Network Manager
nominated by
the European Commission

EUROCONTROL

From Safety-I to Safety-II: A White Paper

DNM Safety

EXECUTIVE SUMMARY

In 2012, the aviation industry experienced the safest year on record, according to IATA, with a very low accident rate. This achievement is despite constant expansion in air traffic, which is expected to continue into the future. With this expansion comes a demand for parallel improvements in safety, so that the number of accidents does not increase.

Most people think of safety as the absence of accidents and incidents (or as an acceptable level of risk). In this perspective, which is termed Safety-I, safety is defined as a state where as few things as possible go wrong. According to Safety-I, things go wrong due to technical, human and organisational causes – failures and malfunctions. Humans are therefore viewed predominantly as a liability or hazard. The safety management principle is to respond when something happens or is categorised as an unacceptable risk. Accordingly, the purpose of accident investigation is to identify the causes and contributory factors of adverse outcomes, while risk assessment aims to determine their likelihood. Both approaches then try to eliminate causes or improve barriers, or both.

This view of safety was developed between the 1960s and 1980s, when performance demands were significantly lower and systems were simpler and more independent. It was assumed that systems could be decomposed and that the components of the system functioned in a bimodal manner – either working correctly or working incorrectly. These assumptions permitted detailed and stable system descriptions and enabled a search for causes and fixes for malfunctions. These assumptions do not fit today's world, where systems such as ATM cannot be decomposed in a meaningful way, where system functions are not bimodal, but rather where everyday performance is (and must be) variable and flexible.

Crucially, the Safety-I view does not explain why human performance practically always goes right. The reason that things go right is not people behave as they are told to, but that people can adjust their work so that it matches the conditions. As systems continue to develop, these adjustments become increasingly important for successful performance. The challenge for safety improvement is therefore to understand these adjustments, beginning by understanding how performance usually goes right. Despite the obvious importance of things going right, safety management has so far paid relatively little attention to this.

Safety management should therefore move from ensuring that 'as few things as possible go wrong' to ensuring that 'as many things as possible go right'. This perspective is termed Safety-II and relates to the system's ability to succeed under varying conditions. According to Safety-II, the everyday performance variability needed to respond to varying conditions is the reason why things go right. Humans are consequently seen as a resource necessary for system flexibility and resilience. The safety management principle is continuously to anticipate developments and events. The purpose of an investigation changes to understanding how things usually go right as a basis for explaining how things occasionally go wrong. Risk assessment tries to understand the conditions where performance variability can become difficult or impossible to monitor and control.

In light of increasing demands and system complexity, we must adapt our approach to safety. While many adverse events may still be treated by a Safety-I based approach without serious consequences, there is a growing number of cases where this approach will not work and will leave us unaware of how everyday actions achieve safety. The way forward therefore lies in moving toward Safety-II while combining the two ways of thinking. Most of the existing methods and techniques can continue to be used, although possibly with a different emphasis. But the transition toward a Safety-II view will also include some new practices to look for what goes right, focus on frequent events, remain sensitive to the possibility of failure, to be thorough as well as efficient, and to view an investment in safety as an investment in productivity. This White Paper helps explains the key differences between, and implications of, the two ways of thinking about safety.

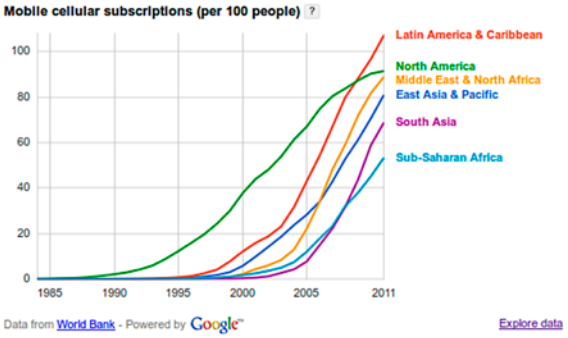
PROLOGUE: THE WORLD HAS CHANGED

To say that the world has changed is not just a phrase; it explains the intention of this White Paper and is also a teaser for the reader's thoughts.

It is a truism that the world we live in has become more complex and interactive and that this change continues at an increasing pace. This is perhaps most easily seen in the ways we communicate, both in the development in mobile phones and in the change from person-to-person calls to social media. Regarding this, Wikipedia says:

"Motorola was the first company to produce a handheld mobile phone. On April 3, 1973 Martin Cooper, a Motorola engineer and executive, made the first mobile telephone call from handheld subscriber equipment in front of reporters, placing a call to Dr. Joel S. Engel of Bell Labs. The prototype handheld phone used by Dr. Martin Cooper weighed 2.5 pounds and measured 9 inches long, 5 inches deep and 1.75 inches wide. The prototype offered a talk time of just 30 minutes and took 10 hours to re-charge."

Indeed, the situation today is that several countries have more mobile phones than people (Ireland, for instance, or Latin America). And according to a UN agency report there will be more mobile subscriptions than people in the world by the end of 2014.



Similar changes have taken place in aviation and ATM. One indication is that the number of flights has increased dramatically in the last 40 years. According to ICAO (2013), world scheduled air passenger traffic reached 5.4 trillion passenger-kilometres performed (PKPs) in 2012, a figure coming from around 5% annual growth over several decades. Similar growth is predicted in the coming years; Airbus (2012) predicts that traffic will double in the next 15 years, as it has every 15 years since 1980 . Just as the use of mobile phones has changed from being a privilege for the few to become a necessity for almost everyone, flying as a means of travelling has become a common good. Forty years ago the number of flights were few, airspace was not crowded, and the level of automation in the air and on the ground was low. System functions and components were loosely coupled, interactions were linear and understandable, and most situations could be fixed locally without major consequences or cascading effects on the overall system.

Today flying has become the normal way of travelling – for business and for private purposes - for most people, at least in the industrialised parts of the world. The increased demands for flights have introduced more competition between airlines and led to the development of larger airports - many the size of small cities. In the air, the increasing demands have led to a crowded airspace and a dense traffic flow, which in turn means more personnel and more extensive automation..

In the early days of commercial flying the focus of attention and of safety - so to say - was almost exclusively on technological failures. Components in the air or on the ground could break or not work as intended. The understanding of the how the technology worked was based on simple cause-effect relations and the models used to explain accidents or incidents were linear, e.g., Heinrich's Domino Model (Heinrich, 1931).

During the second half of the 20th century the technical environment changed and the focus of attention shifted from technological problems to human factors problems and finally to problems with organisations and safety culture. Unfortunately, most of the models actually used to analyse and explain accidents and failures did not develop in a similar way. The result is that safety thinking and safety practices in many ways have reached an impasse. This was the primary driver for the development of resilience engineering in the first decade of this century (e.g., Hollnagel, Woods & Leveson, 2006). Resilience engineering acknowledges that the world has become more complex, and that explanations of unwanted outcomes of system performance therefore no longer can be limited to an understanding of cause-effect relations described by linear models. The importance of resilience engineering for ATM has been explored in a previous White Paper (EUROCONTROL, 2009). This White paper takes a closer look at the consequences of resilience engineering concepts for safety thinking and safety management.

Figure 1: Heinrich's Domino Model (Heinrich, 1931)

SAFETY-I

Safety is a term that is used and recognised by almost everyone. Because we immediately recognise it and find it meaningful, we take for granted that others do the same and therefore rarely bother to define it more precisely. The purpose of this White Paper is to do just that and to explore the implications of two different interpretations of safety.

Safety-I: Avoiding That Things Go Wrong

To most people safety means the absence of unwanted outcomes such as incidents or accidents. Safety is generically defined as the system quality that is necessary and sufficient to ensure that the number of events that can be harmful to workers, the public, or the environment is acceptably low. ICAO, for instance, defines safety as:

The state in which the possibility of harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and safety risk management.

Historically speaking, the starting point for safety concerns has been the occurrence of accidents (actual adverse outcomes) or recognised risks (potential adverse outcomes). Adverse outcomes – things that go wrong – have usually been explained by pointing to their causes, and the response has been to either eliminate or contain the these. New types of accidents have similarly been accounted for by introducing new types of causes – either relating to technology (e.g., metal fatigue), to human factors (e.g., workload, 'human error'), or to the organisation (e.g., safety culture). Because this has been effective in providing short-term solutions, we have through centuries become so accustomed to explaining accidents in terms of cause-effect relations, that we no longer notice it. And we cling tenaciously to this tradition, although it has become increasingly difficult to reconcile with reality.

To illustrate the consequences of defining safety by what goes wrong, consider Figure 2. Here the thin red line represents the case where the (statistical) probability of a failure is 1 out of 10,000. But this also means that one should expect things to go right 9,999 times out of 10,000 – corresponding to the green area in Figure 2. (In aviation, the probability of being in a fatal accident on a commercial flight is 1.4 * 10-7.)

Figure 2: The imbalance between things that go right and things that go wrong

Safety efforts focus on what goes wrong, i.e., the one event out of 10.000, and this focus is reinforced in many ways. Regulators and authorities require detailed reports on accidents, incidents, and even so-called unintended events, and special agencies, departments, and organisational roles are dedicated to scrutinise adverse outcomes. Numerous models claim they can explain how things go wrong and a considerable number of methods are offered to find and address the causes. Accident and incident data are collected in large databases. Accidents and incidents are described and explained in thousands of papers, books, and debated in specialised national and international conferences. The net result is a deluge of information both about how things go wrong and about what must be done to prevent this from happening. The general solution is known as 'find and fix': look for failures and malfunctions, try to find their causes, and then eliminate causes and/or improve barriers.

The situation is quite different for the 9,999 events that go right. Despite their crucial importance, they usually receive little attention in safety management activities such as safety risk management, safety assurance and safety promotion. There are no requirements from authorities and regulators to look at what works well and therefore few agencies and departments that do it. Possible exceptions are audits and surveys, which may include a focus on strengths. However, on the whole, data are difficult to find, there are few models, even fewer methods, and the vocabulary is scant in comparison to that for what goes wrong. There are only few books and papers, and practically no meetings. Looking at how things go right also clashes with the traditional focus on failures, and therefore receives little encouragement. This creates a serious problem because we cannot make sure things go right just by preventing them from going wrong. We also need to know how they go right.

The current state of affairs represents a common understanding of safety, which we shall call Safety-I. This defines safety as a condition where the number of adverse outcomes is as low as possible. Since the purpose of safety management is to achieve and maintain that condition, safety goals are defined in terms of a reduction of the measured outcomes over a given period of time.

Safety-I promotes a bimodal view of work and activities, according to which acceptable and unacceptable outcomes are due to different modes of functioning. When things go right it is because the system functions as it should and because people work as imagined; when things go wrong it is because something has malfunctioned or failed. The two modes are assumed to be distinctly different, and the purpose of safety management is naturally to ensure that the system remains in the first mode and never ventures into the second (see Figure 3).

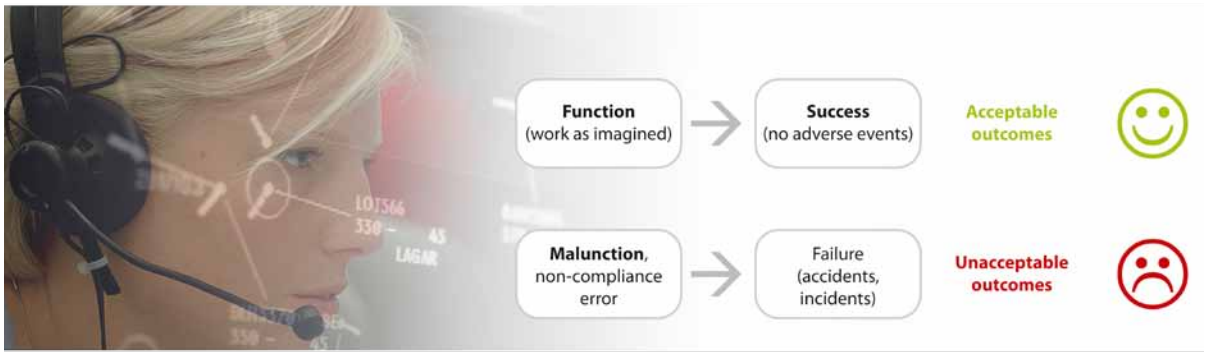


Figure 3: Things that go right and things that go wrong happen in different ways

In Safety-I, the starting point for safety management is either that something has gone wrong or that something has been identified as a risk. Both cases use the above-mentioned 'find and fix' approach. In the first case by finding the causes and then developing an appropriate response, in the second by identifying the hazards in order to eliminate or contain them. Another solution is to prevent a transition from a 'normal' to an 'abnormal' state (or malfunction), regardless of whether this is due to a sudden transition or a gradual 'drift into failure'. This is accomplished by constraining performance in the 'normal' state, by reinforcing compliance and by eliminating variability (see Figure 4). A final step is to check whether the number of adverse outcomes (airproxes, runway incursions, etc.) goes down. If nothing happens for a while, especially if for a long while, then the system is considered to be safe.

In the following, Safety-I will be characterised by looking at its *manifestations*, its underlying *mechanisms*, and its theoretical *foundations*.

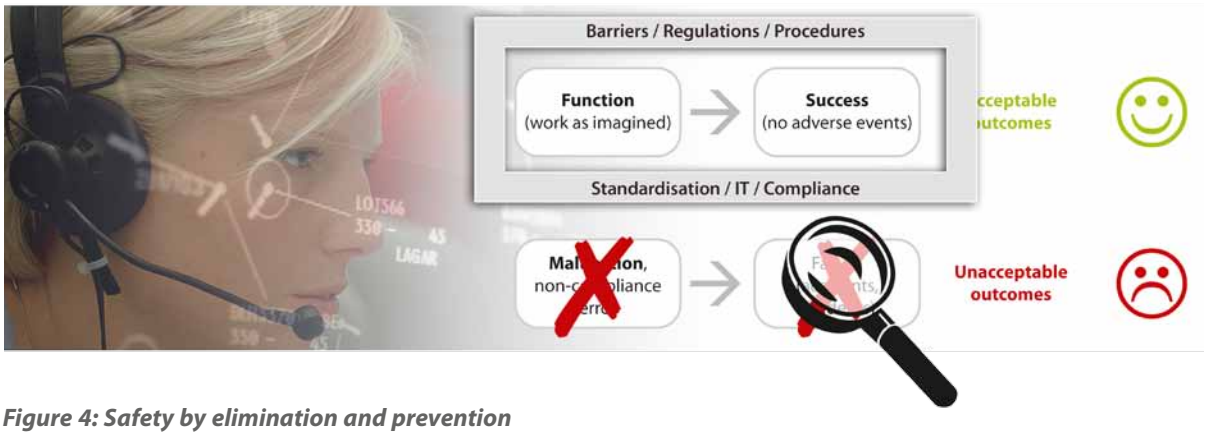


Figure 4: Safety by elimination and prevention

The Manifestations of Safety-I: Looking at what goes wrong

The definition of Safety-I means that the manifestations are adverse outcomes. A system (ATM) is deemed to be unsafe if adverse outcomes occur or if the risk is seen as unacceptable, and safe if no such outcomes occur or if the risk is seen as acceptable. This is, however, an indirect definition because safety is defined by its opposite, by what happens when it is absent rather than when it is present.

The curious consequence of this definition is that the level of safety is inversely related to the number of adverse outcomes. If many things go wrong, the level of safety is said to be low; but if few things go wrong, the level of safety is said to be high. In other words, the more manifestations there are, the less safety there is and vice versa. A perfect level of safety means that there are no adverse outcomes, hence nothing to measure. This unfortunately makes it impossible to demonstrate that efforts to improve safety have worked, hence very difficult to argue for continued resources.

To help describe the manifestations, various error typologies are available, ranging from the simple (omission-commission) to the elaborate (various forms of 'cognitive error' and violations or non-compliance). Note that these typologies often lead to a troublesome confusion between error as outcome (manifestation) and error as cause.

The 'Mechanisms' of Safety-I

The mechanisms of Safety-I are the assumptions about how things happen that are used to explain or make sense of the manifestations. The generic mechanism of Safety-I is the *causality credo* – a globally predominant belief that adverse outcomes (accidents, incidents) happen because something goes wrong, hence that they have causes that can be found and treated. While it is obviously reasonable to assume that consequences are preceded by causes, it is a mistake to assume that the causes are trivial or that they can always be found.

The *causality credo* has through the years been expressed by many different accident models. The strong version of the *causality credo* is the assumption about root causes, as expressed by root cause analysis. This corresponds to the simple linear thinking represented by the Domino model (Heinrich, 1931). While this probably was adequate for the first half of the 20th century, the increasingly complicated and incomprehensible socio-technical environments that developed in the last half – and especially since the 1970s – required more intricate and more powerful mechanisms. The best known accident model is the Swiss cheese model, which explains adverse outcomes as the result of a combination of active failures and latent conditions. Other examples are approaches such as TRIPOD (Reason et al., 1989), AcciMap (Rasmussen & Svedung, 2000), and STAMP (Leveson, 2004). Yet in all cases the mechanism includes some form of causation which allows the analysis to move backwards from the consequences to the underlying causes.

The Foundation of Safety-I: It either works or it doesn't

The foundation of Safety-I represents the assumptions about the nature of the world, so to speak, that are necessary and sufficient for the mechanisms to work. The foundation of Safety-I implies two important assumptions.

- Systems are decomposable.
  We know that we can build systems (including aircraft and airports) by putting things together, and by carefully combining and organising components. The first assumption is that this process can be reversed and that we can understand systems by

decomposing them into meaningful constituents (see Figure 5). We do have some success with decomposing technological systems to find the causes of accidents – the recent impasse of the Boeing 787 battery problems notwithstanding. We also assume that we can decompose 'soft systems' (organisations) into their constituents (departments, agents, roles, stakeholders). And we finally assume that the same can be done for tasks and for events, partly because of the seductive simplicity of the time-line. But we are wrong in all cases.

- Functioning is bimodal.

  It is also assumed that the 'components' of a system can be in one of two modes, either functioning correctly or failing (malfunctioning), possibly embellished by including various degraded modes of operation. System components are usually designed or engineered to provide a specific function and when that does not happen, they are said to have failed, malfunctioned, or become degraded. While this reasoning is valid for technological systems and their components, it is not valid for socio-technical systems – and definitely not for human and organisational components, to the extent that it is even meaningless to use it.

While the two assumptions (decomposability and bimodality) make it convenient to look for causes and to respond by 'fixing' them, they also lead to system descriptions and scenarios with illusory tractability and specificity, and quantification with illusory precision. They are therefore insufficient as a basis for safety management in the world of today.

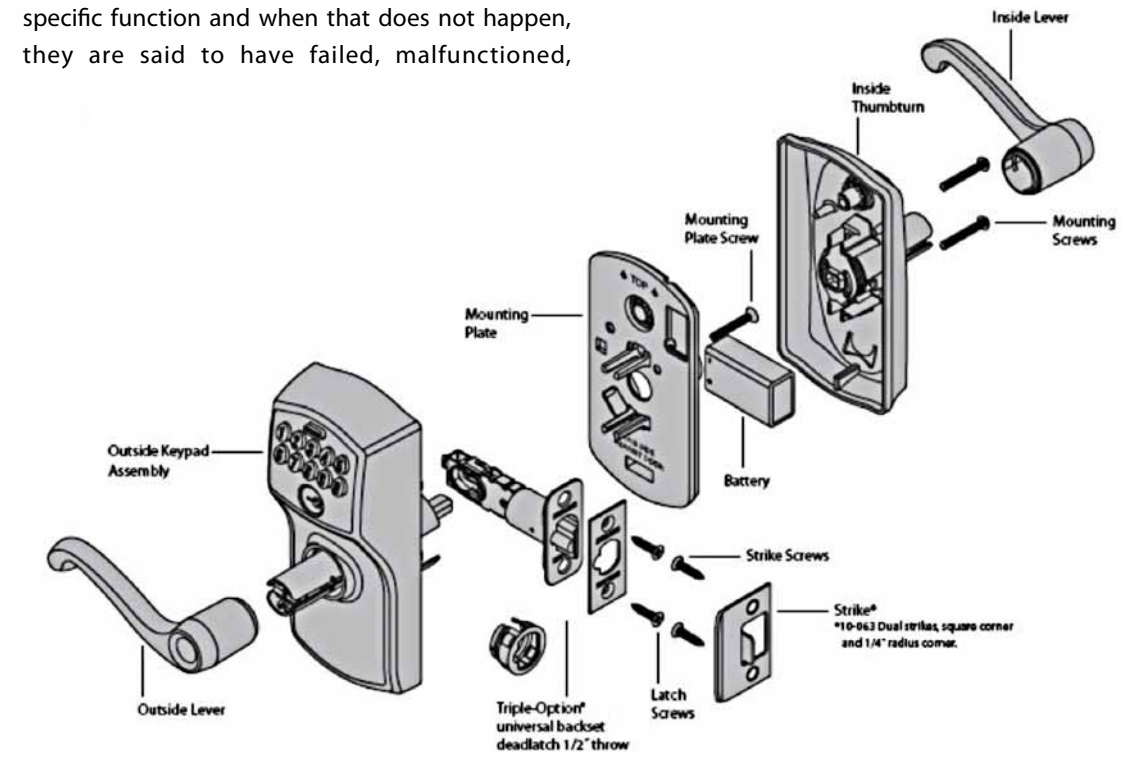
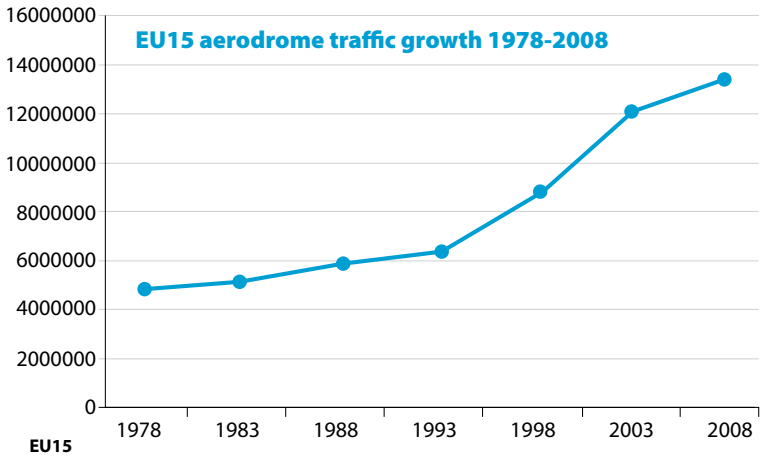


Figure 5: A decomposable system

THE CHANGING WORLD OF ATM

The ever changing demands on work, safety and productivity

Safety-I is based on a view of safety that was developed roughly between 1965 and 1985. The demands and conditions for work, in ATM and elsewhere, were vastly different then compared to now. Consider the growth in traffic in 15 EU states, shown in Figure 6.



Year	1978	1983	1988	1993	1998	2003	2008
	4874379	5210819	5916008	6362527	8804091	12044932	13338594

Figure 6: EU15 Aerodrome Traffic per year (1978 - 2008) (Data courtesy of ICAO)

The working conditions for controllers, as well as for pilots, were also quite unlike those of today. The evolution of ATC training, for instance, can be seen in Figures 7 to 9.

Industrial systems in the 1970s were relatively simple compared to today's world. The dependence on information technology was limited (mainly due to the size and the immaturity of the IT itself), which meant that support functions were relatively few, relatively simple, and mostly independent of one another. The level of integration (e.g., across sectors) was low, it was generally possible to understand and follow what went on, the traffic system was less vulnerable to disruptions, and support systems were loosely coupled (independent) rather than tightly coupled (interdependent). Safety thinking therefore developed with the following assumptions:

- Systems and places of work are well-designed and correctly maintained.
- Procedures are comprehensive, complete, and correct.
- People at the sharp end (operators) behave as they are expected to, and as they have been trained to. (They work as they are supposed or imagined to.)
- Designers have foreseen every contingency and have provided the system with appropriate response capabilities. Should things go completely wrong, the systems can degrade gracefully because the operators can understand and manage the contingencies.

While these assumptions probably never were completely correct, they were considered as reasonable in the 1970s. But they are not reasonable today, and safety based on these premises is inappropriate for the world as it is in the 2010s.

Figure 7: ATC training anno 1970 (Photo: DFS)

Figure 8: ATC training anno 1993 (Photo: DFS)

Figure 9: ATC training anno 2013 (Photo: DFS)

Rampant Technological Developments

Like most industries, ATC is subject to a steadily growing flow of diverse change and improvements. Some changes arise in a response to more powerful technology, but others are a response to increased performance demands. The European Commission has, for instance, set ambitious targets for developments over the next decade or so in relation to four directions: safety, environmental impact, capacity and cost-efficiency. (SESAR has more concretely set the following goals: to enable a threefold increase in capacity, to improve safety by a factor of 10, to cut ATM costs by half, and to reduce environmental impact by 10%. The safety target alone raises the interesting question of how you can measure a ten-fold increase in safety by counting how many things fewer that go wrong.)

The developments of industrial systems since the 1960 have been paced by a number of mutually dependent forces. The first is the development of technology itself, not least IT and 'intelligent' software. Then there is the increase in the needs and demands from users. This seems to follow the so-called 'Law of Stretched systems', which says that every system is stretched to operate at its full capacity and that any improvements, whenever they are made and for whatever reason, will be used to achieve a new intensity and tempo of activity. And finally, since we have long ago passed the point where unaided humans were able to control the processes, there is the ironic attraction of technological solutions, such as automation, to solve the problems. All this is illustrated in Figure 10 below.

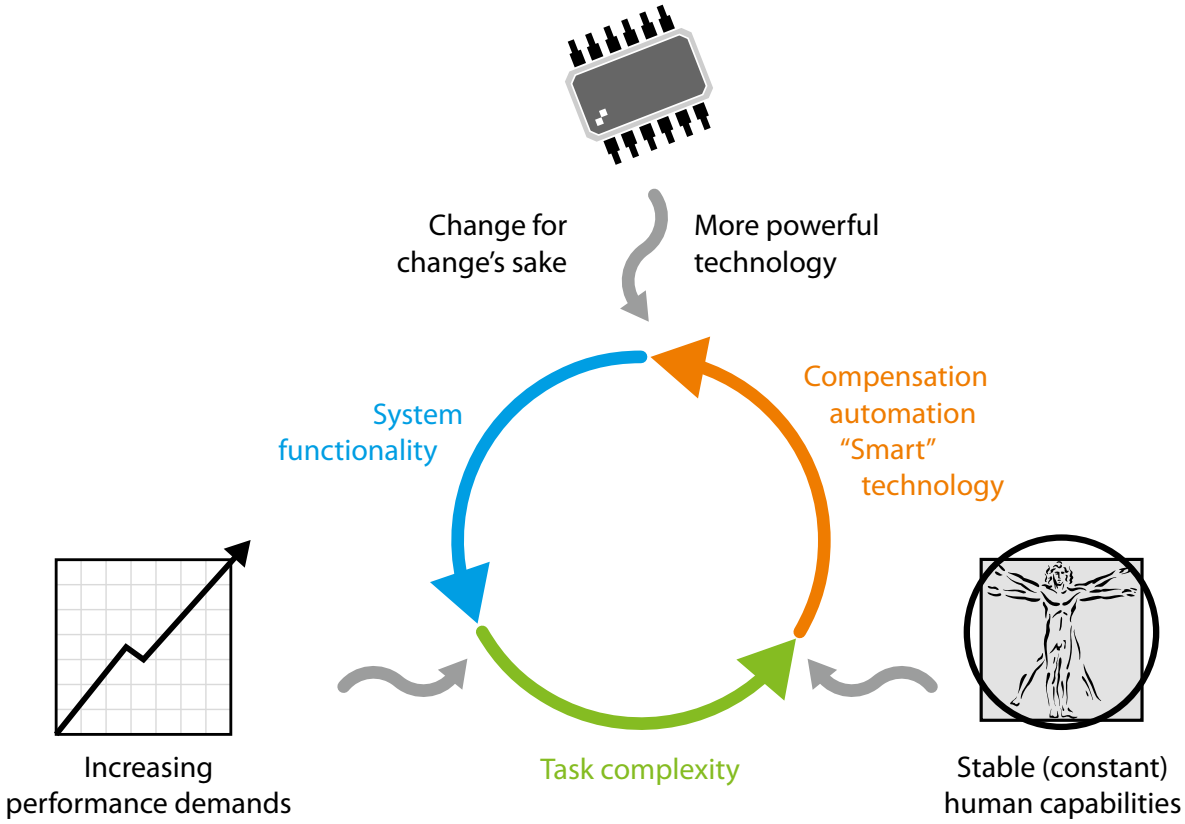


Figure 10: Self-reinforcing cycle of technological innovation

A different trend is the rise in cases where problems are selected based on just one criterion: whether they are 'solvable' with a nice and clean technological solution at our disposal. This has two major consequences. One is that problems are attacked and solved one by one, as if they could be dealt with in isolation. The other is that the preferred solution is technological rather than socio-technical, probably because non-technical solutions are rarely 'nice' and 'clean'.

The bottom line of these developments is that the world as we know it – and as it will be in a few years time in aviation and elsewhere – is a world where few things or issues are independent of each other. Functions, purposes, and services are already tightly coupled and the couplings will only become tighter. Consider, for instance, the four EC targets mentioned above relating to improved safety, reduced environmental impact, increased capacity and better cost-efficiency. While each target may seem plausible on its own, pursuing them in isolation will result in unintended consequences (Shorrock and Licu, 2013). A change to any of them will affect the others in ways which are non-trivial and therefore difficult to comprehend. This clashes with the assumptions of Safety-I, which means that any solution based on Safety-I thinking only will make things worse.

In consequence of rampant technological developments, of the widespread faith in nice and clean technological solutions, and of the general unwillingness to be sufficiently thorough up-front in order to be efficient later, our ideas about the nature of work and the nature of safety must be revised. We must accept that systems today are increasingly intractable. This means that the principles of functioning are only partly known (or in an increasing number of cases, completely unknown), that descriptions are elaborate with many details, and that systems are likely to change before descriptions can be completed, which means that descriptions always will be incomplete.

The consequences are that predictability is limited during both design and operation, and that it is impossible precisely to prescribe or even describe how work should be done. Technological systems can function autonomously as long as their environment is completely specified and as long as there is no unexpected variability. But these conditions cannot be established for socio-technical systems. Indeed, in order for the technology to work, humans (and organisations) must provide buffer functionality to absorb excessive variability.

The reasons why things work – again

Because the systems of today are increasingly tractable, it is impossible to provide a complete description of them or to specify what an operator should do even for commonly occurring situations. Since performance cannot be completely prescribed, some degree of variability, flexibility, or adaptivity is required for the system to work. People are therefore an asset without which the proper functioning would be impossible.

Performance adjustments and performance variability are thus both normal and necessary, and are the reason for both positive and negative outcomes. **Trying to achieve safety by constraining performance variability will inevitably affect the ability to achieve desired outcomes as well and therefore be counterproductive.** Thus rather than looking for ways in which something can fail or malfunction, we should try to understand the characteristics of everyday performance variability.

Work-As-Imagined and Work-As-Done

If the assumption that work can be completely analysed and prescribed is correct, then Work-As-Imagined will correspond to Work-As-Done. Work-As-Imagined is an idealistic view of the formal task that disregards how task performance must be adjusted to match the constantly changing conditions of work and of the world. Work-As-Imagined describes what should happen under nominal working conditions. Work-As-Done, on the other hand, describes what actually happens, how work unfolds over time in a concrete situation.

One reason for the popularity of the concept of Work-As-Imagined is the undisputed success of Scientific Management Theory (Taylor, 1911). Introduced at the beginning of the 20th century, Scientific Management had by the 1930s established time-and-motion studies as a practical technique and demonstrated how a breakdown of tasks and activities could be used to improve work efficiency. Scientific Management thus provided the theoretical and practical foundation for the notion that Work-As-Imagined was a necessary and sufficient basis for Work-As-Done. (Safety was, however, not an issue considered by Scientific Management.) This had consequences both for how adverse events were studied and for how safety could be improved. Adverse events could be understood by looking at the components, to find those that had failed, such as in root cause analysis. And safety could be improved by carefully planning work in combination with detailed instructions and training. This is recognisable in the widespread belief in the efficacy of procedures and the emphasis on compliance. In short, safety could be achieved by ensuring that Work-As-Done was made identical to Work-As-Imagined.

But the more intractable environments that we have today means that Work-As-Done differs significantly from Work-As-Imagined. Since Work-As-Done by definition reflects the reality that people have to deal with, the unavoidable conclusion is that our notions about Work-As-Imagined are inadequate if not directly wrong. This constitutes a challenge to the models and methods that comprise the mainstream of safety engineering, human factors, and ergonomics. A practical implication of this is that we can only understand the risks, if we get out from behind our desk, out of meetings, and into operational environments with operational people.

Today's work environments require that we look at Work-As-Done rather than Work-As-Imagined, hence at systems that are real rather than ideal. When such systems perform reliably, it is because people are flexible and adaptive, rather than because the systems are perfectly thought out and designed. Humans are therefore no longer a liability and performance variability is not a threat. On the contrary, the variability of everyday performance is necessary for the system to function, and is the source of successes as well as of failures. Because successes and failures both depend on performance variability, failures cannot be prevented by eliminating it; in other words, safety cannot be managed by imposing constraints on normal work.

The way we think of safety must correspond to Work-As-Done and not rely on Work-As-Imagined. Safety-I begins by asking why things go wrong and then tries to find the assumed causes to make sure that it does not happen again – it tries to re-establish Work-As-Imagined. The alternative is to ask why things go right (or why nothing went wrong), and then try to make sure that this happens again.

SAFETY-II

According to IATA (2013), close to 3 billion people flew safely on 37.5 million flights in 2012. This means that each day approximately 100,000 flights arrived safely at their destination. Sadly, there were also 75 accidents in 2012, with fifteen of them having fatal consequences (414 fatalities). However, the numbers show that an accident is a rare event – the equivalent of one accident for every 500,000 flights in 2012. For Western-built jets, the accident rate is lower still, with six hull-loss accidents in 2012 - equating to one accident for every 5 million flights.

Pilots, controllers, engineers and others can achieve these results because they are able to adjust their work so that it matches the conditions. (The same, of course, goes for all other industrial domains, although the accident rates may be different.) Yet when we try to manage safety, we focus on the few cases that go wrong rather than the many that go right. But focusing on rare cases of failure attributed to 'human error' does not explain why human performance practically always goes right and how it helps to meet ATM goals. Focusing on the lack of safety does not show us which direction to take to improve safety.

The solution to this is surprisingly simple: instead of only looking at the one case in 10,000 where things go wrong, we should also look at the 9,999 cases where things go right in order to understand how that happens. We should acknowledge that things go right because people are able to adjust their work to the conditions rather than because they work as imagined. Resilience engineering acknowledges that acceptable outcomes and unacceptable outcomes have a common basis, namely everyday performance adjustments (see Figure 11).

Because the ATM of today is intractable, it is impossible to prescribe what should be done in any detail except for the most trivial situations. The reason why people nevertheless are able to work effectively and successfully is that they continually adjust their work to the current conditions – including what others do or are likely to do. As systems continue to become more complicated and expand both vertically and horizontally, these adjustments become increasingly important for successful performance and therefore present both a challenge and an opportunity for safety management.

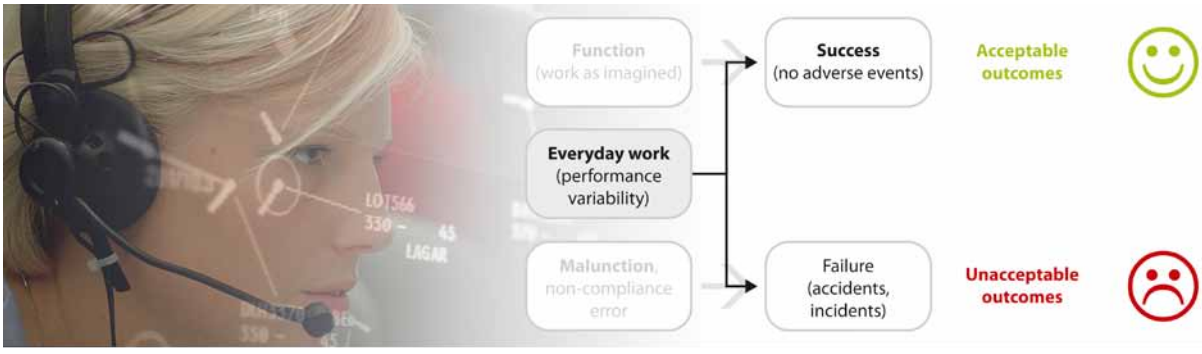


Figure 11: Things that go right and things that go wrong happen in the same way

According to this view we should avoid treating failures as unique, individual events, and rather see them as an expression of everyday performance variability. With the exception of major disasters, it is a safe bet that something that goes wrong will have gone right many times before – and will go right many times again in the future. Understanding how acceptable outcomes occur is the necessary basis for understanding how unacceptable outcomes happen. In other words, when something goes wrong, we should begin by understanding how it (otherwise) usually goes right, instead of searching for specific causes that only explain the failure (see Figure 12).

Work situations are increasingly intractable, despite our best intention to avoid that. One of the reasons for this is ironically our limited ability to anticipate the consequences of design changes or other interventions – both the intended consequences and the unintended side effects. This problem was addressed many years ago in a discussion of automation, where Bainbridge (1983) pointed out that "the designer who tries to eliminate the operator still leaves the operator to do the tasks which the designer cannot think how to automate". This argument applies not only to automation design but also to work specification and workplace design in general. The more complicated a work situation is, the larger the uncertainty about details will be.

The premises for safety management in today's industrial world can be summarised as follows:

- Systems cannot be decomposed in a meaningful way (there are no natural 'elements' or 'components').
- System functions are not bimodal, but everyday performance is – and must be – flexible and variable.
- Outcomes emerge from human performance variability, which is the source of successes as well as failures.
- While some adverse outcomes can be attributed to failures and malfunctions, others are best understood as the result of coupled performance variability.

In consequence of this, the definition of safety should be changed from 'avoiding that something goes wrong' to 'ensuring that everything goes right'. **Safety-II is the system's ability to succeed under varying conditions, so that the number of intended and acceptable outcomes (in other words, everyday activities) is as high as possible.** The basis for safety and safety management must therefore be an understanding of why things go right, which means an understanding of everyday activities.

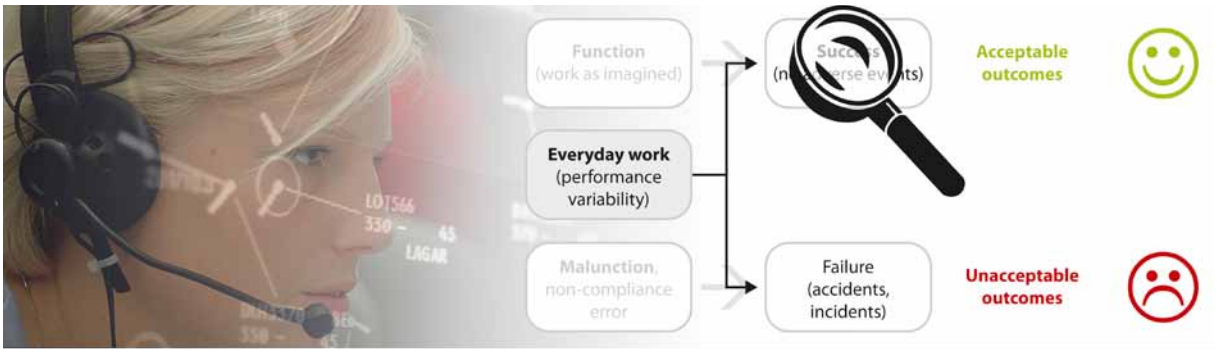


Figure 12: The basis for safety is understanding the variability of everyday performance

Ensuring that as much as possible goes right, in the sense that everyday work achieves its stated purposes, cannot be done by responding alone, since that will only correct what has happened. Safety management must instead be proactive, so that interventions are made before something happens and can affect how it will happen or even prevent something from happening. A main advantage is that early responses, on the whole, require a smaller effort because the consequences of the event will have had less time to develop and spread. And early responses can obviously save valuable time.

In the following, Safety-II is characterised by first looking at its theoretical *foundations*, then its underlying *mechanisms*, and finally its *manifestations*.

The Foundation of Safety-II: Performance Variability rather than bimodality

In contrast to the bimodality principle, the foundation of Safety-II is that performance adjustments are ubiquitous and that performance therefore always is variable. This means that it is impossible as well as meaningless to characterise components in terms of whether they succeed or fail, function or malfunction, etc. The variability should, however, not be interpreted negatively, as in 'performance deviations', 'violations', and 'non-compliance'. On the contrary, the ability to make performance adjustments is an essential human contribution to work, without which only the most trivial activity would be possible.

Technological systems are designed to have little or no performance variability. In order to function reliably, their environment must be highly stable. And in order for this to be the case, humans must ironically adjust their work to meet the demands of machines!

The 'Mechanisms' of Safety-II: Emergence rather than causality

Since performance adjustments and performance variability constitute the foundation of Safety-II, it follows that the mechanisms cannot rely on causality and linear propagations of causes and effects. Although it is still common to attribute a majority of adverse outcomes to a breakdown or malfunctioning of components and normal system functions, there is a growing number of cases where that is not so. In such cases the outcome is said to be emergent rather than resultant. This does not make it impossible to explain what happened, but the explanation will be of a different nature. The meaning of emergence is not that something happens 'magically,' but that it happens in a way that cannot be explained using the principles of decomposition and causality. This is typically the case for systems that in part or in whole are intractable.

The way we usually explain how something has happened is by tracing back from effect to cause, until we reach the root cause – or run out of time and money. This can be illustrated by a representation such as the fish bone diagram shown in Figure 13.

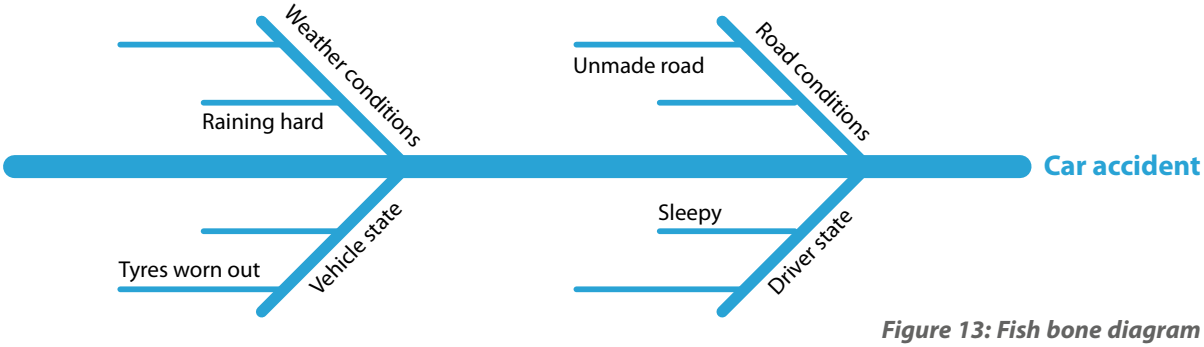


Figure 13: Fish bone diagram

This representation, and this way of thinking, starts from the fact that the final effects are (relatively) stable changes to some part of the system, that they are 'real'. The causes are also assumed to be stable, which means they can be found by going backwards from the outcome. The causes are 'real' in the sense that they can be associated with components or functions that in some way have 'failed,' where the 'failure' is either visible after the fact or can be deduced from the facts.

When something goes wrong, there will be an observable change of something. (Otherwise we could not know that anything had happened.) The outcome may be a runway incursion, a capsized vessel, or a financial meltdown. Safety-I assumes that the causes are real and the purpose of accident and incident investigation is to trace the developments backwards from the observable outcome to the efficient cause. Similarly, risk assessment projects the developments forward from the efficient cause(s) to the possible outcomes.

In the case of emergence, the observed (final) outcomes are also observable or 'real', but the same is not necessarily true for what brought them about. The outcomes may, for instance, be due to transient phenomena or conditions that only existed at a particular point in time and space. These conditions may, in turn, have emerged from other transient phenomena, etc (see Figure 14). The 'causes' are thus reconstructed (or inferred) rather than found. They may therefore be impossible to eliminate or contain in the usual manner, but it may still be possible to control the conditions that brought them into existence, provided we understand how work normally is done.

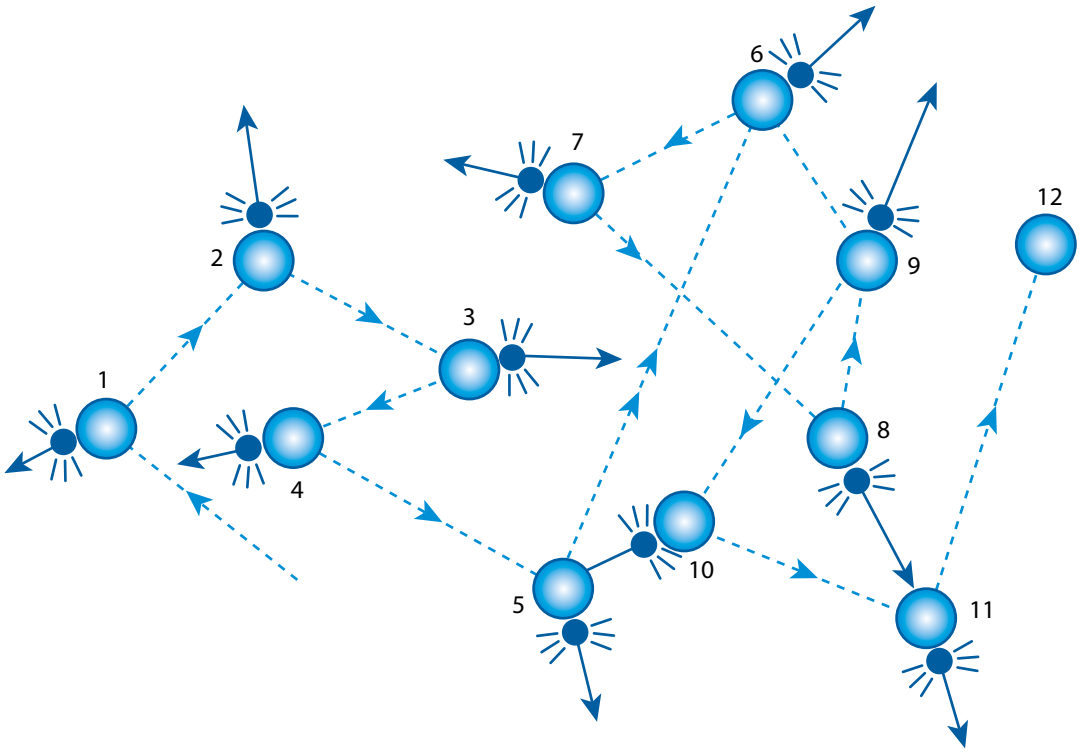


Figure 14: Transient phenomena and emergence

The Manifestations of Safety-II: Things that go right

The definition of Safety-II means that the manifestations are all the possible outcomes, as illustrated by Figure 15, and especially the typical or high frequency outcomes that usually are ignored by safety management. A system (ATM) is still deemed to be unsafe if adverse outcomes occur yet it is more important to understand how it is safe when things go right: safety is consequently defined by what happens when it is present, rather than by what happens when it is absent, and is thus directly related to the high frequency, acceptable outcomes. In other words, the more manifestations there are, the higher the level of safety is and vice versa. This makes it possible to demonstrate that efforts to improve safety have worked, hence easier to argue for continued resources. (It also resolves the possible conflict between safety and productivity, but that is another matter.)

To help describe the manifestations of Safety-II, few typologies are currently available. Even though things go right all the time, we fail to notice it because we become used to it. But since everyday performance is unexceptional, it can be explained in relatively simple terms. For instance everyday performance can be described as performance adjustments that serve to create or maintain required working conditions, that compensate for a lack of time, materials, information, etc., and that try to avoid conditions that are known to be harmful to work. And because everyday performance variability is ubiquitous, it is easier to monitor and manage.

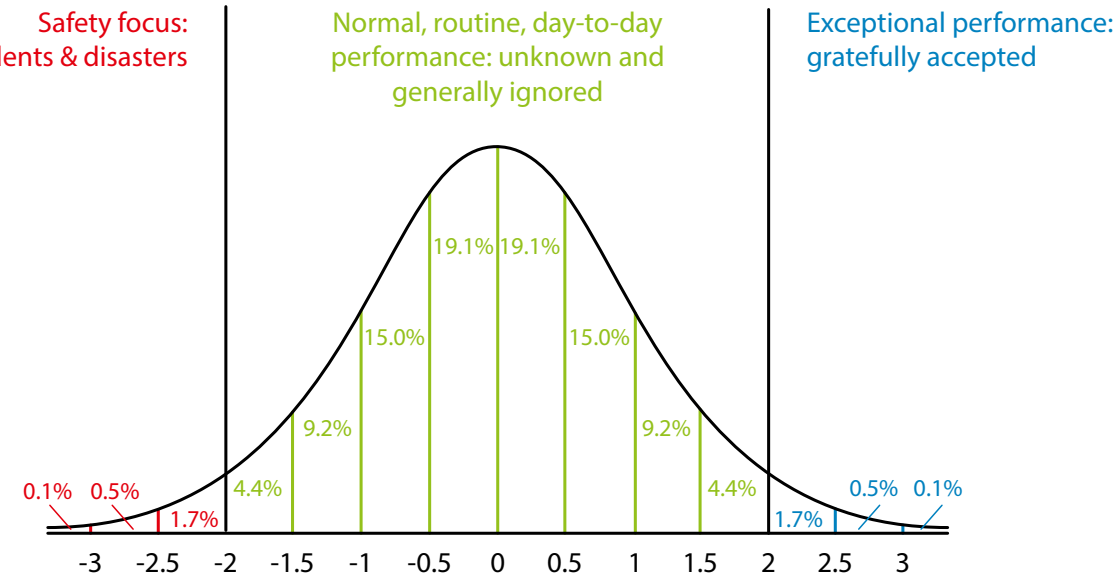


Figure 15: Event probability and safety focus

IV: THE WAY AHEAD

The main reason for juxtaposing Safety-I and Safety-II is to draw attention to the consequences of basing safety management on one or the other. The basic differences are summarised in the table below.

	Safety-I	Safety-II
Definition of safety	That as few things as possible go wrong.	That as many things as possible go right.
Safety management principle	Reactive, respond when something happens or is categorised as an unacceptable risk.	Proactive, continuously trying to anticipate developments and events.
View of the human factor in safety management	Humans are predominantly seen as a liability or hazard.	Humans are seen as a resource necessary for system flexibility and resilience.
Accident investigation	Accidents are caused by failures and malfunctions. The purpose of an investigation is to identify the causes.	Things basically happen in the same way, regardless of the outcome. The purpose of an investigation is to understand how things usually go right as a basis for explaining how things occasionally go wrong.
Risk assessment	Accidents are caused by failures and malfunctions. The purpose of an investigation is to identify causes and contributory factors.	To understand the conditions where performance variability can become difficult or impossible to monitor and control.

Table 1: Overview of Safety-I and Safety-II

What people do in everyday work situations is usually a combination of Safety-I and Safety-II. The specific balance depends on many things, such as the nature of the work, the experience of the people, the organisational climate, management and customer pressures, etc. Everybody knows that prevention is better than cure, but the conditions may not always allow prevention to play its proper role.

It is a different matter when it comes to the levels of management and regulatory activities. Here the Safety-I view dominates. One reason is that the primary objective of management and regulators historically has been to make sure that the customers or the public are not subjected to harm. Another reasons is that these levels are removed in time and space from the actual operation of the systems and services (e.g., ATM), and therefore have limited opportunity to observe or experience how work actually is done. A third reason is that it is much simpler to count the few events that fail than the many that do not – in other words an efficiency-thoroughness trade-off (Hollnagel, 2009). (It is also – wrongly – assumed to be easier to account for the former than for the latter.)

While day-to-day activities at the sharp end rarely are reactive only, the pressure in most work situations is to be efficient rather than thorough. This makes it less legitimate to spend time and efforts to digest and communicate experiences, since this is seen as being non-productive – at least in the short term. Successful safety management nevertheless requires that some effort is spent up front to think about how work is done, to provide the necessary resources, and to prepare for the unexpected. The pressure towards efficiency – such as SESAR's goal to cut ATM costs by half – makes this more difficult to achieve.

It can be difficult to manage safety proactively for the myriad of small-scale events that constitute everyday work situations. Here, things may develop rapidly and unexpectedly, there are few leading indicators, and resources may often be stretched to the limit. The pace of work leaves little opportunity to reflect on what is happening and to act strategically. Indeed, work pressures and external demands often necessitate opportunistic solutions that force the system into a reactive mode. To get out of this – to switch from a reactive to a proactive mode – requires a deliberate effort. While this may not seem to be affordable in the short term, it is unquestionably a wise investment in the long term.

It is somewhat easier to manage safety proactively for large-scale events because they develop relatively slowly – even though they may begin abruptly. (An example would be the eruption of a volcano that may lead to the closure of airspace. In other words, disturbances rather than disasters.) There are often clear indicators for when a response is needed. The appropriate response are furthermore known, so that preparations can be made ahead of time.

It is important to emphasise that Safety-I and Safety-II represent two complementary views of safety rather than two incompatible or conflicting approaches. Many of the existing practices can therefore continue to be used, although possibly with a different emphasis. But the transition to a Safety-II view will also include some new types of practice, as described in the following.

Look for what goes right

Look at what goes right, as well as what goes wrong and learn from what succeeds as well as from what fails. Indeed, do not wait for something bad to happen but try to understand what actually takes place in situations where nothing out of the ordinary seems to happen. Things do not go well because people simply follow the procedures and work as imagined. Things go well because people make sensible adjustments according to the demands of the situation. Finding out what these adjustments are and trying to learn from them is at least as important as finding the causes of adverse outcomes!

When something goes wrong, such as a runway incursion, it is unlikely to be a unique event. It is rather something that has gone well many times before and that will go well many times again. It is necessary to understand how such everyday activities go well – how they succeed – in order to understand how they fail. From a Safety-II view they do not fail because of some kind of error or malfunction, but because of unexpected combinations of everyday performance variability.

The difference between a Safety-I and a Safety-II view is illustrated by Figure 16. Safety-I focuses on events at the tails of the normal distribution, and especially events on the left tail that represent accidents. Such events are easy to see because they are rare and because the outcomes differ from the usual. They are, however, difficult to explain – the attractiveness of root causes and linear models notwithstanding. Because they are rare and because they are difficult to understand, they are also difficult to change and manage.

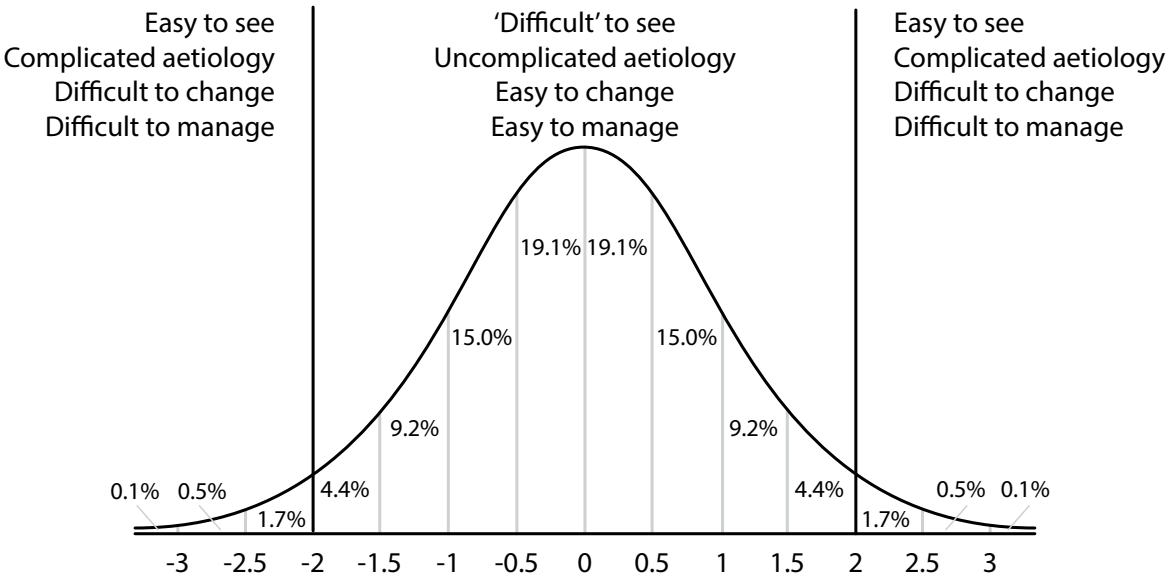


Figure 16: Relation between event probability and ease of perception

Safety-II focuses on events in the middle of the distribution. These are 'difficult' to see, but only because we habitually ignore them in our daily activities. The 'logic' seems to be that if something works, then why spend more time on it? But the fact of the matter is that they usually do not work in the way that we assume, and that Work-As-Done is significantly different from Work-As-Imagined. The events in the middle can be understood and explained in terms of the mutual performance adjustments that provide the basis for everyday work. Because they are frequent, because they are small scale, and because we can understand why and how they happen, they are easy to monitor and manage. Interventions are focused and limited in scope (because the subject matter is uncomplicated), and it is therefore also easier – although not necessarily straightforward – to anticipate what both the main and the side effects may be.

There is of course an evolutionary benefit in not paying attention (or too much attention) to the usual as long as it does not harm us and as long as environment is stable. But in our society, the environment is no longer stable. The work environment, and therefore also work itself, is increasingly unpredictable. This means that the routines that work well today may not work well tomorrow, and that it therefore is important to pay attention to how they work. This is the kind of thoroughness that enables us to be efficient when the time comes to make changes, and to make them rapidly.

Focus on frequent events

Look for what happens regularly and focus on events based on their frequency rather than their severity. Many small improvements of everyday performance may count more then a large improvement of exceptional performance.

The investigation of incidents is often limited by time and resources. There is therefore a tendency to look at incidents that have serious consequences and leave the rest to some other time – that never comes. The unspoken assumption is that the potential for learning is proportional to the severity of the incident/accident.

This is obviously a mistake. While it is correct that more money is saved by avoiding one large scale accident than one small scale accident, it does not mean that the learning potential is greater as well. In addition, the accumulated cost of frequent but small-scale incidents may easily be larger. And since small but frequent events are easier to understand and easier to manage (cf., above), it makes better sense to look to those than to rare events with severe outcomes.

Remain sensitive to the possibility of failure

Although Safety-II focuses on things that go right, it is still necessary to keep in mind that things also can go wrong and to 'remain sensitive to the possibility of failure'. But the 'possible failure' is not just that something may go wrong corresponding to a Safety-I view, but also that the intended outcomes may not be obtained, i.e., that we fail to ensure that things go right. Making sure that things go right requires an ongoing concern for whatever is working or succeeding, not only to ensure that it succeeds but also to counteract tendencies to employ a confirmation bias or to focus on the most optimistic outlook or outcomes.

In order to remain sensible to the possibility of failure, it is necessary to create and maintain an overall comprehensive view of work – both in the near term and in the long term. This can anticipate and thereby prevent the compounding of small problems or failures by pointing to small adjustments that can dampen potentially harmful combinations of performance variability. Many adverse outcomes stem from the opportunistic aggregation of short-cuts in combination with inadequate process supervision or hazard identification. Being sensible to what happens, to the ways in which it can succeed as well as the ways in which it can fail is therefore important for the practice of Safety-II.

Be thorough as well as efficient

If most or all the time is used trying to make ends meet, there will little or no time to consolidate experiences or understand Work-As-Done. It must be legitimate within the organisational culture to allocate resources – especially time – to reflect, to share experiences, and to learn. If that is not the case, then how can anything ever improve?

Efficiency in the present cannot be achieved without thoroughness in the past. And in the same way, efficiency in the future cannot be achieved without thoroughness in the present, i.e., without planning and preparations. While being thorough may be seen as a loss of productivity (efficiency) in the present, it is a necessary condition for efficiency in the future. In order to survive in the long run it is therefore essential to strike some kind of balance.

The cost of safety, the gain from safety

Spending more time to learn, think, and communicate is usually seen as a cost. Indeed, safety itself is seen as a cost. This reflects the Safety-I view, where an investment in safety is an investment in preventing that something happens. We know the costs, just as when we buy insurance. But we do not know what we are spared, since this is both uncertain and unknown in size. In the risk business, the common adage is 'if you think safety is expensive, try an accident'. And if we calculate the cost of a major accident, such as Air France 447 or Asiana flight 214, almost any investment in safety is cost-effective. However, since we cannot prove that the safety precautions actually are or were the reason why an accident did not happen, and since we cannot say when an accident is likely to happen, the calculation is biased in the favour of reducing the investment. (This is something that is typically seen in hard times.)

In Safety-I, safety investments are seen as costs or non-productive. Thus if an investment is made and there are no accidents, it is seen as an unnecessary cost. If there are accidents, it is seen as a justified investment. If no investments are made and there are no accidents, it is seen as a justified saving. While if accidents occur, it is seen as bad luck or bad judgement.

In Safety-II, an investment in safety is seen as an investment in productivity, because the definition – and purpose – of Safety-II is to make as many things go right as possible. Thus if an investment is made and there are no accidents, everyday performance will still be improved. If there are accidents, the investment will again be seen as justified. If no investments are made and there are no accidents, performance may remain acceptable but will not improve. While if accidents occur, it is seen as bad judgement.

Conclusion

Since the socio-technical systems on which our existence depends continue to become more and more complicated, it seems clear that staying with a Safety-I approach will be inadequate in the long run and probably in the short run as well. Taking a Safety-II approach should therefore not be a difficult choice to make. Yet the way ahead lies not in a replacement of Safety-I by Safety-II, but rather in a combination of the two ways of thinking (see Figure 17). It is still the case that the majority of adverse events are relatively simple – or can be treated as relatively simple without serious consequences – and that they therefore can be dealt with in ways that are familiar. But there is a growing number of cases where this approach will not work. For these, it is necessary to adopt a Safety-II view – which essentially means adopting a resilience engineering view. Safety-II is first and foremost a different way of looking at safety, hence also a different way of applying many of the familiar methods and techniques. In addition to that it will also require methods on its own, to look at things that go right, to analyse how things work, and to manage performance variability rather than just constraining it.

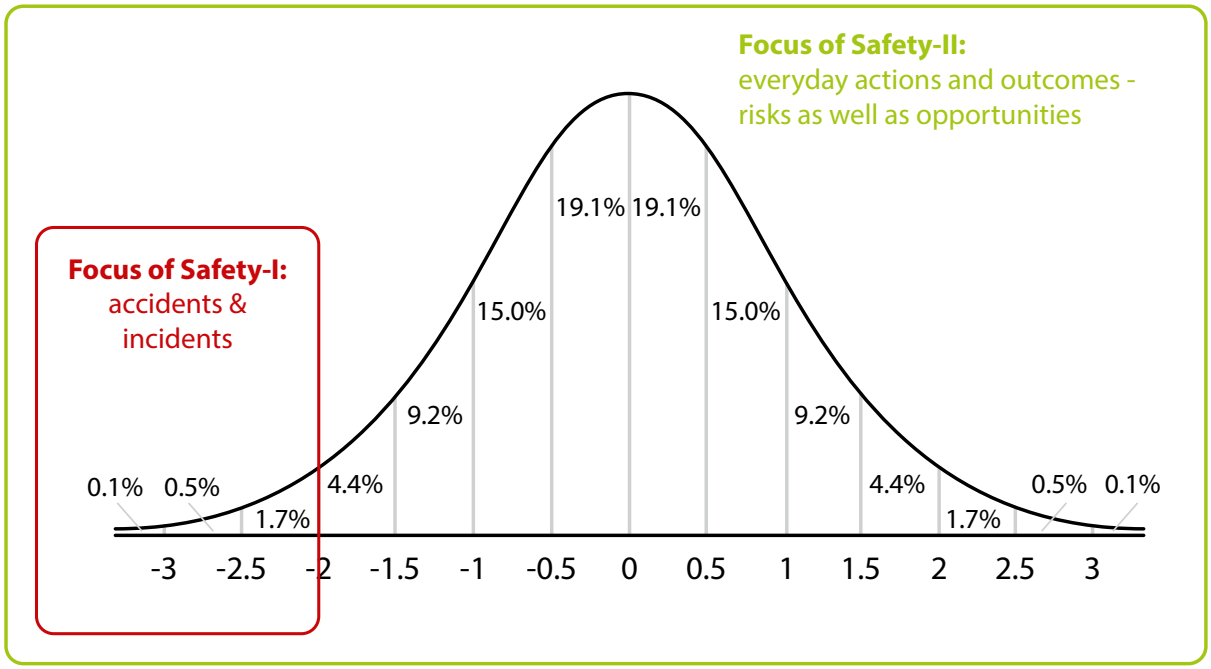


Figure 17: Focus of Safety-I and Safety-II

EPILOGUE

Introducing a different understanding of today's world and of the systems we work in and depend upon may require something akin to a paradigm shift. The safety community has developed a consensus on how things work and how safety can be ensured, but the increase of knowledge has levelled off. We must face the fact that the world cannot be explained by cause-effect models. Incidents and accidents do not only happen in a linear way, but include emergent phenomena stemming from the complexity of the overall aviation system. Asking for "why and because" does not suffice to explain the system in use and does not lead to an improvement in safety.

As a consequence of a paradigm change, safety experts and safety managers need to leave their 'comfort zone' and explore new opportunities. In that change, managers as well as practitioners are looking for models and methods to be used. Some methods already are available and have been applied in different settings.

The new paradigm also means that the priorities of safety management must change. Instead of limiting investigations and learning to incidents, a SMS should allocate some resources to the look at the events that go right and try to learn from them. Instead of learning from events based on their severity, the SMS should try to learn from events based on their frequency. And instead of analysing single severe events in depth, a SMS should explore the regularity of the many frequent events in breadth, to understand the patterns in system performance. A good way to start would be to reduce the dependency on 'human error' as a near-universal cause of incidents and instead understand the necessity of performance variability.

REFERENCES

Airbus (2012). *Navigating the future: global market forecast 2012-2031*.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775-779.

EUROCONTROL (2009). *A white paper on resilience engineering for ATM*. Brussels: EUROCONTROL.

Heinrich, H. W. (1931). *Industrial accident prevention: A scientific approach*. McGraw-Hill.

Hollnagel, E. (2009). *The ETTO principle: Efficiency-thoroughness trade-off. Why things that go right sometimes go wrong*. Farnham, UK: Ashgate.

Hollnagel, E., Woods, D. D. & Leveson, N. G. (2006). *Resilience engineering: Concepts and precepts*. Aldershot, UK: Ashgate.

IATA (2012) *2012 best in history of continuous safety improvements*. Press release no. 8. 28 February 2013.

ICAO (2013) *ICAO predicts continued traffic growth through 2015*. Press release 16 July 2013. Montréal.

Leveson, N. G. (2004). A new accident model for engineering safer systems. *Safety Science*, *42*(4), 237-270.

Rasmussen, J. & Svedung, I. (2000). *Proactive risk management in a dynamic society*. Karlstad, Sweden: Swedish Rescue Services Agency.

Reason, J., Shotton, R., Wagenaar, W. A., Hudson, P. T. W. & Groeneweg, J. (1989). *Tripod: A principled basis for safer operations*. The Hague: Shell Internationale Petroleum Maatschappij.

Shorrock, S. and Licu, T. (2013). Target culture: lessons in unintended consequences. *HindSight 17*. Brussels: EUROCONTROL.

Taylor, F. W. (1911). *The principles of scientific management*. New York: Harper.

GLOSSARY

(Approximate) Adjustments: When working conditions are underspecified or when time or resources are limited, it is necessary to adjust performance to match the conditions. This is a main reason for performance variability. But the very conditions that make performance adjustments necessary also mean that the adjustments will be approximate rather than perfect. The approximations are, however, under most conditions good enough to ensure successful performance.

Bimodality: Technological components and systems function in a bimodal manner. Strictly speaking this means that for every element *e* of a system, the element being anything from a component to the system itself, the element will either function or it will not. In the latter case the element is said to have failed. The bimodal principle does, however, not apply to humans and organisations. Humans and organisations are instead multi-modal, in the sense that their performance is variable – sometimes better and sometimes worse but never failing completely. A human 'component' cannot stop functioning and be replaced in the same way a technological component can.

Efficiency-thoroughness trade-off: The efficiency-thoroughness trade-off (ETTO) describes the fact that people (and organisations) as part of their activities practically always must make a trade-off between the resources (time and effort) they spend on preparing an activity and the resources (time, effort and materials) they spend on doing it.

Emergence: In a growing number of cases it is difficult or impossible to explain what happens as a result of known processes or developments. The outcomes are said to be emergent rather than resultant. Emergent outcomes as not additive, not predictable from knowledge of their components, and not decomposable into those components.

Intractable systems: Systems are called intractable if it is difficult or impossible to follow and understand how they function. This typically means that the performance is irregular, that descriptions are complicated in terms of parts and relations, and that it is difficult to understand the details of how the system works. Intractable systems are also underspecified, meaning that it is impossible to provide a complete specification of how work should be carried out for a sufficiently large set of situations.

Performance variability: The contemporary approach to safety (Safety-II), is based on the principle of equivalence of successes and failures and the principle of approximate adjustments. Performance is therefore in practice always variable. The performance variability may propagate from one function to others, and thereby lead to non-linear or emergent effects.

Resilience: A system is said to be resilient if it can adjust its functioning prior to, during, or following changes and disturbances, and thereby sustain required operations under both expected and unexpected conditions.

Resilience engineering: The scientific discipline that focuses on developing the principles and practices that are necessary to enable systems to be resilient.

Safety-I: Safety is the condition where the number of adverse outcomes (accidents / incidents / near misses) is as low as possible. Safety-I is achieved by trying to make sure that things do not go wrong, either by eliminating the causes of malfunctions and hazards, or by containing their effects.

Safety-II: Safety is a condition where the number of successful outcomes is as high as possible. It is the ability to succeed under varying conditions. Safety-II is achieved by trying to make sure that things go right, rather than by preventing them from going wrong.

Tractable systems: Systems are called tractable if it is possible to follow and understand how they function. This typically means that the performance is highly regular, that descriptions are relatively simple in terms of parts and relations, and that it is easy to understand the details of how the system works.

AUTHORS

Erik Hollnagel is Professor at the Institute of Regional Health Research, University of Southern Denmark (DK), Chief Consultant at the Centre for Quality, Region of Southern Denmark, and Professor Emeritus at the Department of Computer Science, University of Linköping (S). He has through his career worked at universities, research centres, and industries in several countries and with problems from many domains including nuclear power generation, aerospace and aviation, software engineering, land-based traffic, and healthcare. Erik's professional interests include industrial safety, resilience engineering, patient safety, accident investigation, and modelling large-scale socio-technical systems. He has published widely and is the author/editor of 20 books, including five books on resilience engineering, as well as a large number of papers and book chapters. The latest titles, from Ashgate, are "Resilient Health Care", "FRAM – the Functional Resonance Analysis Method," "Governance and control of financial systems", and "Resilience engineering in practice: A guidebook". Erik also coordinates the FRAMily (www.functionalresonance.com). sensei@functionalresonance.com

Jörg Leonhardt is Head of Human Factors in Safety Management Department at DFS - Deutsche Flugsicherung - the German Air Navigation Service provider. He holds a Master degree in Human Factors and Aviation Safety from Lund University, Sweden. He co-chairs the EUROCONTROL Safety Human Performance Su-Group and is the Project leader of DFS-EUROCONTROL "Weak Signals" project. joerg.leonhardt@dfs.de

with

Tony Licu is Head of Safety Unit within Network Manager Directorate of EUROCONTROL. He leads the support of safety management and human factors deployment programmes of EUROCONTROL. He has extensive ATC operational and engineering background and holds a Master degree in avionics. Tony co-chairs EUROCONTROL Safety Team and EUROCONTROL Safety Human Performance Sub-group. antonio.licu@eurocontrol.int

Steven Shorrock is Project Leader, Safety Development at EUROCONTROL and Adjunct Senior Lecturer at the University of New South Wales, School of Aviation. He is a Registered Ergonomist and a Chartered Psychologist with a background in human factors and safety management in several industries, government and academia. He holds a PhD in human factors in air traffic control. steven.shorrock@eurocontrol.int

ACKNOWLEDGEMENTS

The authors kindly acknowledge the comments and contributions of Christoph Peters (DFS), Shahram Etminam (DFS), Radu Cioponea (EUROCONTROL) and Seppe Celis (EUROCONTROL), as well as the reviewers of the document.

A NOTE FROM EUROCONTROL: HOW TO FIND OUT WHAT GOES RIGHT

A common question about Safety-II concerns the methods and approaches that can be used in practice. As mentioned in the Conclusion, many familiar methods can still be applied, but may need to be applied differently. They may need to be adapted or extended to widen the focus to include what goes right. In addition, other methods will be needed. The following gives some brief guidance on approaches for finding out what goes right. This will be expanded in a future guidance document. Note that these are not 'Safety-II methods', but are approaches that can be used to examine what goes right.

Safety observation. Several observation methods focus on observing everyday work and can be used to better understand what goes right. Examples include the 'Day-to-day safety survey' approach (used in NATS) and EUROSS (used in MUAC). Such approaches can be used to understand performance variability and adjustments in Work-As-Done. Time spent observing everyday work (for instance in operational areas), with or without a particular method, is particularly important for safety specialists to understand how things really work.

Safety investigation. Occurrence investigation focuses on what went wrong, but investigations can also focus on what goes right in the context of 1) adverse events (what went right during the event?), 2) ordinary work (how do things usually go right?), and 3) exceptional performance (why do things sometimes go exceptionally well?). The investigator's skills can therefore be employed to understand why things go right. Methods, taxonomies and language should be modified to be less failure-oriented and find out what goes right as well as wrong.

Safety assessment. Safety assessments can (and should) also focus on success as well as failure. Some safety assessment methodologies (such as the EUROCONTROL 'Safety Assessment Made Easier') have the capability to do this, for instance with an integrated 'success approach'. FRAM (Functional Resonance Analysis Method) can also be used proactively for safety assessment, using the idea of resonance arising from the variability of everyday performance.

Safety culture. Safety culture surveys or assessments, by whatever methods, should focus on strengths as well as weaknesses. More generally, they should help to understand everyday Work-As-Done. One such approach is the EUROCONTROL safety culture survey, which includes questionnaires, workshops, interviews and informal observations. The EUROCONTROL safety culture discussion cards can also be used to understand everyday work and to identify what goes right, for instance via storytelling or analysis of strengths and opportunities.

Safety development. Several organisational development approaches can be used with a focus on safety and what goes right. *Action research* is a participatory and collaborative approach to inquiry to generate new insights and bring about change in organisations. *Appreciative inquiry* (a particular form of action research) looks at what already works well in a system, rather than what does not work. *Storytelling* and *narrative approaches* provide a natural and memorable way of sharing experiences and knowledge about safety. Everyday experiences can be captured via written material, interviews, workshops, etc.

Team resource management (TRM). TRM involves strategies for the best use of all available resources - information, equipment and people - to optimise safety and efficiency. It should therefore be an ideal opportunity to understand how things go right in operations. Some time spent in TRM could be redirected to understand adjustments, performance variability and the efficiency-thoroughness trade-offs in operations.

More generally, many human factors methods for data collection, task and system analysis, and data analysis can be used to look at what goes right. Some of these are included within the EUROCONTROL HIFA website (Human factors Integration in Future ATM systems) See also the EUROCONTROL White Papers on 'Human Performance in Air Traffic Management Safety' , as well as 'Resilience Engineering for ATM'.

ANSPs will certainly already be undertaking some of the activities and using some of the approaches mentioned above. However, the question is whether the aim is that as few things as possible go wrong, or that as many things as possible go right. To focus on things going right, it is necessary to evolve our approaches to safety.

For further information contact:
Steven Shorrock steven.shorrock@eurocontrol.int
Tony Licu antonio.licu@eurocontrol.int

NOTES

EUROCONTROL

www.eurocontrol.int