Prologue: The Scope of Resilience Engineering

Erik Hollnagel

The focus for safety efforts is usually, and traditionally, the unwanted outcomes, injuries, and losses that are the result of adverse events. This matches the common understanding of safety as 'freedom from unacceptable risk.' Resilience Engineering, however, defines safety as the ability to succeed under varying conditions. It is a consequence of this definition that it is equally important to study things that go right as things that go wrong. For Resilience Engineering, the understanding of the normal functioning of a socio-technical system is the necessary and sufficient basis for understanding how it fails. And it is both easier and more effective to increase safety by improving the number of things that go right, than by reducing the number of things that go wrong. The definition of resilience can be made more concrete by pointing to four abilities that are necessary for a system to be resilient. These are the ability to respond to events, to monitor ongoing developments, to anticipate future threats and opportunities, and to learn from past failures and successes alike. The engineering of resilience comprises the ways in which these four capabilities can be established and managed.

Introduction

In the world of safety, comprising issues such as accident investigation, risk assessment, safety management, and safety culture, the focus has traditionally been on that which has gone wrong or could go wrong. This is illustrated by the commonly used definition of safety as 'freedom from unacceptable risk.' The

focus on what could go wrong obviously makes practical sense, since clearly it is important for every enterprise to understand both what has gone wrong and what may go wrong, in order to develop measures either to prevent it from happening (again) or to protect against the outcomes.

This line of thinking is well illustrated by the traditional risk matrix, an example of which is shown in Figure P.1. The risk matrix characterises the risk level of possible outcomes by considering both their probability, that is, the likelihood that they will happen, and the severity of the consequences, that is, the magnitude of the possible consequences.

The risk matrix, however, only looks at things that can go wrong. Yet if we consider the possible outcomes of something (an event, a function, or a process), it is clear that things can go right as well as wrong. It is furthermore reasonable to expect that things normally will go right, that they will turn out as planned or intended, and that it is unusual for things to go wrong. We are therefore unpleasantly surprised when it happens. In view of this, it seems reasonable to propose that a description of possible outcomes should go beyond the traditional risk matrix and extend the 'consequence' dimension to include both positive (wanted) and negative (unwanted) outcomes. This can be shown as in Figure P.2.

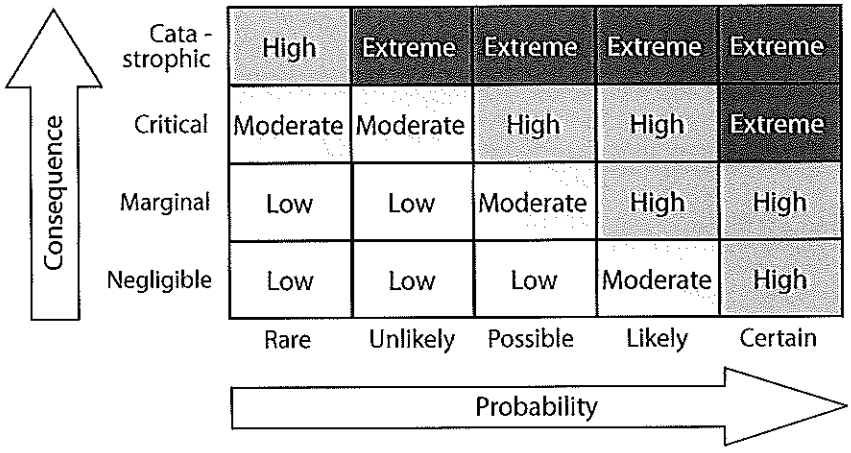

Figure P.1 A traditional risk matrix

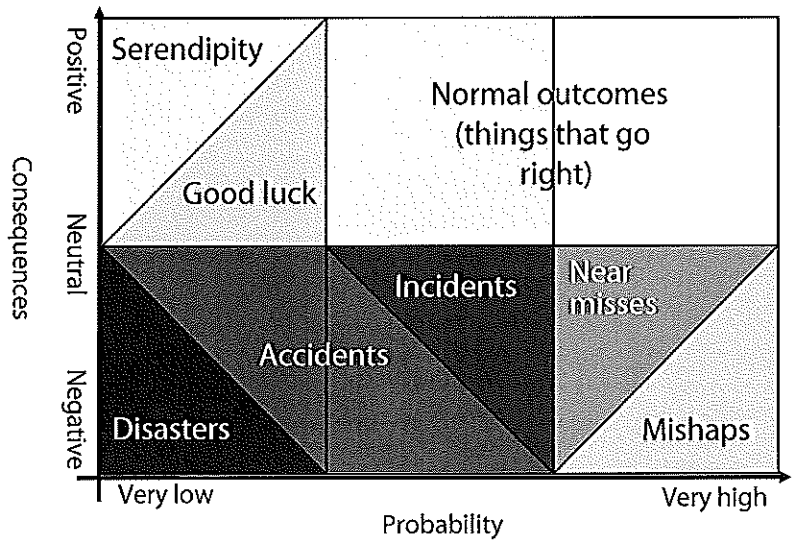

Figure P.2 Range of outcomes

Safety efforts have traditionally focused on unwanted or negative outcomes, and have furthermore been limited to outcomes with a relatively low probability such as incidents and accidents. (Unwanted negative outcomes with high probability, for example, mishaps, will normally have been eliminated, since otherwise the system would be unable to maintain its required functioning.) If we for a moment assume that there is a simple causal relation between events and outcomes, it becomes possible to characterise several characteristic subsets of outcomes as follows:

1. Positive outcomes that have a high probability. This subset represents the successes or 'normal' actions, that is, the things that not only go right but that also are intended and expected to go right—in other words, normal work or normal functioning. Indeed, if normal work either did not result in wanted outcomes, or was not highly predictable, something would be seriously wrong.
2. Positive outcomes that have a low probability. This subset represents the 'good' things that can happen, but that happen unexpectedly. There is no commonly recognised terminology for these, but terms such as *serendipity* or even *good fortune* represent at least some of them.

3. Negative or unwanted outcomes that have a low probability, that is, things that go wrong and which are unexpected—although not unimaginable. This is the subset of outcomes that usually is associated with safety—or rather, with a lack of safety—particularly outcomes that are serious (in terms of causing significant losses) and that are hard to predict. This subset includes the commonly used categories of incidents and accidents. It also includes disasters, although these rarely are covered by industrial safety.
4. Negative or unwanted outcomes that have a high probability. This basically means outcomes that realistically must be expected to happen frequently or even regularly. In practice most of these outcomes have only minor negative consequences, because they otherwise would have been eliminated (cf., the As Low As Reasonably Practicable (ALARP) principle). They are commonly described as near misses or 'almost accidents,' or as unsafe actions. Near misses are usually benign but may lead to serious negative consequences. Another subset is the mishaps, that is, 'near misses' with serious outcomes. Predictable events that may result in serious unwanted outcomes can, however, normally be are assumed to have been eliminated.

A more condensed description of the four sets of outcomes is shown Table P.1.

Table P.1 The sets of possible outcomes

	Things that go right (wanted outcomes)	**Things that go wrong (unwanted outcomes)**
Outcomes with high predictability	This is the set of outcomes that represent the normal functioning of a safe system. Ought to be governed by an As High As Reasonably Practicable (AHARP) principle.	The serious outcomes in this set are normally eliminated; the minor unwanted outcomes are usually tolerated, as described by the ALARP principle.
Outcomes with low predictability	These outcomes are not normally considered in system management, but should obviously be facilitated as far as possible. They are gratefully accepted if and when they occur.	These outcomes are the focus of traditional safety efforts. They are the subject of risk assessment, prevention, and protection. Many efforts are made to calculate how 'unexpected' they are, hence transfer them to the set above.

As already mentioned, safety efforts have usually focused on outcomes that are both unwanted (i.e., with significant negative consequences) and unexpected or difficult to predict, corresponding to the categories of accidents and incidents in Figure P.2 or the high to extreme risks in Figure P.1. The common understanding is that safety can be achieved if accidents, incidents (and mishaps) either can be prevented or if their number (or frequency) can be reduced. Disasters must, of course, not be neglected although their predictability usually is so low that it is difficult to do much to prepare for them. (In relation to the terminology proposed by Westrum (2006), disasters can be seen as irregular threats or even improbable events.) Since the 1980s, the safety focus has occasionally been extended from incidents and mishaps to include near-misses also. But the practical problem is that there are so many near misses, that they happen so frequently, and that the consequences usually are negligible so that it is not considered cost-effective to do much about them.

More importantly, the traditional approaches to safety usually disregard what lies 'above' the middle of Figure P.2, that is, the ways in which things can go right. This is due to the unspoken assumption that we can best learn about things that go wrong by studying only things that go wrong. It is nice when things go right, but there is no need to pay much attention to them precisely because they go right. It is also due to the fact that as we get used to something, we tend not to notice it any longer. (The technical term for this is habituation, which denotes the psychological process in humans that leads to a decrease in response to a stimulus after repeated exposure over a specified duration of time.)

Resilience Engineering, however, takes a different position. Resilience Engineering sees the 'things that go wrong' as the flip side of the 'things that go right,' and therefore assumes that they are a result of the same underlying processes. In consequence of that, 'things that go right' and 'things that go wrong' should be explained in basically the same way. It therefore makes as much sense to try to understand why things go right as to understand why they go wrong. In fact, it makes more sense because there are many more things that go right than things that go wrong, the

ratio depending on how (im)probable an accident is considered to be. If, for instance, the probability of failure is 10E–4 (meaning 10–4 or 1/10.000), then humans are usually blamed for 80–90 percent of the one case out of 10.000 when things go wrong. By the same 'logic,' humans should be praised for a similar 80–90 percent of the 9.999 cases where nothing goes wrong. (In both cases humans should actually be seen as accountable for the full 100 percent, since it would otherwise be necessary to postulate some *deus ex machina* to account for the remaining 10–20 percent.) Resilience Engineering proposes that we should try to understand a system's performance in general, rather than limit ourselves to the things that go wrong, that is, try to understand all the outcomes shown in Figure P.2 rather than only the negative ones—with the possible exception of 'good luck.'

Both Figures P.1 and P.2 use the probability of an outcome as a descriptive dimension, but neither considers the frequency of outcomes. While probability and frequency are closely linked, they do not mean the same, and for the safety of everyday work, the frequency of outcomes is perhaps the more important. Following the argument made above, if things go wrong one time out of every 10.000, then things go right the remaining 9.999 times. This is illustrated in Figure P.3, where a third dimension, representing frequency, has been added to the diagram shown in Figure P.2.

As Figure P.3 tries to illustrate, there are many more things that go right than things that go wrong. Even for ultra-performing systems, the ratio is around 1:1.000 (Amalberti, 2006). For ultra-safe systems it may be 1:1.000.000 or even lower, meaning that the number of normal outcomes is at least six orders of magnitude larger than the number of failures. It is the set of normal outcomes that rightly ought to represent the safe performance of a system or process, just as the set of accidents and incidents represent unsafe performance. It may therefore be said that safety efforts, almost paradoxically, have focused on unsafe functioning rather than on safe functioning. This may, as noted above, be due to the psychological fact that safety is nearly invisible while a lack of safety is highly visible. We notice that which is unusual while we become habituated to that which is usual. Resilience Engineering recognises this paradox and argues that safety should deal with

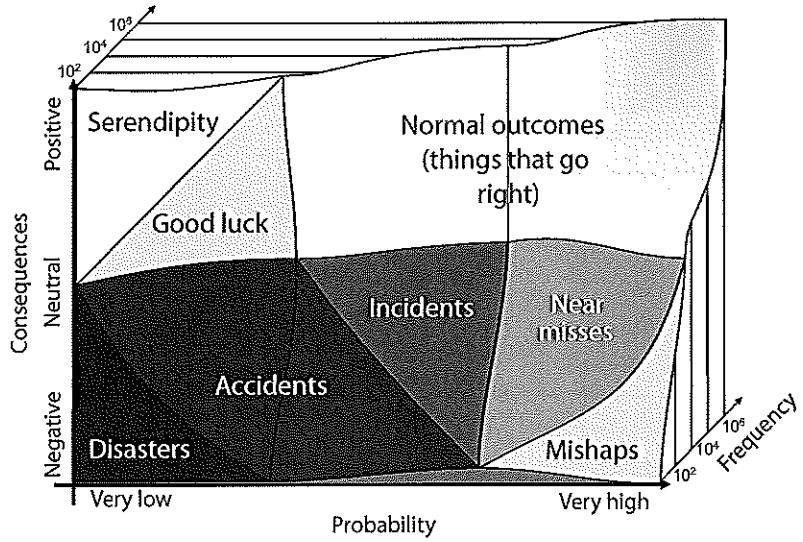

Figure P.3 The frequency of various outcomes

safe performance as well as unsafe performance—with things that go right as well as things that go wrong.

According to this line of reasoning, Resilience Engineering is not a simple replacement for safety (management). Safety (management) has traditionally, and with good reason, focused on a subset of the possible events and outcomes. This was in many ways sufficient as long as systems and processes were manageable, or tractable, so that normal functioning could be ensured by limiting or constraining performance variability (cf., Hollnagel, 2009). The developments in socio-technical systems during the last 20 years or so have, however, created an increasing number of systems and processes that are intractable, and where performance variability consequently is a necessity and an asset rather than a liability. Resilience Engineering argues that it is necessary to look at success as well as at failures precisely in order to understand failures or why things wrong. The argument is that there are no special 'error producing' processes that magically begin to work when an accident is going to happen, but which otherwise lie dormant. On the contrary, there are no fundamental differences between performance that leads to failures and performance that leads to successes. We are therefore best served by trying to understand performance in general, regardless of whether we focus on individual, collective, or organisational performance.

The difference between 'classical' safety management and Resilience Engineering is demonstrated by the differences between the definitions. A common definition of safety was mentioned above. Resilience can in the same manner be defined as:

> The intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions.

This definition obviously comprises the classical definition of safety, since 'the ability to sustain required operations' is tantamount to the 'freedom from unacceptable risks.' But the definition of resilience also makes clear that safety cannot be seen independently of the core process (or business) of the system, hence the emphasis on the ability to *function* under 'both expected and unexpected conditions' rather than just to avoid failures. It is this ability that makes the system both safe and efficient, and Resilience Engineering deals with both.

The difference between the two views is illustrated by Figure P.4, which uses a balance to show two different ways to improve safety. One is to reduce the number of things that go wrong, which obviously will tip the scale in favor of safety. The other is to increase the number of things that go right, which will achieve the same effect, but which at the same time will contribute to productivity and the core business processes. Resilience Engineering favors the second approach. The goal of Resilience Engineering is to increase the number of things that go right rather than to reduce the number of things that go wrong, noting that the latter will be a consequence of the former.

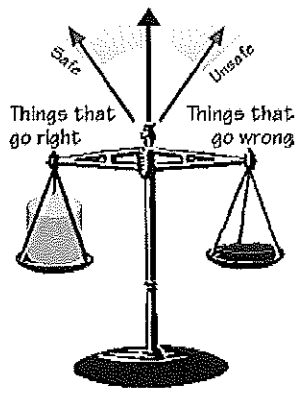

Figure P.4 A resilience engineering view of safety

The Four Cornerstones of Resilience

If we define resilience as proposed above, the goal of Resilience Engineering becomes how to bring about resilience in a system. The key term of this definition is the system's ability to *adjust* its functioning. This working definition of resilience can be made more detailed by noticing that it implies four main factors, each representing an essential system ability. The four factors, or four essential abilities are (cf., Figure P.5):

- Knowing what to *do*, that is, how to respond to regular and irregular disruptions and disturbances either by implementing a prepared set of responses or by adjusting normal functioning. This is the ability to address the *actual*.
- Knowing what to look for, that is, how to *monitor* that which is or can become a threat in the near term. The monitoring must cover both that which happens in the environment and that which happens in the system itself, that is, its own performance. This is the ability to address the *critical*.
- Knowing what to *expect*, that is, how to anticipate developments, threats, and opportunities further into the future, such as potential changes, disruptions, pressures, and their consequences. This is the ability to address the *potential*.
- Knowing what *has happened*, that is, how to learn from experience, in particular how to learn the right lessons from the right experience—successes as well as failures. This is the ability to address the *factual*.

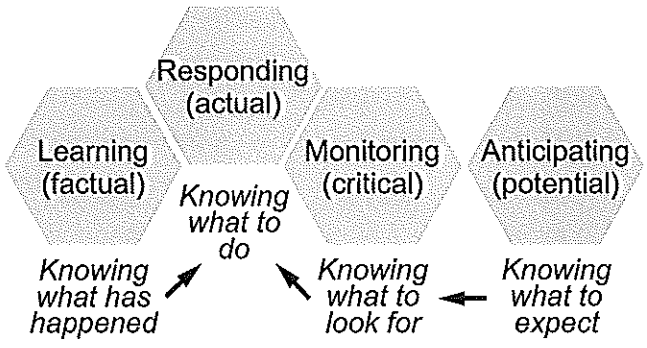

Figure P.5 The four cornerstones of resilience

If the resilience of a system is defined by the abilities to respond to the actual, to monitor the critical, to anticipate the potential, and to learn from the factual, an obvious question is how this can be brought about? This is really the question of how resilience can be *engineered* or the question of what Resilience Engineering is in practice. A detailed answer, or rather detailed answers, can be developed by considering each of the four factors in a more operational perspective. This quickly leads to a number of issues that can serve as the starting point for more concrete measures (cf., Epilogue). The focus on the issues arising from each of the four factors provides a way to think about Resilience Engineering in a practical manner. Starting from the level of the system as a whole this soon leads to the development of operational details and specific steps to be taken on a concrete level. This can, however, only be done by referring to a specific domain or field of activity, or even to a specific organisation at a certain time. Much of that may obviously make use of existing methods and techniques, although seen from a resilience perspective and in some cases be supplemented by new methods and techniques. For any given domain or organisation it will also be necessary to determine the relative weight or importance of the four main abilities, that is, how much of each is needed. The right proportion cannot be determined analytically, but must be based on expert knowledge of the system under considerations and with due consideration of the characteristics of the core business. Yet the minimum requirement is that none of the four can be left out if a system wants to call itself resilient.

Reading Guide

The chapters of this book have been organised in four main sections that correspond to the four main resilience abilities. While this organisation serves to emphasise how each ability can be considered in more detail, the chapters also make clear that Resilience Engineering cannot work by focusing on each of the four abilities in isolation. The four abilities depend on each other, and it is necessary to acknowledge and understand the dependencies or couplings among them in order successfully

to 'engineer' resilience. Corresponding to the 'new' definition of safety as the ability to succeed under varying conditions, the four abilities represent functions that can be improved, hence something that grows as the safety of a system gets better. Taken together, strengthening the abilities to respond, to monitor, to anticipate, and to learn is the best way to ensure that more things go right and that fewer things go wrong.