other. On the other hand, the introduction of a new species of earthworm is connected with much more uncertainty than the option not to introduce the new species. Such asymmetry is a necessary but

insufficient condition for taking the issue of unknown dangers into serious consideration.

2. *Novelty*: Unknown dangers come mainly from new and untested phenomena. The emission of a new substance into the stratosphere constitutes a qualitative novelty, whereas the construction of a new bridge does not.

An interesting example of the novelty factor can be found in particle physics. Before new and more powerful particle accelerators have been built, physicists have sometimes feared that the new levels of energy might generate a new phase of matter that accretes every atom of the earth. The decision to regard these and similar fears as groundless has been based on observations showing that the earth is already under constant bombardment from outer space of particles with the same or higher energies. (Ruthen 1993)

3. *Spatial and temporal limitations*: If the effects of a proposed measure are known to be limited in space or time, then these limitations reduce the urgency of the possible unknown effects associated with the measure. The absence of such limitations contributes to the severity of many ecological problems, such as global emissions and the spread of chemically stable pesticides.

4. *Interference with complex systems in balance*: Complex systems such as ecosystems and the atmospheric system are known to have reached some type of balance, which may be impossible to restore after a major disturbance. Due to this irreversibility, uncontrolled interference with such systems is connected with a high degree of uncertainty. (Arguably, the same can be said of uncontrolled interference with economic systems; this is an argument for piecemeal rather than drastic economic reforms.)

It might be argued that we do not know that these systems can resist even minor perturbations. If causation is chaotic, then for all that we know, a minor modification of the liturgy of the Church of England may trigger a major ecological disaster in Africa. If we assume that all cause-effect relationships are chaotic, then the very idea of planning and taking precautions seems to lose its meaning. However, such a world-view would leave us entirely without guidance,

even in situations when we consider ourselves well-informed. Fortunately, experience does not bear out this pessimistic worldview. Accumulated experience and theoretical reflection strongly indicate that certain types of influences on ecological systems can be withstood, whereas others cannot. The same applies to technological, economic, social, and political systems, although our knowledge about their resilience towards various disturbances has not been sufficiently systematized.

## **6. The role of experts**

The distinction between risk and uncertainty relates to the epistemic situation of the individual person. If you are absolutely certain that current estimates of the effects of low-dose radiation are accurate, then decision-making referring to such exposure may appear to you as decision-making under risk. On the other hand, if you are less than fully convinced, then such a decision will have the characteristics of decision-making under uncertainty. Even when exact estimates of risk are available, to the extent that there is uncertainty about these estimates, there is also uncertainty involved in the decision.

The exact estimates of risks that are available in some fields of knowledge have in general been made by highly specialized experts. If these experts have full confidence in their own conclusions, then decisions based on these estimates are, from their own point of view, decisions under risk (known probabilities). However, the situation may look different from the viewpoints of managers, other decision-makers, and the general public. Experts are known to have made mistakes. A rational decision-maker should take into account the possibility that this may happen again. Suppose for instance that a group of experts have studied the possibility that a new micro-organism that has been developed for therapeutical purposes will mutate and become virulent. They have concluded that the probability that this will happen is 1 in 100 000 000. For the decision-makers who receive this report the crucial issue need not be whether or not a risk of that magnitude should be accepted. Instead it could be how certain the conclusion is.

Insufficient attention has been paid in risk management to the fallibility of experts. Experts often do not realize that for the non-expert, the possibility of the experts being wrong may very well be a dominant part of the risk (in the informal sense of the word) involved e.g. in the use of a complex technology. On some (but not all) occasions when experts have complained about the public's

irrational attitude to risks, the real difference is that the public has focused on uncertainties, whereas experts have focused on quantified risk estimates.

Clearly, when there is a wide divergence between the views of experts and those of the public, this is a sign of failure in the social system for division of intellectual labour. Such failures have not been uncommon in issues of technological risks. However, it should not be taken for granted that every such failure is located within the minds of the non-experts who distrust the experts. It cannot be a criterion of rationality that experts are taken for infallible. Other issues must be raised, such as whether or not experts have been able to deal satisfactorily with the concerns of the public.

## 7. Conclusion

We started out by noting that knowledge about risk has a mildly paradoxical flavour, since it is knowledge about the unknown. We have indeed found severe limitations in what knowledge can be obtained about technological risks. Technology is associated not only with quantifiable risks but also with non-quantifiable uncertainties that may be quite difficult to come to grips with. Unexpected events do happen, not least in technological systems.

For the engineer, this insight has important practical consequences. Even if risk estimates show that a fire in the plant is so improbable that sprinklers are uneconomical, it may be a good idea to install those sprinklers, since there may be unknown causes of fire that have not been included in the calculations. Even if a toxic gas is so well contained that the possibility of a leakage has been virtually excluded, measures should be taken to protect workers and others who may be affected in the (unlikely) event of a leakage.

This is, of course, not a new insight, but rather epistemological underpinnings for well-established principles in safety engineering, such as contingency planning and redundant safety barriers. Safety does not mean measures only against those hazards that are known and quantified, but also, as far as possible, against those that are unknown and unexpected.

## References

- Bondi, H. (1985). Risk in perspective, in *Risk*. Edited by M.G. Cooper. Oxford: Clarendon, 8-17.
- Cohen, B.L. (1985) Criteria for Technology Acceptability. *Risk Analysis* 5:1-3.

Donahoe, F.J. (1969) "Anomalous" Water. *Nature* 224:198.

Hansson, S.O. (1996) Decision-Making Under Great Uncertainty. *Philosophy of the Social Sciences* 26:369-386.

Lichtenstein, S. et al. (1982) Calibration of probabilities: The state of the art to 1980. In *Judgment Under Uncertainty, Heuristics and Biases*. Edited by Kahneman et al. Cambridge: Cambridge University Press, 306-334.

Luce, R.D. and H. Raiffa (1957) *Games and Decisions*. New York: Wiley.

O'Riordan, T. and J. Cameron, eds. (1994) *Interpreting the Precautionary Principle*. London: Earthscan.

O'Riordan, T., J. Cameron, and A. Jordan, eds. (2001) *Reinterpreting the Precautionary Principle*. London : Cameron.

Otway, H. (1987) Experts, Risk Communication, and Democracy. *Risk Analysis* 7:125-129.

Porter, A.L., et al. (1991) *Forecasting and Management of Technology*. New York: John Wiley & Sons.

Ruthen, R. (1993) Strange Matter. *Scientific American* (August): 10.