Safe Design

Sven Ove Hansson
Department of Philosophy and the History of Technology
Royal Institute of Technology, Stockholm
soh@infra.kth.se

Abstract: Safety is an essential ethical requirement in engineering design. Strategies for safe design are used not only to reduce estimated probabilities of injuries but also to cope with hazards and eventualities that cannot be assigned meaningful probabilities. The notion of safe design has important ethical dimensions, such as that of determining the responsibility that a designer has for future uses (and misuses) of the designed object.

Keywords: safety, risk, safe design, safety barrier, ethics.

1. Safety – An ethical issue in design

In the small literature that is available on the ethics of engineering design, there is consensus that safety is an essential ethical requirement. It is generally agreed that designers have an ethical responsibility to make constructions that are safe in future use. However, it is far from clear how far this responsibility extends. It needs to be specified in at least two respects.

The first of these consists in answering the question "*safe against what?*" Safety is concerned with avoiding certain classes of events that it is morally right to avoid. In engineering design, safety always includes safety against unintended human death or injuries that occur as a result of the intended use of the designed object. But does it include the avoidance of accidents in foreseeable but unintended uses of the object? Does it include protection against malevolent use of the object by criminals or terrorists? (Kemper 2004) The prevention of long-term health effects? The prevention of damage to the environment?

We can use the design of bridges as an example of this problem. Designers of bridges are normally held responsible for the structural reliability of their constructions. If a bridge collapses, then we hold the engineers who designed it responsible. However, there are other types of safety issues in connection with bridges. Accidents happen when people climb and walk on arches, dive from the bridge, or throw objects on ships or vehicles passing below the bridge. Dark and inaccessible parts of bridges can be used for criminal activities. Some people commit suicide by jumping from bridges. Most of these issues are not traditionally taken to be the responsibility of bridge constructors. (van Gorp 2005, pp.104-110) Should the concept of safe design be so wide that it covers these and other potential negative events in addition to the

traditional issues of structural reliability?

It can be argued in favour of a wide definition of the designer's responsibility that what she does has a lasting influence on safety. The designer can often solve safety problems that are virtually impossible for future users to solve. However, against this it can be argued that the designer is not in a position to solve all problems that may arise from future uses. It is impossible to predict all future uses and misuses of a product. How can the designer be responsible for future events that she has no means to foresee?

The other aspect of safety that needs to be specified is *what it means to be safe against something*. This is the subject of the present contribution. I will approach it by studying some major practices in engineering design.

2. Practices in Safe Design

There are many treatments of safe design in particular fields of engineering, but I am not aware of any fully general account of principles for safe design. However, the following four design principles are in general use in many fields of engineering. They can therefore be taken as representative of the engineering practices of safe design:

1. *Inherently safe design*. A recommended first step in safety engineering is to minimize the inherent dangers in the process as far as possible. This means that potential hazards are excluded rather than just enclosed or otherwise coped with. Hence, dangerous substances or reactions are replaced by less dangerous ones, and this is preferred to using the dangerous substances in an encapsulated process. Fireproof materials are used instead of inflammable ones, and this is considered superior to using flammable materials but keeping temperatures low. For similar reasons, performing a reaction at low temperature and pressure is considered superior to performing it at high temperature and pressure in a vessel constructed for these conditions.

2. *Safety factors*. Constructions should be strong enough to resist loads and disturbances exceeding those that are intended. A common way to obtain such safety reserves is to employ explicitly chosen, numerical safety factors. Hence, if a safety factor of 2 is employed when building a bridge, then the bridge is calculated to resist twice the maximal load to which it will in practice be exposed.

3. *Negative feedback*. Negative feedback mechanisms are introduced to achieve a self-shutdown in case of device failure or when the operator loses control. Two classical examples are the safety-valve that lets out steam when the pressure becomes too high in a steam-boiler and the dead man's handle that stops the train when the driver falls asleep. One of the most important safety measures in the nuclear industry is to ensure that reactors close down automatically in critical situations.

4. Multiple independent safety barriers. Safety barriers are arranged in chains. The aim is to make each barrier independent of its predecessors so that if the first fails, then the second is still intact, etc. Typically the first barriers are measures to prevent an accident, after which follow barriers that limit the consequences of an accident, and finally rescue services as the last resort. One of the major lessons from the Titanic disaster is that an improvement of the early barriers (in this case: a hull divided into watertight compartments) is no excuse for reducing the later barriers (in this case: lifeboats).

Safety engineering includes many more principles and practices than the four mentioned above. Education of operators, maintenance of equipment and installations, and incidence reporting are examples of safety practices of general importance. However, I believe that the four mentioned above cover at least a large part of the practices that are central in engineering design.

3. SAFETY, RISK, AND UNCERTAINTY

Is there a common notion of safety underlying the four general safety practices outlined in the previous section? One obvious answer could be that safety is understood in this context as the antonym of risk, so that a design is safe to the extent that it reduces risk. In probabilistic risk analysis (PRA; also called probabilistic safety analysis, PSA), risk is defined in exact numerical terms. Therefore, safe design could tentatively be defined as design that reduces or minimizes risk in the standard sense of this term, as it is used in PRA. In what follows I will show that this is not a workable definition of safe design. To see this, we need to introduce the decision-theoretical distinction between risk and uncertainty.

In decision theory, "risk" and "uncertainty" are the two major categories of lack of knowledge. In decision-making under risk, the probabilities of possible outcomes are known, whereas in decision-making under uncertainty, probabilities are either unknown or only known with insufficient precision. Hence, decisions at the roulette table are decisions under risk, whereas a choice between two dinner parties is a decision under uncertainty. Uncertainty also covers the cases in which the possible outcomes, not only their probabilities, are unknown. (Hansson 1996)

Few if any decisions in actual life are based on probabilities that are known with certainty. Strictly speaking, the only clear-cut cases of "risk" (known probabilities) seem to be idealized textbook cases that refer to devices such as dice, coins, or roulette wheels that are supposedly known with certainty to be fair. More typical real-life cases are characterized by uncertainty that does not, primarily, come with exact probabilities. Hence, almost all decisions are decisions "under uncertainty". To the extent that we make decisions "under risk", this does not mean that these decisions are made under conditions of completely known probabilities. Rather, it means that we have chosen to simplify our description of these decision problems by treating them as cases of known probabilities.

This ubiquity of uncertainty applies also in engineering design. An engineer performing a complex design task has to take into account a large number of hazards and eventualities. Some of these eventualities can be treated in terms of probabilities; the failure rates of some components may for instance be reasonably well-known from previous experiences. However, even when we have a good experience-based estimate of a failure rate, some uncertainty remains about the correctness of this estimate and in particular about its applicability in the context to which we apply it. In addition, in every design process there are uncertainties for which we do not have good or even meaningful probability estimates. This includes the ways in which humans will interact with new constructions. As one example of this, users sometimes "compensate" for improved technical safety by more risk-taking behaviour. Drivers are known to have driven faster or delayed braking when driving cars with better brakes. (Rothengatter 2002) It is not in practice possible to assign meaningful numerical probabilities to these and other human reactions to new and untested designs. It is also difficult to determine adequate probabilities for unexpected failures in new materials and constructions or in complex new software. We can never escape the uncertainty that refers to the eventuality of new types of failures that we have not been able to foresee.

Of course, whereas reducing risk is obviously desirable, the same may not be said about the reduction of uncertainty. Strictly interpreted, uncertainty reduction is an epistemic goal rather than a practical one. However, by reducing uncertainty we place ourselves in a situation in which we can make more well-informed practical decisions, e.g. about risk reduction. In the choice between decision alternatives that differ in their degrees of uncertainty about possible dangers, by choosing an alternative with low uncertainty we ensure that risks are within stricter bounds than if we choose an alternative with greater uncertainty in this respect.

In summary, engineering design always has to take into account *both* uncertainties that can be meaningfully expressed in probabilistic terms *and* eventualities for which this is not possible. The former are no less ethically relevant than the latter. In the next two sections, I will discuss the implications of uncertainty for two of the safe design strategies mentioned above, namely safety factors and multiple safety barriers.

4. Safety Factors

Probably, humans have made use of safety reserves since the origin of our species. They have added extra strength to their houses, tools, and other constructions in order to be on the safe side. However, the use of numerical factors for dimensioning safety reserves seems to be of relatively recent origin, probably the latter half of the 19th century. The earliest usage of the term recorded in the Oxford English Dictionary is from WJM Rankine's book *A manual of applied mechanics* from 1858. In the 1860s, the German railroad engineer A. Wohler recommended a factor of 2 for tension. (Randall 1976) The use of safety factors is now since long well established in structural mechanics and in its

many applications in different engineering disciplines. Elaborate systems of safety factors have been developed, and specified in norms and standards.

A safety factor is typically intended to protect against a particular integrity-threatening mechanism, and different safety factors can be used against different such mechanisms. Hence one safety factor may be required for resistance to plastic deformation and another for fatigue resistance. As already indicated, a safety factor is most commonly expressed as the ratio between a measure of the maximal load not leading to the specified type of failure and a corresponding measure of the applied load. In some cases it may instead be expressed as the ratio between the estimated design life and the actual service life.

In some applications safety margins are used instead of safety factors. A safety margin differs from a safety factor in being additive rather than multiplicative. In order to keep airplanes sufficiently apart in the air a safety margin in the form of a minimal distance is used. Safety margins are also used in structural engineering, for instance in geotechnical calculations of embankment reliability. (Duncan 2000)

According to standard accounts of structural mechanics, safety factors are intended to compensate for five major categories of sources of failure:

1) higher loads than those foreseen,
2) worse properties of the material than foreseen,
3) imperfect theory of the failure mechanism in question,
4) possibly unknown failure mechanisms, and
5) human error (e.g. in design).

(Knoll 1976. Moses 1997.)

The first two of these refer to the variability of loads and material properties. Such variabilities can often be expressed in terms of probability distributions. However, when it comes to the extreme ends of the distributions, lack of statistical information can make precise probabilistic analysis impossible. Let us consider the variability of the properties of materials. Experimental data on material properties are often insufficient for making a distinction between e.g. gamma and lognormal distributions, a problem called *distribution arbitrariness*. (Ditlevsen 1994) This has little effect on the central part of these distributions, but in the distribution tails the differences can become very large. This is a major reason why safety factors are often used as design guidance instead of probabilities, although the purpose is to protect against failure types that one would, theoretically, prefer to analyze in probabilistic terms.

> Theoretically, design by using structural system reliability is much more reasonable than that based on the safety factor. However, because of the lack of statistical data from the strength of materials used and the applied loads, design concepts based on the safety factor will still dominate for a period. (Zhu 1993)

The last three of the five items on the list of what safety factors should protect against all

refer essentially to errors in our theory and in our application of it. They are therefore clear examples of uncertainties that are not easily amenable to probabilistic treatment. In other words: The eventuality of errors in our calculations or their underpinnings is an important reason to apply safety factors. This is an uncertainty that is not reducible to probabilities that we can determine and introduce into our calculations. It is for instance difficult to see how a calculation could be accurately adjusted to compensate self-referentially for the possibility that it may itself be wrong. However, these difficulties do not make these sources of failures less important from an ethical point of view. Safety factors are used to deal both with those failures that can be accounted for in probabilistic terms and those that cannot.

5. Safety Barriers

Some of the best examples of the use of multiple safety barriers can be found in nuclear waste management. The proposed subterranean nuclear waste repositories all contain multiple barriers. We can take the current Swedish nuclear waste project as an example. The waste will be put in a copper canister that is constructed to resist the foreseeable challenges. The canister is surrounded by a layer of bentonite clay that protects the canister against small movements in the rock and "acts as a filter in the unlikely event that any radionuclides should escape from a canister". This whole construction is placed in deep rock, in a geological formation that has been selected to minimize transportation to the surface of any possible leakage of radionuclides. The whole system of barriers is constructed to have a high degree of redundancy, so that if one the barriers fails the remaining ones will suffice. With usual PRA standards, the whole series of barriers would not be necessary. Nevertheless, sensible reasons can be given for this approach, namely reasons that refer to uncertainty. Perhaps the copper canister will fail for some unknown reason not included in the calculations. Then, hopefully, the radionuclides will stay in the bentonite, etc. In this particular case, redundancy can also be seen as a means to meet public scepticism and opposition (although it is not self-evident that redundant safety barriers will make the public feel safer).

For another example, we can again consider what is possibly the most well-known example of technological failure in modern history, the Titanic that sank with 1500 persons in April 1912. It was built with a double-bottomed hull that was divided into sixteen compartments, constructed to be watertight. Four of these could be filled with water without danger. Therefore, the ship was believed to be unsinkable, and consequently it was equipped with lifeboats only for about half of the persons onboard.

We now know that the Titanic was far from unsinkable. But let us consider a hypothetical scenario. Suppose that tomorrow a ship-builder comes up with a convincing plan for an unsinkable boat. A probabilistic risk analysis shows that the probability of the ship sinking is incredibly low. Based on the PRA, a risk-benefit analysis has been performed. It shows that the cost of life-boats would be economically indefensible. The expected cost per life saved by the life-boats is above 1000 million dollars, a sum that can evidently be more efficiently used to save lives elsewhere. The risk-benefit analysis therefore clearly

shows us that the ship should not have any lifeboats.

How should the naval engineer respond to this proposal? Should she accept the verdict of the economic analysis and exclude lifeboats from the design? My proposal is that a good engineer should not act on the risk-benefit analyst's advice in a case like this. The reason for this is obvious from what has already been said: The calculations may possibly be wrong, and if they are, then the outcome may be disastrous. Therefore, the additional safety barrier in the form of lifeboats (and evacuation routines and all the rest) should not be excluded, in spite of the probability estimates showing them to be uncalled for.

6. Conclusion

Many of the most ethically important safety issues in engineering design refer to hazards that cannot be assigned meaningful probability estimates. It is appropriate that at least two of the most important strategies for safety in engineering design, namely safety factors and multiple safety barriers, deal not only with risk (in the standard, probabilistic sense of the term) but also with uncertainty.

Currently there is a trend in several fields of engineering design towards increased use of probabilistic risk analysis (PRA). This trend may be a mixed blessing since it can lead to a one-sided focus on those dangers that can be assigned meaningful probability estimates. PRA is an important design tool, but it is not the final arbitrator of safe design since it does not deal adequately with issues of uncertainty. Design practices such as safety factors and multiple barriers are indispensable in the design process, and so is ethical reflection and argumentation on issues of safety. Probability calculations can often support, but never supplant, the engineer's ethically responsible judgment.

References

Clausen, Jonas, Sven Ove Hansson and Fred Nilsson, "Generalizing the Safety Factor Approach", *Journal of Reliability and Engineering System Safety*, in press.

Ditlevsen, O. 1994. "Distribution arbitrariness in structural reliability" in Schuëller, G. Shinozuka, M. and Yao, J. (1994) *Proc. of ICOSSAR'93: Structural Safety & Reliability* 1241-1247.

Duncan, J.M. 2000. "Factors of safety and reliability in geotechnical engineering". *Journal of Geotechnical and Geoenvironmental Engineering* 126:307-316.

Hansson, Sven Ove. 1996 "Decision-Making Under Great Uncertainty", *Philosophy of the Social Sciences* 26:369-386.

Kemper, Bart. 2004. "Evil Intent and Design Responsibility" *Science and Engineering Ethics* 10(2): 303-309.

Knoll, F. 1976. "Commentary on the basic philosophy and recent development for safety margins", *Canadian Journal of Civil Engineering.* 3:409-416.

Lloyd.Peter and Jerry Busby.2003. "Things That Went Well—No Serious Injuries or Deaths": Ethical Reasoning in a Normal Engineering design Process" *Science and Engineering Ethics* 9:503-516.

Martin, Mike W. and Roland Schinzinger. 2005. *Ethics in engineering*, 4th ed., Boston: McGraw-Hill, 2005.

Moses, F. 1997. "Problems and prospects of reliability-based optimisation", *Engineering Structures* 19:293-301.

Palm, Elin and Sven Hansson, "The Case for Ethical Technology Assessment (eTA), *Technological Forecasting and Social Change*, in press.

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

Rothengatter, Talib. 2002 "Drivers' illusions – no more risk", *Transportation Research*, part F, 5:249-258.

van de Poel, Ibo. 2001 "Investigating Ethical Issues in Engineering Design" *Science and Engineering Ethics* 7: 429-446.

van Gorp, Anke. 2005. *Ethical issues in engineering desing; Safety and sustainability*, PhD thesis, Delft University 2005.

Zhu, T.L. 1993. "A reliability-based safety factor for aircraft composite structures", *Computers & Structures* 48:745-748.