1

INTRODUCTION

Niklas Möller, Sven Ove Hansson, Jan-Erik Holmberg, and Carl Rollenhagen

Principles for action have a much more important role in safety engineering and safety management than in most other disciplines. In practical safety work, we refer to principles such as fail-safe, safety barriers, safety factors, system redundancy, resilience, inherent safety, and many others. Much of the academic literature on safety, safety standards, and regulations recommends the use of one or other such principle. Many of the best-known contributors to the safety literature owe their fame to their roles as originators or promoters of one or other safety principle.

1.1 COMPETITION, OVERLAP, AND CONFLICTS

But the field is not characterized by consensus on which safety principles we should use. To the contrary, the literature on these principles abounds with divergent and sometimes conflicting recommendations. The overall picture is a rather confused one, due to competition, overlap, and conflicts among the principles.

Handbook of Safety Principles, First Edition. Edited by Niklas Möller, Sven Ove Hansson, Jan-Erik Holmberg, and Carl Rollenhagen.

It is not uncommon to hear presentations in which one of the safety principles is expanded to include all aspects of risk and safety enhancement so that it becomes *the* overarching principle under which the others can be subsumed. Quite a few of the principles have been presented with such ambitions—general quality principles, integrated risk management, and safety culture to name just a few—but obviously at most one of the principles can be superordinate to all the others. There seems to be a certain element of modishness in the coming and going of safety principles, and the field may not be entirely devoid of factionalism.

More often than not, one and the same safety measure can be presented as based on various principles. Terminologies also differ between industry branches and engineering specialties. For instance, what is called "inherent safety" in the chemical industry is called "substitution principle" in many industries that use chemical products, "passive safety" in the nuclear industry, and "primary prevention" in health-related applications. These principles seem to be close in meaning, but how large is the overlap? Can they perhaps even be described as one and the same principle but under different names?

Conflicts between the principles are far from uncommon. The principle of cost–benefit optimization tells us not to reduce low radiation doses if the reduction is costly, but at least some interpretations of the ALARA ("as low as reasonably achievable") principle tell us to reduce them. The principle of multiple safety barriers sometimes recommends an extra layer of safety that the principle of simplicity would dissuade us from since it makes the system more complex and difficult to manage in a safety-critical situation. Sometimes, even two applications of the same safety principle can lead to a conflict. For instance, the substitution principle recommends that we replace flammable substances by less flammable ones and toxic substances by less toxic ones. In the choice between two substances, one of which is less flammable and the other less toxic, this will lead to a conflict.

1.2 A NEW LEVEL IN THE STUDY OF SAFETY PRINCIPLES

As we see it, the study of safety principles has to be taken to a new and more comprehensive level. It is not sufficient to study the principles one at a time, and promotion of single principles needs to be replaced by unbiased comparative investigations. There is no lack of topics for such studies. We need to find out the relationships between the different principles, not least how they overlap and how they may run into conflict with each other. We also need to learn how they are conceived and applied by those who are supposed to be helped by them in their daily work (not only how they are conceived by their most fervent champions). And most importantly, we need to know if they make a difference in practice. What effects, if any, does their application have on safety outcomes? In short, the academic literature on safety principles should become much less advocacy-based and much more evidence-based.

We see this book as a first step toward that new level in the study of safety principles. Most of the major safety principles are presented, and they are all dealt with

on an equal basis. We have asked the authors to compare the principle(s) they present to other safety principles. We have also asked them to clarify the limitations and weaknesses of the various principles, and to inform the reader of whatever empirical evidence there may be of the effects of using the principles in practice. The book contains a significant amount of comparative material, and we hope that it will also serve as an inspiration for more comparative studies of safety principles in the near future.

1.3 METAPRINCIPLES OF SAFETY

Does it make any difference which safety principle(s) we appeal to, and which of them we choose as an overarching principle for safety management? We believe that it can indeed make a difference. The reason for this is that the different safety principles put emphasis on different components of safety management. There are many possible "metaprinciples" which may be used for bringing out differences in emphasis between safety principles. We have found the following simple list of basic tasks in safety management useful, and will in this section illustrate how it brings forward an interesting pattern for the principles of safety covered in this handbook:

1. *Inventorize*. Identify and assess specific safety problems in planned or existing systems.
2. *Capacitate*. Investigate what capacities the system has to deal with safety-related problems and how those capacities can be improved. Many of these principles are applied in the design phase but can also be implemented as a consequence of applying problem-finding principles in existing systems.
3. *Prioritize*. Set priorities among the potential improvements.
4. *Integrate*. Make safety management coherent and comprehensive, for example, by using general quality principles and integrated safety management principles.

Each of these tasks is an important component in safety management. Therefore, the safety principle(s) applied in safety management should sustain the performance of each of them. We will call them metaprinciples since they will be used to evaluate many of the common safety principles.

In Figure 1.1, we have placed three of the safety metaprinciples at the vertices of a triangle, and we have introduced some well-known safety principles at different places in the triangle. The diagram illustrates how these three principles give rise to different approaches to practical safety work. Notably, some safety principles are close to one of the vertices. Such a safety principle will in practice only support one of the metaprinciples, and it is therefore in obvious need of supplementation. One example is the principle of experience feedback. This is a principle with a strong focus on inventorizing. By studying previous accidents, incidents, and other events, we can learn much on how to avoid similar events in the future. But obviously,

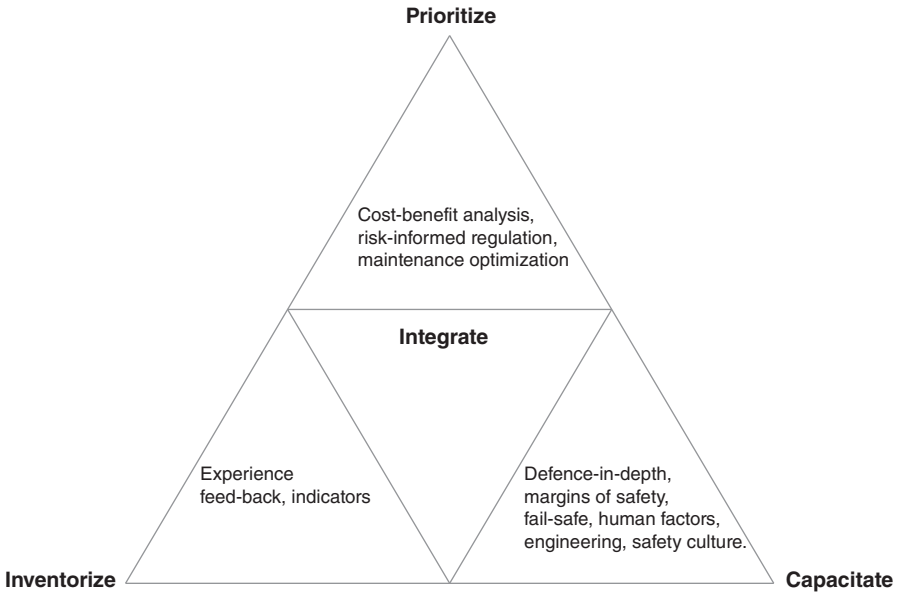

Figure 1.1. Four metaprinciples of safety. Specific safety principles can be characterized according to whether they have a dominant focus on finding real or potential safety problems (inventorize), providing capacity and resources to cope with real or potential problems (capacitate), or to make priorities (prioritize). At the center of the figure are principles that describe how to integrate components of the other principles.

this does not necessarily teach us about the capacity of the system to deal with the events (particularly unforeseen ones), and neither does it tell us much about how to prioritize among different problems in need of solution.[1] Therefore, the principle of experience feedback is not sufficient to guide safety management as a whole. We will have to add other principles that provide guidance for capacitation and prioritization.

Similarly, the principle of cost–benefit optimization is very close to the vertex representing prioritization. It is a rather complete (but of course not uncontestable) principle for priority-setting of safety measures, but it does not help us much in identifying safety problems or pinpointing general capacity improvements that can increase our preparedness for unidentified hazards. Therefore, cost–benefit optimization is not suitable as an overarching principle to cover all safety work.

In the middle triangle of Figure 1.1, we have positioned safety principles that are very general and usually contain parts of the other three metaprinciples. These principles often give general advice about how to structure safety management and the other metaprinciples. For example, general quality principles give advice that

[1] It should, however, be recognized that the principle of experience feedback is also used to identify good practices and solutions to previously identified safety problems.

can be applied to most other specific safety principles. Integrated safety management principles are also of this type.

Not surprisingly, principles that are close to each other in the diagram are more likely to overlap than principles at larger distance from each other. For instance, inherent safety and the substitution principle are close to each other in the diagram. This is because they both provide means to avoid both known and unknown dangers. These two principles tend to yield the same recommendations. For instance, both tell us to replace a flammable substance by a non-flammable one, which is an advantage both in known and unknown accident scenarios. The two principles are quite close to each other, and one might sensibly ask if they coincide or if one of them is a variant of the other.

In other cases, principles that are close to each other in the diagram run into conflict rather than overlapping. This applies for instance to cost–benefit optimization and best available technology. These two principles tell us how to prioritize, but they give us different advice on that topic. When a new, safer technology becomes available, best available technology will advise us to implement it, but cost–benefit optimization will often give contradictory advice. Other potential relations and potential conflicts between safety principles will be discussed further below.

1.4 OTHER WAYS TO CHARACTERIZE SAFETY PRINCIPLES

While our list of metaprinciples provides a way to see how safety principles overlap and what parts of safety management they cover, it is far from the only way to describe and categorize safety principles. As we see it, there is a need for a considerable methodological pluralism in the characterization of safety principles. For the purpose of dividing this book into main sections, we have adopted a more traditional approach, collecting the chapters *thematically* into five parts: Safety Reserves, Information and Control, Demonstrability, Optimization, and Organizational Principles and Practices. Some of these categories seem to correlate rather closely to our suggested metaprinciples. All of the principles treated in the first part, Safety Reserves, have a rather strong capacitating focus, for example, whereas most of the principles covered in the Optimization part focus on prioritizing. Other parts correlate more diversely and can be placed in the middle. The chapters in the Organizational Principles and Practices part, for example, have some emphasis on capacitation, but include many safety principles with a wholesale approach which cover all the three components positioned at the corners.

There are several alternative ways of categorizing safety principles. One common division is procedural. Different phases in the lifecycle of a system require different principles; typically, the three basic system phases selected are *design*, *operation*, and *decommissioning* (IAEA, 1986).

Another categorization of safety principles attempts at dividing hazard reduction into categories according to their priority. The basic idea here is that a hazard should if

possible be *eliminated*. If that is not possible, it should be *isolated*, *controlled*, or *limited*. Haddon (1980a, 1980b; cf. also Runyan, 2003; Saleh et al., 2014) arguably provides the classic account of this strategy. On his model, safety is analyzed through the three basic concepts of *threat*, *barrier*, and *object of value,* and his strategy, in which he utilizes the so-called energy model of accidents (Saleh et al., 2010), includes:

1. Reducing the energy in the system
2. Controlling the energy in the system
3. Separating the energy source temporally and spatially from the object of value
4. Enhancing the damage resistance of the objects of value (Saleh et al., 2014)

Bahr (1997) provides a more recent example of a similar strategy (in turn based on NASA, 1993), suggesting the following taxonomy:

1. "Designing out" the hazard
2. Safety devices
3. Warning devices
4. Special procedures and training

First, Bahr writes, we should "design out" the hazard from the system. If that is not possible, we should control the hazard using various fail-safe devices; for example, pressure valves relieving the system of dangerous pressure build-up. When designing out or controlling is not an option, warning devices (e.g., smoke alarm) and procedures (e.g., emergency shutdown) and training should be used (Bahr, 1997).

Another suggested list of covering principles, similar in that it also focuses on substantially different strategies of risk reduction, is given by Möller and Hansson (2008). They divide a large number of engineering safety principles into four covering principles:

1. Inherently safe design
2. Safety reserves
3. Safe fail
4. Procedural safeguards

Inherently safe design is the design strategy to minimize the inherent dangers in the process as far as possible. The general idea here is that potential hazards are excluded rather than just enclosed or otherwise coped with (cf. Chapter 17). Safety reserves is the strategy of making constructions strong enough to resist loads and disturbances exceeding those that are intended; for example, by employing explicitly chosen, numerical safety factors. (Chapters 3–6 treat different aspects of this strategy.) The covering principle of safe fail entails that the system should fail "safely"; internal components may fail without the system as a whole failing, or the system fails

without causing harm. "Fail-safe," "fail-silence," and "negative feedback" denote different variants of this principle. (Hammer, 1980). Procedural safeguards refer to control mechanisms for enhancing safety, ranging from general safety standards and quality assurance to training and behavior control of the staff. (Several chapters in this handbook, and Part V in particular, treat procedural safeguards.)

Another categorization focuses on the temporal dimension involved in decision-making. Here, a fundamental division is between principles focusing on *passed experience* (such as in Chapter 7), *current states of the system*, or *projections to the future*.

Safety principles may also be categorized in relation to the object of regulation. Here, a distinction can be drawn between principles directed at the *technical system*, the *human agent*, and the *organization*.

Many other ways to characterize safety principles are available in the literature. For further categorizations, see, for example, Saleh et al. (2014), Jackson and Ferris (2013), Jackson (2010), Khan and Amyotte (2003), and Kletz (1978, 1998).

1.5 CONFLICTS BETWEEN SAFETY PRINCIPLES

Safety principles are not conflict-free. However, we should distinguish between on the one hand conflicts between principles as such and on the other hand conflicts between applications of principles. To exemplify the former type of conflicts, we can suppose that a company has adopted the following two principles: All employees have the right to report any safety concern to the chief safety manager, and all communications on safety must be checked by the responsible foreman before they are disseminated outside of the department. These two principles are obviously in conflict. The example is contrived, and the reason why it is contrived is that this type of conflict appears to be unusual. Safety principles that are used in practice tend not to be in conflict in this way.[2]

The other type of conflict, between applications of principles, is much more common. Two principles may seem to be perfectly compatible, but there may still be practical cases when they cannot both be satisfied. For instance, in workplaces with risks of poisonous gas leakage, we may wish to implement the principle that it should be possible to evacuate the building in a very short time. In workplaces where a terrorist attack is comparatively likely and can have disastrous effects, we may wish to implement the principle that unauthorized access should be virtually impossible. It is not difficult (although sometimes expensive) to implement one of these two principles. However, implementing them both is often very difficult. Therefore, when they both

[2] Note though that safety and security principles often conflict, since they may have different goals. Security procedures may, for example, attempt to keep an intruder from getting away in case of an incident (locking down a site after a breach), while safety procedures should do the exact opposite, that is, help people to abandon a site in case of an incident.

need to be applied, we have a conflict, but strictly speaking, it is not a conflict between the principles but between their applications in a particular situation.

In practice, it is the latter type of conflicts (between applications of principles) that we have to deal with, rather than conflicts of the former type (between principles as such). We can further distinguish between three types of conflicts here.

First, there are conflicts between two applications of the same safety principle. We have already given an example of this: The principles of inherent safety requires both that we replace toxic substances by less toxic ones and that we replace flammable substances by less flammable ones. These two specifications run into conflict if the least toxic alternative is not also the least flammable one.

The second type concerns applications of different safety principles. For instance, we may have one safety principle requiring that all safety-critical procedures should follow pre-determined protocols, and another safety principle requiring that all employees should be encouraged to take initiatives and continuously improve the safety of work processes in which they take part. Although these two principles can be combined it is difficult to do so, and in most practical cases they will give rise to conflicts.

The third type of conflict concerns applications of a safety principle and some other principle that is supposed to be upheld on the workplace. Safety measures are sometimes costly (and the savings they induce tend to be difficult to demonstrate beforehand). Therefore, conflicts between safety principles and the ubiquitous principle of cost minimization are common.

Other conflicts that are rather typical in many sociotechnical systems are represented by the following examples:

- The principle of automation of controlling important safety functions may be in conflict with the safety principle that operators should have full control and situation awareness about what happens in a technical process.
- The principle of diversified safety systems may be in conflict with the principle of simplicity in designs.
- Principles of clear lines of accountability in a line organization may be in conflict with the principle that decisions shall be taken in those groups that hold the strongest expertise in a certain domain.
- Principles of documentation and details in descriptions of a safety management system may be in conflict with the principle of striving for oversight and simplicity.
- The principle of applying detailed step-by-step instructions may be in conflict with principles that strive for learning and competence in safety related work.

1.6 WHEN CAN SAFETY PRINCIPLES BE BROKEN?

Since conflicts involving safety principles are so common, it seems unavoidable that sometimes safety principles can legitimately be broken. But on the other hand, many

if not most, severe accidents resulted from violations of safety principles. Therefore, we need guidelines (or metaprinciples) for legitimate violations of safety principles. We propose four such guidelines:

1. Barring exceptional circumstances, violations of safety principles can only be legitimate if the purpose is to *improve safety*. In a conflict between two safety principles, one of them will have to yield, at least in the short run. In a conflict between a safety principle and some other principle (such as cost minimization), the safety principle should be upheld.
2. Conflicts between safety principles should as far as possible be solved in a *risk-minimizing way*. For instance, in the above-mentioned conflict between the aims to reduce toxicity and risks of fire, an analysis showing that one of these two risks is larger than the other provides a weighty argument on how to proceed.
3. Conflicts between safety principles are often a sign that *more thorough changes are needed* that will make it possible to satisfy all of the conflicting principles. For instance, having to choose between a highly toxic and a highly flammable substance is an unsatisfactory situation. The short-term choice between conflicting safety principles should therefore be followed by development work aiming at finding ways to comply fully with both of them.
4. All violations of safety principles should be done *openly*, and discussed with everyone who is concerned. Such open discussions diminish the risk that safety principles are given up for no good reasons.

1.7 SAFETY IN CONTEXT

In this introductory chapter, our aim has been to put the notion of safety principles in context, suggesting a simple set of metaprinciples as well as pointing to the multitude of ways in which safety principles may be characterized. By covering a large number of safety principles this book actualizes how principles may conflict, and we have sketched a number of different types of conflicts between principles, as well as addressed the question when safety principles can rightly be broken. Our treatment has by necessity been on the abstract side of things, and before turning to the main content of this book, we would therefore like to remind the reader that while principles may provide action guidance and structure, the actual decision situation in which we find ourselves comes with a unique context (cf. Jackson and Ferris, 2013). In the individual case, we always have to look carefully at the context and take a stand on the salient factors. There will always be a judgment involved on which principles to apply, how, and when. This judgment can be fine-tuned by carefully studying our most informed accounts of the principles of safety. To this we now turn.

REFERENCES

Bahr, N. J. (1997). *System Safety Engineering and Risk Assessment: A Practical Approach*. Washington, DC: Taylor & Francis.

Haddon, W. (1980a). Advances in the epidemiology of injuries as a basis for public policy. *Public Health Reports*, 95(5), 411–421.

Haddon, W. (1980b). The basic strategies for preventing damage from hazards of all kinds. *Hazard Prevention*, 16, 8–11.

Hammer, W. (1980). *Product Safety Management and Engineering*. NJ: Prentice-Hall.

IAEA (1986). *General design safety principles for nuclear power plants: A safety guide*. International Atomic Energy Agency, Vienna.

Jackson, S. (2010). *Architecting Resilient Systems: Accident Avoidance and Survival and Recovery from Disruptions*. Wiley Series in Systems Engineering and Management, A.P. Sage (Series Ed.). Hoboken, NJ: John Wiley & Sons.

Jackson, S., and Ferris, T. (2013). Resilience principles for engineered systems. *Systems Engineering*, 16(2), 152–164.

Khan, F. I., and Amyotte, P. R. (2003). How to make inherent safety practice a reality. *The Canadian Journal of Chemical Engineering*, 8(1), 2–16.

Kletz, T. (1978). What you don't have, can't leak. *Chemistry and Industry*, 6, 287–292.

Kletz, T. (1998). *Process Plants: A Handbook for Inherently Safer Design*. Taylor & Francis.

Möller, N., and Hansson, S. O. (2008). Principles of engineering safety: Risk and uncertainty reduction. *Reliability Engineering & System Safety*, 93(6), 776–783.

NASA (1993). Safety policy and requirements document. NHB 1700.1 (V1-B). NASA, Washington, DC.

Runyan, C. W. (2003). Back to the future – revisiting Haddon's conceptualization of injury epidemiology and prevention. *Epidemiologic Reviews*, 15(1), 60–64.

Saleh, J. H., Marais, K. B., and Favaró, F. M. (2014). System safety principles: A multidisciplinary engineering perspective. *Journal of Loss Prevention in the Process Industries*, 29, 283–294.

Saleh, J. H., Marais, K. B., Bakolas, E., and Cowlagi, R. V. (2010). Highlights from the literature on accident causation and system safety: review of major ideas, recent contributions, and challenges. *Reliability Engineering and System Safety*, 95(11), 1105–1116.