Chapter 3
The Keys to a Successful Systemic Approach to Risk Management

Abstract It is universally admitted that an approach to safety applied to our complex industries (nuclear, chemical, construction and skilled trades) and services (medicine, banking and finance, public and private transport), can no longer be limited to finding local technical solutions; it absolutely must be systemic and global. How should these concepts be fleshed out? This chapter seeks to answer this question from various different perspectives, using examples taken from many contrasting areas, breaking down bias and prejudice and offering practical keys.

On Safety, Systems, Complexity … and the Structure of this Chapter

The management of risk in an enterprise is not only about avoiding or reducing accidents (affecting the system or those who work in it). It also concerns everything that may compromise survival, whether the threat is economic, political, social or damage to the image of the enterprise, particularly following an accident.

In order to understand a systematic approach, one must accept that risk management covers all the risks that could "kill" the enterprise, whether they are social, technical or financial.

The reduction of risks in a socio-professional system is therefore a complex concept, which can be defined differently depending on which perspective is adopted: fewer industrial accidents, fewer accidents affecting the installation, fewer risks of harm to social conditions and operations (no redundancies, protecting careers) or fewer risks to the business model (debt, profits, economic vulnerabilities).

> Risk management covers all the risks that could "kill" the enterprise. Safety, in terms of avoiding accidents, is only one such risk: other (economic and strategic) risks can sometimes kill the enterprise more quickly and may therefore take priority over safety when it comes to short-term investment.
>
> In this case, the art of risk management is to set priorities, make trade-offs, manage emergencies (doing effectively what one has decided to do),

R. Amalberti, *Navigating Safety*, SpringerBriefs in Applied Sciences and Technology,
DOI: 10.1007/978-94-007-6549-8_3, © The Author(s) 2013

but also remembering that some areas are neglected as a result and managing these in a specific way within the relevant divisions (a sound awareness of what is not being done, through the development of an awareness of these temporary vulnerabilities among both managers and operators, for example by strengthening detection and recovery systems when it is not possible to make further investments in prevention).

All these dimensions are legitimate, even though they often conflict with each other: the economic survival of the *business model* often involves increased exposure to the risk of an accident, which is managed more or less rationally and effectively after a setback (Fukushima is the extreme example of this).

This text endorses the perspectives of industrial safety and safe service provision (reduction of accidents, patient safety in medicine) by showing how the literature and experience on the ground are now making it possible to build a systemic approach which is effective, coherent and maintains control of the compromises that are made in relation to other areas of risk within the enterprise.

The key to the success of the systemic approach can be summarised by three complementary key points: (1) controlling the four stages of the trade-offs which are always present in building the safety structure of a complex system, (2) doing well what one has decided to do, and knowing and controlling what one has decided not to do, (3) future thinking rather than past thinking.

This text is structured around those three key points

The Swiss Cheese Model as the Archetype for Systemic Models … and Its Current Limitations

When speaking about the systemic model of risk management, everyone immediately thinks of the slices model set out by Reason in the 1980s [1, 2].

This model, based on three tenses, is simple and tells us: (1) that one cannot completely eliminate (patent) errors by people who are directly engaged in work, (2) that deep defences are needed to avoid the propagation of these errors as far as an accident, and (3) that it is necessary to be aware of organisational and management errors (latent errors) which, without being the immediate cause of accidents, increase the vulnerability of the individuals and defences directly engaged in the work by not giving them all the resources they need to be effective.

This model is always a heuristic one, and its author, whom I count as both a teacher and a friend, truly deserves his reputation and his global place in the pantheon of those who have contributed towards safety. It must also be admitted, however, that this model is not now sufficient to create a systemic approach that can offer effective safety in complex professional activities. It has four major defects:

- it reflects a linear model of the accident which is based on the propagation of failures in the structures and components making up the model; in this sense it

harks back to ideas that are already very old and are inspired by the domino model of Heinrich [3] or to the chain of errors, although it is more complex because it introduces the role of organisations and design[1];

- the model is still profoundly Cartesian, since it breaks down the universe of professional work into parts (structures and components, slices) and then attempts to find the vulnerabilities in each one of these parts; it explains the accident in terms of local vulnerabilities and leads to a search for "the error". The model certainly offers a vision of the whole by referring to the interactions and distortions in the alignment of the slices and vulnerabilities (it is important that failures should not be aligned), but it does not effectively take into account the risks of accidents where there are no failings in the individual parts but which are instead associated with weak links between parts that are not defective and emerging properties and risks of the "whole" (typically the global perspective on the system)[2];
- the model suggests that the identification and complete elimination of latent causes and exposures to risk are the (only) way to make progress in terms of safety. By doing this, it points us towards a model of avoiding or reducing exposure to risk in order to improve safety and eliminate all vulnerabilities, while living with voluntary exposure to risk (voluntarily creating holes in the slices) is a realistic factor that promotes survival for many enterprises. The world of safety must now accept this analysis and not reject it, since sociotechnical systems most commonly die because of their poor economic, organisational and political choices. Reason's model therefore provides keys to action which are valuable and are centred on a simple choice in favour of safety, but which are still inadequate in the real industrial world;
- finally, to reiterate the points above, he remains within the lines of traditional ideas and implicitly supports the idea that the best thing (for the enterprise) is always to achieve more safety, up to the total or virtually total elimination of accidents and incidents. This vision, which is acceptable and makes good sense in less safe systems, paradoxically reaches its limit in systems that have become very safe. The safer a system becomes, the more difficulty it has surviving these last accidents, and the more these last accidents tend to be exceptional accidents which are largely provoked by the system itself, which has become too safe, too rigid, too proceduralised and has, in short, lost its native resilience. We see this every day: ultra-safe industries have much greater difficulty justifying their safety policy (which everyone acknowledges to be effective, despite the very rare accidents that still occur), while fishing and road traffic have difficulty managing safety associated with their continuing succession of daily accidents [6]. Worse still, the search for less and less (accidents, incidents and errors) inspired by the nuclear industry, the aviation industry and the (small number of) ultra-safe industries, ultimately leads us to forget that there are other authentic models of safety (for example the HR0 model or resilience model) which have their own rules and their own contexts for progress, which are very well suited to the thousands of professional activities that are not subject to the demands imposed by our ultra-safe systems.

[1] This criticism is debated particularly effectively in Hollnagel et al. [4].

[2] This criticism is discussed particularly well in Dekker [5].

On this last point, it is also important to understand our own bias when it comes to building our knowledge base in the area of safety, which is associated with the professional applications that are used and studied in order to build these models of safety. Over 20 years (1990–2010) I have reviewed more than 2,072 conceptual articles addressing models of safety and their characteristics, published in eight international journals specialising in the areas of industrial safety, safety at work and safety in services (particularly medical, banking and road safety).[3] More than two-thirds of these articles (1,547) repeatedly address safety (safety at work or process safety) in five major industrial domains (nuclear industry, aviation industry, chemical industry, offshore and construction); almost a quarter of them (483), although in many cases these simply repeat and validate the models from the major industries, refer to medicine, banking and road safety (note from the author: the work done in this last field is rather more original). Very few (42) are original monographs on safety applied to skilled trade activities with low or very low levels of safety (mining, professional fishing, miscellaneous skilled trades). To some extent our understanding of risk and our models of safety largely result from the model of major industry, which is certainly responsible for the disasters highlighted in the media but not for the largest number of deaths, and which is ultimately very homogeneous in terms of its limitations and preoccupations (a level of safety between 10-5 and 10-7, highly developed regulatory requirements, priority given to human and organisational factors etc.); this model taken from major industries ultimately represents only a very small proportion of human professional activity on the planet. This recruitment bias accounts for many of our errors when we make generalisations about ideas relating to the safety of complex systems, by limiting our perspective to a narrow field and almost to a single specific case.

Controlling Systemic Safety: Four Key Steps for Building Safety in a Complex System

Improving safety in a complex system by adopting a systemic approach always requires a procedure that follows the same four steps: (1) knowing where the risks are, prioritising and building an ad-hoc system of defences, (2) setting this paper model alongside the real situation and making adjustments accordingly, particularly to various shifts in practices, (3) carrying out the analysis at a higher level and considering the macroeconomic and political constraints, (4) once all the preceding steps have been completed and the system is much more restricted and constrained, it is still necessary to ask how much resistance it still has to exceptional circumstances; so the question of resilience becomes central.

[3] Search of Google and (review summaries) in December 2011, limited to eight journals: Human factors, Safety Science, Ergonomics, Accident analysis and Prevention, Journal of Safety research, Journal of Risk research, International Journal Quality in Health Care, British Medical Journal Quality and Safety.

The ultimate level of safety observed at time T0 in a socio-technical system is always the result of a four-step construction process. These four steps are successive, but they do have feedback loops.

- The first step is always to identify the risks and establish an ideal model of defence. This is the classical field of risk mapping, decision-making matrices and more broadly systems reliability, extended to include human reliability. Once the risk has been identified and prioritised, this step leads to the definition of lines of defence (barriers) to reduce the occurrence of the accidents that are of concern.
- The second step is to set this ideal model alongside the real situation. In many circumstances, operators do not comply with this model and suffer no particular penalty, at least for a long time; many divergences occur for many different reasons, and it is useful to understand these. This migration in practice will sooner or later lead to incidents and accidents. The end of this phase therefore necessary involves a feedback loop to adjust the ideal model, but the way in which the ideal model is interrogated and corrected in a feedback loop is far from always being identical. In the majority of cases those concerned make the mistake of considering that the ideal model should not be called into question and that it is only necessary to strengthen its defences or the authority exerted over the operator to induce him to follow the instructions and procedures. The idea of strengthening the operator's "safety culture" so that he will follow the script set out in the ideal model that one believes in, is one path which is often adopted in parallel with a purely procedural and regulatory hardening of the ideal model to force compliance with it. Perhaps, however, some of the fundamentals of this ideal model should be questioned rather than seeking to make it stick to the real situation.
- The third step is systemic in nature. No-one imagines that a complex system can be made completely safe simply by relying on putting procedures and recommendations in place and forcing only front-line operators to comply strictly with these good practices and recommendations. A further step needs to be taken to strengthen what one might call the "system", and this step is based on a strategy of "safe governance" of this system, and action involving the middle and top management: how to conceive a safe structure for the system, the relationships between bodies, professions, the specific interests of each one, directorates, divisions, branches and subcontractors, the levels at which each actor should be independent or perhaps dependent, what risks should be taken and how, when controlling the compromises that are made between economy, profitability, certainty and safety.
- The fourth step concerns the ultimate resilience of the resulting model. Safety is often a problem that is even more difficult to control once all the preceding steps have been taken, once the system has become safe or ultra-safe, once the procedures, safety instructions and protective measures have reasonably been followed and adopted by front-line operators, once the management is personally involved in the decisions and trade-offs in favour of safety, sometimes accepting sacrifices in terms of profitability and once a culture of total reporting is established. In fact no safe system can be completely protected from disasters, and although these are

certainly rarer, they are infinitely more damaging to the image of an enterprise or activity that has become safe, to the point that it more easily results in a crisis and sometimes in the death of the enterprise. If the same accident had occurred earlier in the history of the same enterprise, at a time when it was less safe, it no doubt would not have resulted in the same consequences. In the end the excellence of the level of safety that has been reached becomes the mirror for the model which is predicted and looked for; but the system will die of something that has not been predicted. Adaptation to exceptional circumstances is never written into the strictly procedural models. The system is no longer robust in the face of rare challenges and loses its "resilience". This section explains everything that is paradoxical about this final step. The more proceduralisation occurs, the more conformity and safety is achieved in relation to an ideal model and, unfortunately, the more professionals and managers become "deskilled". They are rarely exposed to difficult situations; they lose the habit of making decisions that involve sacrifices in contradictory dimensions (which are characteristic of short-term survival under these difficult conditions). In brief, resilience is a property which is relatively native to less safe systems in which the operators are exposed to highly variable situations. It reduces once the system is made safe using the previous three steps, and at the end of the process it diminishes to the point where it needs to be reinforced using specific mechanisms in systems that have become ultra-safe. Unfortunately this phase of re-inserting resilience is often delicate or even impossible to achieve because it runs counter to the solutions that have been used to strengthen the ideal model and achieve the current level of safety.

We will now look in detail at the content of each one of these steps, what can be achieved by them and the respective challenges facing each one.

Step 1: Evaluate the Risk and Build a Defensive Fortress

It is impossible to initiate a safety process without evaluating the risk and protecting oneself against what appears to be a threat.

The tools used to evaluate risk, both a priori (on the basis of a systematic analysis of the vulnerabilities of the system in question) and a posteriori (based on the analysis of accidents and incidents that have really occurred) are the toolkit of reliability engineers; they are well-known and they are not described in this work. We refer the reader back to the plethora of literature that exists on this subject and we will only cite a small selection of the references, which is no doubt incomplete but is nevertheless sufficient to obtain an overview of the principal methods used.[4]

[4] Numerous references exist on this subject, most of them dating back some time. Many different summaries exist on a number of websites (although this list is not exclusive) http://pachome1.pacific.net.sg/~thk/risk.html accessed on 27 décembre 2011 http://www.statcart.com/ viewed on 27 December 2011or the remarkable summary with reference to the medical field, which was published in French in five articles by a group of authors, specifically [7, 8].

Principal methods of risk analysis

A priori analysis

Analysis of processes	Tool	Usefulness	Limitations
	Process analysis	A prerequisite for other processes (FMEA/FMECA, PRA, HACCP) in order to describe an activity that brings together all the functional constraints (flows, resources etc.) in order to identify critical points and improve the steps in its functions, particularly with regard to interfaces between services	Requires a knowledge of the need that has to be met in relation to the available resources
	Functional analysis	A prerequisite for other processes (FMEA/FMECA, PRA, definition of the need to be met, value analysis) involved in introducing and implementing a new activity (service or product), setting out the functions needed to achieve this, starting from design, applicable constraints and performance criteria	Complex, time-consuming process suitable for designing a new activity forgetting a function
Example of the use of WHAT if type techniques	HAZOP *Hazard and operability study*	Breaking down the system that is to be constructed into parts, referred to as "nodes", then, using a keyword-based guide, brainstorming on possible deviations, assessment of the potential risks to human safety, property and the environment	This process is well adapted to anticipatory work on new systems. Extensively used in the chemical industry. The quality of the descriptions reflects the quality of the participants
Example of use of WHAT if type techniques	HACCP *Hazard analysis control critical point*	Interactive method involving experts, well codified with 7 steps, applicable to both existing and prospective systems	Authoritative in the food industry
	PRA *Preliminary Risk Analysis*	Identifies accident scenarios in the presence of danger, evaluates them, establishes a hierarchy between them and makes deductions concerning risk control in a relatively exhaustive way	Complex and time-consuming process Applications in ultra-safe or very high-risk industries (space travel)

(continued)

A priori analysis			
Analysis of processes	Tool	Usefulness	Limitations
	FMECA Failure mode effect and criticality analysis FMEA Failure mode and effect analysis	A method to ensure safe functioning which allows methodological analysis of a critical process, from description of the steps involving risk to measurement of the criticality of causes (frequency multiplied by severity) FMEA: a simplified approach to the FMECA method which can be used in the absence of quantified data	Selection of critical processes A lengthy process, reserved for intrinsically very high-risk processes
Instrument panel monitoring	Instrument panel warning indicators	Possible method for monitoring an action to control risks Makes it possible to detect risks, anticipate and make decisions	Data gathering can be burdensome
Surveys in comparison with a frame of reference	Audit	Verifies the implementation and effectiveness of defined measures Simple to understand *Specific example: the LOSA Line-Oriented Safety Audit: which is the method recommended by the ICAO, relies on auditors seated behind the pilots on a jump seat, entering data on errors made by the crew* Other examples: Audits of human and organisational factors	Requires programming and preparatory work (auditor training) and a validated, professional frame of reference. Not well suited for organisational analysis The subsequent methodological analysis is essential in order to benefit from any observation of a dysfunction
	Risk visits, Walk rounds Health And Working Conditions Monitoring Committee	Lever for possible triggering of an institutional process (for example insurance) and action by the management One particularly well-known version of such visits is applicable to action taken in partnership (between trade unions and management) to improve safety in working conditions	Requires follow-up of the observations and a detailed organisation in order to be useful

(continued)

(continued)

A priori analysis			
Analysis of processes	Tool	Usefulness	Limitations
Computer-assisted techniques	Go method Markov modelling Dynamic Event logic Analytical methodology	All these techniques use mathematical and computerised formalisation of the process to automatically create graphs showing the propagation of risk starting from an event	Complex and reserved for very specific applications
A posteriori analysis			
Accident analysis or analysis of near misses	In-depth analysis	Looking at the causes of accidents One example which is used particularly frequently in medicine: the ALARM (Association of Litigation and Risk Management) method, which is inspired by Reason's model and focuses on looking for latent errors	Simple but leaves considerable scope for interpretation by analysts
	Tree based techniques: fault tree analysis, event tree analysis, MORT-Management Oversight Risk tree	A series of methods that start from the event and use logical links to establish the causes; the analysis may also start from a sentinel event and seek to understand its potential consequences	Quite time-consuming if the analysis goes beyond the superficial level, requires a favourable environment and sufficient authority to push through the conclusions

In this phase 1 we will only debate a small number of points that give rise to difficulties, which are both specific to this initial phase and strategic in terms of decision-making and trade-offs.

What reference framework should be used to characterise and measure risk?

The analysis, characterisation and measurement of risk call into question our understanding of the scope of the domain that is considered to be relevant for the purpose of explaining risk. To some extent our scientific analysis of risk, which is intended to allow us to measure risk using the most formalised methods possible, depends on subjective values from the outset.

To use a simple (simplistic) example: a bank wishes to produce a map of its financial risks for its "financial products" division. The traditional analysis will cover the range of processes that are used by the trading departments to interact in the markets: organisation of the department, order flows, rules of engagement, relevance of the mathematical risk models used, IT tools, delegation of engagement, supervision and control. If only this domain is covered, one can only have a relatively limited, technical perspective on an internal process within the bank which is specific to trading. It is easy to imagine, however, that the real risk may depend more on global political balances than on trading techniques within the bank. Extending the perimeter of technical risk analysis to include the analysis of political risk at the national level or even globally changes the model of the processes that need to be considered to feed into the charting process, alters the results of the HAZOP, FMECA and PRA models that characterise risk and ultimately changes part of the measurement that is considered to be relevant to this risk.

We understand that an excessively narrow scope of analysis can quite easily result in an analysis of risk that covers only a fraction of the real risk and sometimes quite a marginal fraction.

Taking this one step further: changing to a different risk analysis

Medicine is constantly evaluating the risks and the efficacy of its strategies. Let us consider the strategies for the management of obesity in children. An analysis is carried out of the risk and benefit of the medicines that are given and the management of this problem. It is known, however, that lifestyle habits in underprivileged social contexts and industrial pressure from soft drinks and confectionery manufacturers (which cumulatively promote unbalanced nutrition in schools and at home) represent a potential source of much greater risks in terms of obesity than poor medical management. In this case, the perimeter that needs to be considered in order to prevent the risk of obesity will benefit from consideration of the wider range of risks associated with social action rather than only the risks associated with the limited domain of medical action.

This simple example points to one of the core principles of cost-benefit analysis and economic analyses in the area of safety: how to consider safety from a perspective of greater efficiency, how to achieve better and safer production, at the same cost or if possible even at a lower cost.

This type of analysis is not new, but it is always quite difficult to put into practice because it specifically forces us to extend the perimeter of risk analysis far beyond the limited technical domain that triggers the analysis of risk to its own process in order to evaluate its own work.

Typically this leads to a systemic approach.

For example, an article written some time ago [9] proposes an inventory of 500 high-risk human activities which are evaluated in terms of QUALYs (life-years lost engaging in the activity or passively exposing oneself to risk—those living close to industrial installations or others). Without discussing in detail the method that is used in the comparison, the article does address a fundamental question: everyone carries out their own risk analysis within their own small domain (of benefit) and therefore creates local solutions for risk reduction and risk contingency planning that are sometimes extremely complex and expensive. Taking a "bird's eye view", however, this landscape appears to be divided into hundreds of separate compartments, each one defended from its own risks, which raises questions about the relevance of an approach to the risks that exist within the compartment.

When analysing risk, it should be possible to consider the offsetting of risks and the comparison of risks with other compartments within the same domain, with other alternative solutions (in the case of obesity above… or in a more difficult case with an even more sensitive taboo, the case of risks associated with commercial activity carried out by tele-presence in a virtual environment, as compared with the risk that is accepted in terms of aviation safety, to deal with the same problem by physically transporting operators). Reticence to engage with these thought-processes is understandable, since each model of intra-compartment risks corresponds to a business model that has little interest in a more global perspective that could be harmful to its business. The example of nuclear risks or difficult very deep water oil operations, as compared with alternative energy technologies, shows how it is necessary to go as far as disasters and beyond in order to truly accept this global perspective in a fully transparent way.

The same is true when considering the time-horizon and the capacity for recovery and attenuation. Risk analyses take little account of the often positive trade-offs involved in taking short-term risks in order to safeguard the long term more effectively. Let us imagine a risk matrix which considers an intervention in a factory in a difficult context where there is a leak from a pressurised valve. Acceptance of the immediate risk will affect the long-term risk. Even if the intervention results in an industrial accident in the short term, it may have a considerable cost benefit overall in the long term, since it avoids a more significant breakdown in the plant and no doubt other severe consequences as well. Is it necessary to prohibit taking immediate risks in order to achieve long-term safety in this way? And where is the critical time-horizon?

On closer inspection it becomes clear that managing risks does not always mean reducing them, but it often means exchanging them for other risks and for risks at different times.

These trade-offs result in both gains and losses, depending on the perimeter and the time-horizon under consideration.

For the safety division in the factory, the industrial accident that occurs during a difficult intervention during which safety codes are not complied with, will almost always be considered to be the result of poorly controlled risk management, even if it safeguards the long-term situation. The only exception is in the emotional interpretation of the event, if it is clearly established that it was a heroic action (in other words the benefit is clearly and immediately identifiable in the very short term, for example saving an injured employee in a toxic environment without protection, where the person saving him failed to comply with instructions). Exactly the same logic is applicable to lymphangitis in the arm caused by an infusion of medication which is intended to treat cancer but which one day "bypasses" the vein. The immediate effect is disastrous for the patient: extreme pain, a swollen arm, disability, a number of weeks taken recover, but the long-term effect will be absolutely minor. The overall effect of treating the cancer will be on a much larger scale than this incident along the way which has no longer-term consequences.

Ultimately the difficulty consists in having a system that accepts the dynamic and intelligent nature of such trade-offs. This widening of the perimeter is generally impossible due to the fact that the system primarily consists of human beings, careers, individual attitudes to be justified, conflicting financial interests and power bases, and it is important to recognise that the ultimate beneficiary of the long-term exchange of risks is rarely the person who has to make the decision to accept the short-term risk.

The question of choosing the perimeter can be asked in different forms, including the question of the social judgement that will alter the analysis of the risk acceptance matrix.

Let us imagine an amateur mountaineer, who is walking in the mountains and climbing risky and technically difficult slopes. He exposes himself to a risk which he knows to be very high because his cost-benefit analysis (pleasure) is positive. If this analysis is extended to his family circle, the cost-benefit assessment by his wife will necessarily be different, giving a much lower weighting to the pleasure and putting a higher weighting on the effect of his absence and the risk of an accident. If it is extended to society, however, the cost-benefit analysis will result in a very low positive value (positive economic influence on high mountain resorts), for a very large negative value: the cost of providing assistance, cost of disability. Depending on the chosen perimeter, the analysis therefore leads to different results in terms of acceptance of the risk matrix.

What is the place of voluntary (or compulsory) reporting of incidents during this first phase?

There is a huge amount of literature on difficulties associated with reporting in all industries and services,[5] [10]; the observations are often the same: massive underreporting, due to (1) fear of consequences (sometimes legal, but above all internal

[5] Obstacles to participation: the top 9 reasons why workers don't report near misses, 2011, http://ehstoday.com/safety/news/9-reasons-near-miss-reporting/.

within the enterprise, personal image and sanctions), (2) chronically poor understanding of what has to be reported (a representation of what the managers or the enterprise expects… which filters out many problems because they are judged not to be relevant for these purposes because no-one is affected, they are not severe enough, they have been recovered, they are too common etc.) and (3) ineffective use of the results obtained.

Some solutions have been found to the problem of protection for employees who make reports. A number of legislative frameworks (no blame, no shame) protect staff who make reports in certain countries (particularly the United States and Denmark[6,7]) in many industrial and service sectors. For example, for 25 years the national aviation safety reporting system in the United States (ASRS—Aviation Safety Reporting System) has protected the aviation professionals who report their errors by guaranteeing them anonymity and rendering legal prosecution impossible.[8] Similar systems now exist in the medical domain.

Most of the difficulties associated with the management of voluntary reporting from a safety perspective are different in nature and are still current.

Reporting risk is still very much bound up with the concept of the safety culture and to a lesser extent with the concept of improving results in the area of safety

A system has to agree to be transparent in relation to errors as a prerequisite for becoming safe. Reason [11] mentions four essential features when constructing an effective safety culture, all of which are more or less linked to error reporting: the ability to refrain from punishing those who make reports except in cases of intentional violations with serious consequences (just culture), the ability to share these events that are reported (informed culture), the ability to draw lessons from these reports (learning culture) and the ability to change the organisational model whenever reporting shows the ineffectiveness of the current model (flexible culture). These ideas are now well established and have been dealt with by many other authors[12, 13].

Having said this, reporting and the measurement of the safety culture (which is often firmly centred on this reporting aspect) do raise a fundamental problem in relation to the real link between the amount of reporting and the benefits in terms of the level of safety.

This link is obvious in aviation and in the nuclear industry but more open to debate in other industries.

In fact there is a bias towards specific cases resulting from the model used in civil aviation and in the nuclear industry, characterised by its powerful global, regional and national supervisory bodies (ICAO, EASA, National Aviation Authorities,

[6] Danish act on patient safety, http://www.patientsikkerhed.dk/admin/media/pdf/133907d0940e4d5f751852ec8f6b1795.pdf.

[7] US patient safety and quality improvement act, 2005, http://www.ahrq.gov/qual/psoact.htm.

[8] http://asrs.arc.nasa.gov/overview/immunity.html#, accessed on 26 December 2011.

Nuclear Energy Agency, NISA, IAEA etc.), the reality of total, permanent external surveillance (air traffic control and black boxes). In brief, these are relatively specific systems in which reporting incidents does not leave much of a margin for professionals since they will be seen and read by the supervisors in any case if they cause the slightest consequences. In this sense, the model of civil aviation or the nuclear industry is actually an exemplary model of voluntary reporting. In fact the density of reporting is correlated with the safety of individual airlines in civil aviation, since it expresses a function which is absolutely essential in that environment.

How many other industrial models, however, are similar to that model? Almost none… There is less supervision, actors have greater autonomy; it is not surprising that transverse studies on models of the safety culture whose key aspects are taken from aviation, the nuclear industry and the chemical industry, do not always yield such convincing results in other industries, particularly in medicine.

Reporting and the safety culture: what is the link between the reporting culture and safety performance within the industry? [14–16]

At the organisational level, there are nine dimensions that are repeatedly tested in questionnaires on the safety culture: the first is the policy on risk management and incident reporting, followed by the quality of the technical platform, the quality of maintenance, procedures, the quality and quantity of staff and planning, skills, collective commitment, communication and monitoring change. The use of questionnaires reveals a degree of aggregation and overlapping between the values of these nine dimensions, with the management largely predicting all the other values.

In the end, a number of questions remain unresolved, in particular the formal link between the measurement of a specific type of organisation and the risk of accidents. Values are measured if they are likely to be significant in terms of safety, as in the case of incident reporting, which is largely passed up to the highest level within the organisation, so that there is little independence of scale between the macro, meso and micro levels. One might ask whether, in view of the limitations of these questionnaires, audit techniques would perhaps be more appropriate.

The willingness to submit reports is even more variable in medicine than in the rest of industry [17–22]

All the literature indicates that systems that rely on reporting by health care professionals are subject to massive under-reporting. This is nothing new. In 1995, a study carried out over a 6 month period in a Harvard hospital showed that the reporting rate represented barely 6 % of the actual SAI (Serious Adverse Incident) rate as estimated through retrospective file analysis [18]. In 1998 similar results were found at the Brigham & Women Hospital in Boston: a system that automatically detected SAIs on the basis of electronic patient records had detected 2,620 alerts [19]; after verification, 365 SAIs were identified. On retrospective analysis of the records (which was carried out independently of the previous process) it was possible to identify 385 SAIs while health care professionals had declared 23 SAIs during the same period. Of the 617 separate SAIs detected using at least

one of the three methods, 65 % were identified by retrospective analysis, 45 % from electronic records and only 4 % from reports in the official reporting system. The literature consistently reports similar figures (between 3.5 and 10 %), which testifies to the very poor performance of such spontaneous reporting systems. Another result that is often seen [17, 23] is the massive over-representation of reports made by nurses (70–80 % of the databases) as compared to those made by doctors. Among this group, senior doctors were almost entirely missing from the database of staff making reports [21]. The most recent results [22] confirm these difficulties.

Paradoxically, the "no blame no shame" condition and questions about voluntary reporting may become less relevant to the introduction of risk mapping in future

As we have just seen, the question of the lack of legal protection for staff making reports, staff hierarchies and supervisory authorities has been an issue that has been obsessively addressed by the literature.[9] At present, reports from the actors involved represented virtually the only source of information.

Things should be different in future: incident reports contributed by actors on the ground will be marginal in comparison with other means of finding out about deviations and incidents. Due to computer technology and continuous systems supervision (black boxes), reporting has begun to be superseded by automated procedures. In this case, the initial difficulty involves "extracting the failure automatically from a stream of data". The real difficulty once this has been put in place will be knowing what to do with the no doubt impressive number of deviations catalogued by the automatic tracking system (out of all proportion to the number that are voluntarily reported by the actors today).

Automation of incident detection. A study carried out in 2010 [25] compares the results achieved using three methods used to catalogue serious adverse incidents (SAIs) in medicine in the United States: (1) a system of voluntary national reporting for medical professionals (the AHRQ or Agency for Healthcare Research and Quality system); (2) a system based on compulsory reporting by professionals of all incidents relating to a list of 20 national Patient Safety indicators (PSIs); and (3) an automated method for analysing the contents of all electronic medical records of patients admitted to hospital (global trigger tools method). The three methods were used for the same cohort of 795 patients from three general hospitals in 2004; the automated medical records monitoring system revealed 10 times more SAIs than the two other methods. A total of 393 SAIs were detected and 355 of these were only detected using the automated method.

The aviation industry has pioneered this process with its regulatory provisions for systematic analysis of on-board flight recorders (the tape for each flight is read and all abnormal values outside a normality envelope are subjected to additional manual analysis[10]). The analysis of medical records using the trigger tools method has been

[9] A good summary of this debate is found in Dekker [24].

[10] http://www.iata.org/ps/intelligence_statistics/pages/fda.aspx.

inspired by almost the same values [26, 27]: automated searches for abnormal values which trigger a subsequent manual analysis to understand the event.

The benefit offered by burdensome mapping methods is obvious for major industries but quite limited in innovative industries

The benefit offered by burdensome mapping methods (as compared with simple methods such as meetings between experts) is real, but ultimately it is quite fragile in terms of the time devoted to these formalities, particularly in highly innovative industrial systems.

To keep people's minds focused, a simple exchange of experiences over a few days involving a panel of carefully selected professionals who are carefully guided and actually work within the sector (guided brainstorming) will identify around 50 to 60 % of the total risk within a particular area; a feedback analysis (incident reporting) which is not carried out in depth and does not penetrate below the surface of anecdotes and immediate causes will yield almost nothing further (the stories that are reported are often tautological and confirm risks that are already known). On the other hand, an in-depth analysis of the same incidents yields 10 % more (i.e. 60–70 % of the real risk, once added to the initial brainstorming process) by identifying the systematic vulnerabilities (latent factors), but this takes a few days longer and experts once again have to be involved. Finally, formal methods (process analysis, functional analysis, FMECA, PRA etc.) add 15–25 % additional knowledge about the risk, but at the cost of a large investment of time (usually weeks or months). If all these steps have been carried out perfectly, which is rare and is only done in a small number of ultra-safe industries because of the time and resources required, the mapping results from combining the different methods may cover up to 90 to 95 % of the real risk. This is a very good result on paper but it has to be adjusted downwards to account for the natural obsolescence of the picture that results from it, which loses between 2 % (nuclear industry) and 20 % (medicine or software industries) of its relevance per year, depending on the pace of innovation and restructuring in the economic market for the system in question; due to the cost of the formal mapping process, it is very rarely conducted again at the same frequency…

Controlling safety in a context of high innovation: the case of medicine, with a knowledge turnover of 5.5 years.

Sjohania and his colleagues [28] analysed 100 reviews of questions published between 1995 and 2005 on recommended treatment strategies in multiple medical specialties, limiting themselves to the best randomised or semi-randomised controlled trials. They used two assessment criteria: quantitative, defined as whether or not a change in the clinical result by more than 50 % occurred in relation to at least one criterion as compared with the initial review, and qualitative, which considered efficacy, the identification of

new complications or new gaps in knowledge that were not known at the time of the initial review. They found one of the two signals in 57 % of the reviews that were published. The average value for the half-life of knowledge before an obsolescence signal appears is 5.5 years. In 7 % of cases, the review already had an obsolescence signal at the time of its publication and 11 % of reviews had one within less than 2 years. Medicine is one of the few professional human activities with such a high rate of innovation (it is only really exceeded by the software industry). This frenetic pace of innovation is in contrast to the quality model that medicine has chosen to import, modelled on ultra-safe systems (particularly aviation), with a total time taken for deployment of the method for an innovative item extending to 10 years on average: two years to identify the problem, two years to come up with local solutions, one year to address it in a centralised way at the supervisory level, two years to come up with a consensual solution (a national recommendation) and two to three years to train all operators nationwide in the use of the solution. This is as good as saying that the quality cycle is never completed in medicine because of innovation. It will be realised that there is a need to take a special approach rather than using the methods that are known from industry, since innovation, even more than quality, is a force for progress in safety that no-one would wish to arrest—for example the shift to day surgery using minimally invasive surgical techniques and natural routes of access results in a fall in the number of hospital acquired infections (e.g. the abdominal wall is not breached, cross-infection is reduced to a minimum), while the methods proposed by quality approaches have had difficulty maintaining this (very high) rate for decades. This makes it clear that no health care professional will choose to reject this innovation and continue to rely on the quality effect, even if its introduction gives rise to other problems. What will be the use of maps created five years ago to address the risks of traditional surgery if we know that during the next five years a massive transition will take place to a different type of surgery in all Western countries?

The process of building defences after the results of mapping have been obtained is still a strategic node which is not easy to resolve

The last point (what safety strategy should be adopted, on the basis of what is known about risk mapping) is clearly even more strategic when it comes to action planning. That is the final point in this initial phase: specific risks are selected which it has been decided to protect, and barriers (defensive mechanisms) are then built against those risks. We should recall that there are three complementary types of barriers available [29] (prevention, recovery and attenuation barriers), and each of these makes use of a combination of intangible tools (training, laws and recommendations) and physical tools (failsafe systems, locking systems, access blocks etc.). We will now take a few moments to consider these points.

Selecting the risks being protected against: the relevant decision-making matrices

The mapping process provides a list of risks but no priorities for addressing them; a decision-making strategy is therefore needed which accepts certain risks and protects itself against others.

When it comes to the risk of accidents, the solution generally comes from using a frequency*consequence decision-making matrix.[11] The results of the mapping process are arranged in boxes along the frequency axis (from very frequent to exceptional) and along the consequence axis (from minor to disastrous). The decision-making process involves accepting certain risks (which occur very frequently but have no consequences, or which are extremely exceptional even though their consequences are disastrous) and protecting oneself against all the other risks. It is possible to protect oneself either by reviewing the permitted design or working conditions (prevention), or by increasing the capacity to recover from and attenuate risks (mostly through training).

This approach results in two risks: actually protecting oneself effectively from the risks that haec been identified as a priority and acknowledging the impasse situations that have been accepted. This last area is clearly the most difficult, and brings us back to the problem of dealing with the management of weak signals.

Weak signals are an attractive concept but one that often turns out to be illusory from a management perspective

The rationale described above leads to the building of defences against all those parts of the risk matrix that are judged to be acceptable to the system.

The weak point in this rationale is the risks that are excluded, particularly the weak signals that are often discussed in the literature and at conferences and which need to be listened to and taken into account more effectively [30, 32].[12]

This is because analysing weak signals means nothing other than analysing those parts of the current matrix that it has been decided not to analyse. What appears to be simple when expressed in this way actually turns out to be very complicated, for a number of reasons:

- the question of choosing what to include. Unlike the part of the matrix that one is dealing with, which corresponds to a finite number of elements, the part that is not dealt with or which is set aside, consists of an infinite number of elements (since it represents all the members of an infinite set minus those that are being addressed). It seems obvious that all our resources would not be enough to deal with an infinite set of potential risks; it is therefore necessary to make choices on what to include… but how? In this context, the problem of choosing what to

[11] In the area of production line quality, decision-making methods tend to be used that give priority to frequency (the Pareto method is the best known of these).

[12] Ostberg [31] RISCResearch Paper No. 3, http://www.wisdom.at/Publikation/pdf/RiskBerichte/RRR_GOestberg_SomeIntangibles_09.pdf.

include comes back to the accident model that is developed to deal with these events. This choice in turn can be analysed into two other questions;

- the question of the accident model that is chosen. Weak signals cannot be dealt with using standard accident models, because it is precisely these (weak) signals which are (quite reasonably) set aside because they are not sufficiently severe or frequent. It is necessary to use models using percolation or coinciding conditions that find minor signals and events when they coexist in the same context; when these signals occur together, this constitutes a risk event. It goes without saying that the management and complexity of these percolation models has nothing in common with the simple models; these aspects take time, consume resources (the processing of weak signals that are rejected creates a considerable additional workload) and above all a high level of competency in the underlying field (the qualified individuals are usually university staff and are rarely employed in enterprises);
- the third limiting factor is the macroeconomic cost of extending monitoring to include weak signals. Thanks to a number of studies carried out in various places, the additional cost can be estimated to be 5 to 10 times the current cost of safety [33]. There are two types of added cost: of course the signals have to be included in the analysis, but there are also the indirect costs resulting from the protection strategies that would be developed to combat these low risks (which would cause some industrial initiatives and innovative risk-taking to grind to a halt) [34].

Although the concept of weak signals is very attractive in theory, it is easy to establish that it is quite unrealistic when it comes to the Cartesian management of everyday risks.

Fundamentally, weak signals have their greatest social utility through the actions of whistle-blowers. This existence of a counterbalancing power and a form of activism that asks questions about risks that are rejected or neglected, even if no in-depth analysis occurs of the Cartesian fundamentals in response to their revelations, at least maintains the feeling in society and among risk managers that not everything—indeed far from everything—is being controlled by the risk and accident models that are worked out on rational grounds.

Step 2: Comparing the Paper Model with the Reality on the Ground

Once the theoretical defence model has been built on the basis of the mapping process, it has to be put into practice and made sustainable: theories are called into question by real experiences and this raises discrepancies which have to be understood and either refuted (or accepted and changes made to the model) for the model to retain its relevance.

The most difficult problem to deal with is that of the progressive migration of the system and the automatic increase in violations as safety is improved [35].

Migration of practices is the norm in all systems [36]. Technical and economic conditions regularly impose new constraints on working situations: in many cases more (performance) has to be done with less (personnel, resources). Such degraded conditions, which initially only arise occasionally or during critical commercial periods, are not immediately punished by poor results or incidents (generally the opposite happens, as positive results are achieved thanks to the increase in performance); as a result the migration becomes standard and is accepted by everyone. This "illegal-normal" standard is accompanied by beneficial feedback (for workers) and this is perceived as corresponding to and compensating them for their efforts in terms of production: the hierarchy often gives them greater autonomy (specific initiatives are contested less, timetables and replacements are allowed to organise themselves, in short more is tolerated, bonuses are awarded for severe staff shortages and these deviations are gradually omitted from the feedback process). It would create problems if these migrations did generate signals, because they are providing a service both to the enterprise (performance) and to the workers (secondary benefits). The proportion of the enterprise that is operating within an "illegal-normal" range can easily rise to as high as 40–50 % of existing procedures in systems that are under economic pressure.

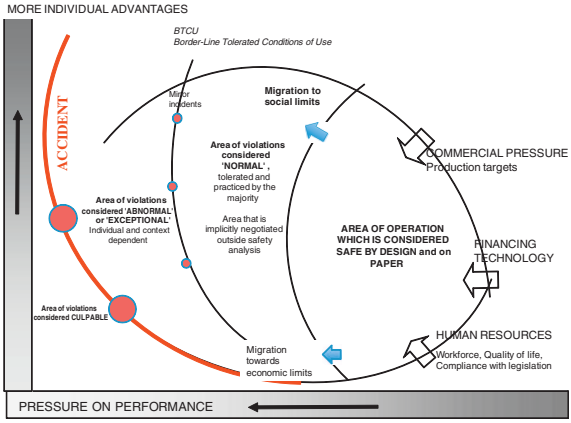
The size of this hidden area that does not conform to the prescribed model, is proportional to the appearance on paper of the margin that was built into create safety buffers: the violation or deviation are in fact only defined in relation to a regulatory requirement (whether this is inside or outside the enterprise itself). If the regulatory requirement contradicts an increasing demand for performance, the number of violations will automatically increase and the system will begin to migrate.

Two paradoxical points to remember about violations:

(a) violations are a characteristic of safe systems (which have procedures); they do not exist in activities that have no legislative framework and no rules;
(b) procedures that are poorly designed and too demanding automatically create violations.

A typical model of migration in practice (Amalberti et al. [36, 37], inspired by Rasmussen [38]): professional practices are controlled and limited from their conception by a combination of rules and formal and informal barriers to avoid the pressure of production causing rapid migration towards a high-risk area. These pressures on production, however, are so strong that the system will migrate, particularly if the system has been locked into very (excessively) prudent practices since it was first designed. Practices will migrate towards greater productivity and also towards greater benefits for operators. These "normal" violations, which are tolerated by everyone, may affect up to 50 % of procedures and in some cases will

continue to be amplified, creating real risks. It is the risk of secondary migrations that constitutes the priority target for interventions in safety.



How can migration be controlled? There are three solutions available to specialists:

- the first is a good idea which is nevertheless false, in relation to training: it involves front-running migration through unlimited increases in the skills of professionals so that they are able to deal with the exceptional peaks in production demand but return to conformity mode once the peaks have passed. As we saw in the last chapter, this solution contradicts the real situation: the more operators are technically trained and become even more expert, the more they will integrate their past successes, become confident and continue to migrate "one step further" even under normal conditions;
- the second is typically systemic and relates to design. We have just seen that more violations take place the more ideal and unrealistic the safety model imagined in the design on paper has become and the more it fails to take into account the economic constraints on production. At this stage we should remember that it is better to have a safety model which is designed to be less ambitious and compatible with the required performance rather than the other way round. We must also remember two other points: no safety model can absorb all the future changes that will occur in economic conditions from the time when it is designed. It is therefore quite normal for a model of practices to migrate in order to adapt to these new economic conditions, and it is therefore also normal to adapt the safety model regularly, sometimes relaxing its constraints and not necessarily systematically strengthening them. Worse still, since all violations and all migrations are a sign of dissociation between the requirements of the safety model and the performance model, the response to migration beginning must never be to reinforce the procedure or the constraint, since such a response will enlarge the difference between what is prescribed and what actually happens and will automatically increase the number of violations;

- the third is typically sociodynamic, and concerns monitoring of individuals who deviate more than others: migration of practices towards the illegal-normal opens up possible breaches in terms of social tolerance of non-compliance with the procedure [39]. If nothing is done, certain individuals who are more prone than others to take an autonomous stance will quickly extend these local permissions to all their practices. These individuals can never be controlled by orders from the hierarchy, at least not if they have had no accidents; worse still, they are often brilliant individuals whose accumulated successes and performance records are often presented as examples and who hold an envied status within the group. The only way to contain these individuals is to put them under the "authority of the group", an old historic solution inspired by the functional protection that groups and families provide to the most fragile individuals, and which is often found today in modern ethno-psychiatry. Setting up a reciprocal, consensual surveillance system between the members of the group, with the involvement of local managers and under their supervision, has proven to be highly effective in reducing extreme forms of individual deviation (but of course not those forms that have become a consensus for the whole group). It is specifically these techniques that underlie the success of "sellers of regulatory conformity for safety at work", who promote the techniques of BBS (Behaviour-Based Safety) such as the DuPont method. These methods are used in non-typical ways to reduce deviations from rules on wearing safety equipment in industry [40]. Nevertheless, making these techniques work requires a powerful team dynamic (all deviations have to be mutually reported, this has to be done politely, time must be taken to understand why etc.); this means more time and more personal involvement, as well as a renunciation of the cult of success and hero-worship. It is not surprising that these techniques are rarely introduced, except in the limited case of wearing safety equipment in factories, and if they are introduced they also turn out to be very difficult to maintain over time.

Step 3: Widening the Angle of Attack, Addressing Safety Through Macro-Scale Organisation

Making the operator's immediate environment safe is not enough to make a whole system safe. The first two steps are almost exclusively located at the micro level (the workplace and the area nearby) and at the meso level (the nearby enterprise).

A system consists of much more than this.

The general economic conditions governing trades, professions and regulators are very important when it comes to determining the existing opportunities to make the system safe [41].

A difference in safety by a factor of 100 to 1000 between the safest and least safe systems

The differences between safety levels in industry rise to a factor of 100 in terms of safety at work and a factor of 1,000 in terms of process and plant safety. For

example, in the area of safety at work, 1 in 1,000 professional fishing skippers die at work every year, as compared with 1.5 deaths per year for every 10,000 workers in the construction industry, and less than 5 deaths per year for every 100,000 workers in all other sectors [42]. The safety figures for industrial installations and processes are even more encouraging. One in 1,000 patients dies from an unexpected complication associated with an unwanted effect of his treatment in hospital, as compared to a risk of 1 premature death (associated with risk exposure) per million residents living for 5 years in the immediate neighbourhood of a nuclear power station.[13] or one premature death per million passengers who take an airline flight.

The effect of these differences is that there is little capacity to make relative changes in the level of safety between the major classes of human activity. For a change to be visible on this scale, it would be necessary for medicine, for example, to improve safety by a factor of 10 (which is a real challenge, since it has achieved virtually no progress in 10 years), and even such a change would still be negligible in comparison with systems such as civil aviation, which are still 1,000 times safer.

The overriding importance of macro-scale challenges in relation to local actions to improve safety

The systemic factors and central management of the enterprise often impose major limitations in terms of safety initiatives that can be pursued at the local level [38].

As a minimum, these may involve a willingness to coordinate all policies centrally while putting a brake on isolated local initiatives; more frequently still, the management are aiming to preserve the contradictory tensions between the real need for safety, to maintain their licenses by responding in accordance with administrative requirements imposed by regulators, while protecting the desire for exposure to acceptable risks in order to strengthen the image, maximise production, support the economic model or promote the personal success of managers.

Safety specialists have to learn to deal with this paradoxical framework and understand the trade-off mechanisms that are used in their industry.

The nuclear industry, for example, has a very low social tolerance of voluntary exposure to risk. The same is true in civil aviation, so here the trade-offs tend to be made in favour of safety initiatives, while targeting full and centralised coherence (and working to prevent isolated local actions). The same has not been true (at least until recently) in international finance, medicine, fishing and motoring, to cite just a few examples; in these cases, senior management—or the institution in the broad sense—gives priority to exposure to risk and safety activities are mostly conducted at the local level, and it is accepted that they are local and limited in scope.

It is understood that a safety specialist may move from one system to another during his career, using the same tools and the same knowledge and yet obtaining very different results. The context in terms of risk acceptance serves to buttress the level of safety, while local actions are only flimsy supports, whose leverage has a

[13] Wilson [43], or http://mullerlbl.gov/teaching/physics10/old%20physics%2010/physics%2010%20notes/Risk.html.

very limited duration and geographical scope and which are less likely to be rolled out by the management the more they are perceived as incompatible or detrimental to the overall economics of the system.

We will come back to these points later on when we analyse the work that has been done on types of safety and margins of action, and also on ways of building trade-offs and compromises.

Five barriers that account for the differences between less safe and ultra-safe systems [44].

The first barrier is the lack of any limitation on performance imposed by the authorities regulating the system. Fishing skippers at sea, for example, are able to remain at sea even in force 10 conditions, professional mountaineers may decide to embark on an attempt on the summit regardless of the risk conditions, and the on-call doctor in the emergency team has to see every patient who comes in, with no hope of relief even if he is exhausted. In the absence of regulations that act as a brake on exposure to risk, these professionals cannot hope to have a safety level higher than 10-3. The second barrier involves the autonomy of the actors themselves. If the only factor that matters to you is the realisation of your personal goal, and you do not take into account the consequences for those who are following after you, there is no chance of that you will reach a safety level that is higher than 10-4 for your activity. A surgeon who is running 4 h late with his list will be able to complete his list without making a mistake, but the risk will result from a patient who has had an operation returning to the ward at midnight or 1 am, in conditions where there are fewer nursing staff and at a time that creates greater risks in terms of post-operative monitoring. A safe approach would require him to take into account the limitations on the ward and no doubt to cancel his surgical list once it had become delayed in this way. The third barrier "or the artisan's barrier" summarises the effects above in social terms. One does not ask the name of the airline pilot or the biologist, but one chooses a jeweller or a surgeon; they are artisans; they market their differences and their personal knowledge. There is no skilled trade in the world, however, with a level of safety higher than 10-4. The reason for this is simple: safety is based on reducing differences and maintaining a stable service 24 h a day, while the artisan markets his services precisely on the basis of the difference between himself and his neighbour and competitor: the instability of the system as a whole is therefore structurally inherent in his work. The fourth barrier, which is referred to as "overprotection", occurs in systems that are already safe. The safer the system becomes, the more paradoxically exposed it is to enquiries and to blaming of those responsible whenever an accident does happen; the response is often to create rapid growth in protective procedures that are intended to provide cover for management. These protective procedures place unnecessary restrictions on work, make it more burdensome and promote deviation and migration of working practices among front line operators,

while paradoxically increasing the risks that they are intended to protect against. Finally, the fifth barrier relates to the loss of rationality in communication about safety in systems that have become very safe. Fewer accidents occur, but they have a higher media profile. One important component of communication by managers and the actions that are taken therefore involves reassuring the media. Statements and assurances are often issued precipitately (there is a tendency to read the surface-level statistics and boast about progress without taking a step back) along the lines that the press wishes to hear (and not necessarily along the lines of the safety model). Together with this approach of making decisions at multiple levels with no assessment of their objective results, the safety system becomes a huge, multi-layered system of procedures and requirements, where nobody knows which ones actually contribute towards achieving the level of safety; as a result the system becomes incapable of self-cleaning, becomes excessively complex and creates an asymptotic trend in safety [45].

Step 4: Once All the Steps Have Been Taken, is it Still Possible to Withstand Exceptional Events?

Managed safety and controlled safety

The safety of complex systems is the sum of two aspects: on the one hand the safety achieved through all the prohibitions, limitations, legal requirements (referred to as controlled safety) and on the other, the safety supported by the adaptive intelligence of the operators and professionals within the system (managed safety). These two concepts, which are now widely discussed in the literature, were introduced for the first time in 2008 in the leading article published in Human Factors on the work done on professional fishing skippers in the thesis by Morel et al. [6].

$$\textbf{Total safety} = \textbf{controlled safety} + \textbf{managed safety}$$

Let us consider the terms of this equation so that we can apply it to different fields of professional work.

Skilled trades are subject to few regulations; their whole of their rather modest safety structure is based mostly on the quality and skills of the operators themselves, with the high level of variation that is inherent in individual quality. On the other hand, for these experts adaptation to exceptional conditions is their everyday work and they are remarkably good at it (at least this is true of the best of them, who survive both economically and physically).

$$\textbf{Total safety (skilled trade systems)} = \textbf{controlled safety} + \textbf{managed safety}$$

(*the font size metaphorically shows the strength of each term in the equation*)

Very safe systems, on the other hand, have large numbers of procedures and prohibitions. The level of safety is high, but the adaptive expertise of operators in these areas is automatically reduced because they are no longer exposed to exceptional situations and they are no longer trained to work outside their procedural framework.

Total safety (ultra-safe systems) = controlled safety + managed safety

There is currently no known solution that can preserve both the expertise of the operators in exceptional situations and the benefit of achieving maximum system safety by procedural means.

It should be remembered that resilience, or the art of adapting to exceptional conditions, is a native feature of human systems that rely on their own autonomy to survive; it automatically disappears as a result of using the traditional tools that enhance safety in industry and service sectors. In many cases it would be preferable to have less of this in less safe systems (because it is associated with frequent improvisation and habitual non-compliance) and also to reintroduce it in ultra-safe systems (because it would permit adaptation to exceptional situations, an ability which by this stage has largely been lost).

The safety choice in the aviation industry: suppressing heroes and prohibiting training in situations that are too exceptional.
In 1995 two serious accidents and incidents occurred involving Airbus A310 aircraft which suggested that pilots do not have enough training in difficult manoeuvres. The first was an accident which occurred on takeoff from Bucharest, where the crew were distracted or busy and the aircraft was allowed to go into overbank and was unable to recover from this unfamiliar situation; a few months later, when coming into land in Paris, the crew of another Airbus A310 were surprised to find the aircraft in an unexpected attitude and it was only with great difficulty and considerable good luck that they recovered the aircraft without damage. The two inquiry commissions emphasised the lack of training given to these pilots who fly modern aircraft but who are clearly better trained to manage the computer and use the automatic pilot than to fly manually themselves. The response from the civil aviation authorities was uncompromising: there was no question of resuming training in manual flying, which only occurs in exceptional circumstances. This would re-open the door to the “heroes” who depart from standard procedures too frequently; the aviation industry has worked long and hard to eliminate them. The chosen solution will be to aim to use warning signals to provide better indications that the aircraft is moving outside its normal flight path, and to rely more on automatic systems and on the safety net to recover from these unusual situations automatically. This attitude has been seen in relation to every single unusual accident.

This is an exaggerated example of a system (civil aviation) which relies completely on procedures and supervision and which benefits from these every day in terms of safety results (one of the best safety standards in the world), but is now locked into that approach. Any attempt to reintroduce improvisation in areas where there are no procedures is initially seen as putting the system at risk and is consequently not allowed.

It should also be noted that the very well-known case of the successful landing on the Hudson River by the US Airways flight in 2009 is simply viewed as a stroke of luck in the world of aviation regulation and is not leading to any questioning of the model that has been described above (there has been no collective and professional learning experience from this accident in the aviation sector, even though it has supported a number of research groups who wish to re-use it for their own purposes).

Institutional resilience: surviving an accident can prove to be as important as avoiding the accident

Safety specialists have a simple and unique model in their minds: reduce the number of accidents and incidents. This means forgetting a whole aspect of safety which is typically systematic: knowing how to survive accidents.

This approach may appear to be cynical, but it should be integrated as a fully-fledged component of the modern approach to safety in an enterprise or complex system.

If this is done, it adds a critical aspect to the two well-known themes of crisis management (knowing how to respond to the accident and knowing how to manage communication) by emphasising "the ability to continue production after the accident". When considered more closely, this third aspect may even be the most important of all. It encourages the continuation of the crisis cell long after the accident, to manage "long-term damage to the image".

The accident is not just an isolated incident in the life of a system. It may be (and often is) repeated at regular intervals. Every accident is like a heart attack or a relapsing cancer for a human being. It calls into the question the survival of the whole system.

A model of institutional resilience: the example of Air France.
The airline Air France has suffered five major accidents in 15 years (747 in Papeete in 1993, 747 cargo plane in Madras in 1999, Concorde in Paris in 2000, A340 in Toronto in 2005, and A330 between Rio and Paris in 2009) and a series of major incidents during the same period. This makes Air France the riskiest major airline in the world. What is more, unlike its competitors which have had fewer accidents (in some cases only one) leading (in almost all cases) to their economic demise (examples include TWA

and Swiss Air), Air France managed to save its image from being destroyed after these accidents and in some cases even gained market share! This is an example of resilience on the part of the enterprise. One example: on the night in 2005 when its A340 crashed at Toronto airport as a result of poor judgement by its crew, the Air France management successfully created a positive reading of its company image, both immediately and in the longer term, thanks to a "positive" angle on the accident which was broadcast by all press outlets and international news agencies, since it demonstrated the airline's ability to evacuate an aircraft in difficulty. This was the abiding image in the public mind. The ability to do this owes nothing to chance. It is prepared in advance, is managed directly and forms part of a highly rehearsed post-accident plan. This ability is part of the company's know-how.

Three Models of Balanced Safety Rather Than Just One

The idea of a single model of safety that applies in every context and aims to have zero accident is naïve. Safety is a social construct and it adapts to demand. As this section shows, there are many different responses to safety, which describe a number of different models of safety (resilience, HRO, ultrasafe), each with their own approach, advantages and limitations. These models take different approaches to the trade-off between the benefits of adaptability and the benefits of the level of safety.

Three Very Different Strategies in Terms of Exposure to Risk

Everyone will agree that writing a safety plan offers no guarantee that the plan will be put into practice.

Everyone will also agree that it is rare for a safety plan to mention what is not going to be done because of the trade-offs that have been agreed.

These two aspects (doing well what one has decided to do, and knowing what one has decided not to do) are strategic in terms of managing risk on the ground.

The compulsory pencil and paper risk mapping exercises and the risk prevention protocols developed for regulators are no doubt essential, but these are often only one-time efforts (for demonstration purposes). They need to be rooted in everyday reality over a long period of time… which presents a completely different problem.

In the absence of concerted action and trade-offs at a higher level, front-line management and front-line staff will be the level at which all the (contradictory) constraints issued by the various divisions are integrated: produce more, under conditions requiring quite high levels of situational adaptation and in accordance

with the safety instructions and rules. There is every reason to fear erroneous local interpretations and new vulnerabilities.

Three families of safety plans or solutions are always available to sociotechnical systems for the management of everyday risks.

- **PLAN A. The first family of safety solutions involves eliminating or delaying exposures to risk. We will call this plan A**: the aviation industry excels at this strategy. Thanks to its global coverage and its absolute authority, air traffic control is able to prevent situations where aircraft are exposed to difficult conditions. The same applies to the nuclear industry, which has very robust incident procedures, all of which tend to ensure the immediate safety of the installation and shut it down temporarily. In return, this level of supervision promotes economies of scale in training: there is no point training pilots to fly aircraft in hurricane conditions if one knows it is possible to avoid all hurricanes. This prudent strategy, however, also requires a high standard of systematic supervision which is often outside the scope of fragmented, deregulated and/or highly competitive industrial activities and therefore also of skilled trade systems.

The ability to implement plan A depends on the way the system is organised. Let us take the example of a comparison between two health care systems in France and the United Kingdom, where a hip replacement is carried out for a patient with comorbidities (diseases other than the hip problem, such as diabetes or hypertension). In these cases the replacement operation is never an emergency procedure and it is better to wait for conditions in which the comorbidities are perfectly under control before the operation takes place, to avoid postoperative complications. This strategy of waiting for favourable conditions works quite well in the United Kingdom, because the system of access to health care is highly controlled; on the other hand it is quite ineffective in France, because there is too much private provision and too little regulation; patients can consult as many surgeons as they wish and these consultations are reimbursed, allowing them to get an operation date very close to what is convenient for them. It is not surprising that in this context French surgeons take more risks than their English equivalents to avoid the patient leaving, and use plan A less often.

- **PLAN B. The second family of safety solutions involves accepting exposure to risk while complying with all the recommended standards and procedures (we will call this plan B)** but while maintaining the ability to detect changes in the context and maintaining local, intelligent adaptation of these procedures (the procedure at the heart of the system). Strict and standardised implementation of all professional recommendations under standard working conditions will minimise the number of accidents. These plan B and plan A approaches generally feed into the responses that are made to regulators.

There is no such thing as a working environment without incidents and above all without surprises. In order to gain the maximum benefit from this approach, it is therefore necessary either to be able to stop the system and make it safe quickly (no go), or to have unambiguous procedures for dealing with incidents (specific

procedures to deal with each catalogued type of abnormal event). In this case the operator does have to be aware of how the situation is evolving and must be able to identify the problems and apply these procedures.

It should be noted that poor system ergonomics can easily compromise this aim. Dave Woods, setting out the pioneering work by Bainbridge [46], has extensively defined the risk of "surprises" in standardised situations in ultra-safe environments where changes do not occur often; such surprises are mostly associated with designs and supervisory systems that do not incorporate the ethical rules to ensure the stability of the tool being used (see the example below). From this, Woods and Hollnagel [47] deducted a number of principles of ergonomics that were to characterise the design of safe systems; in particular they reiterated, and popularised under the term Joint Cognitive System the old idea that had been put forward by French ergonomists since the 1960s[14]: "ergonomic analysis is on the wrong track when it breaks down the work in a Cartesian way and analyses it in separate parts (work analysis)"; the activity of the operator is embodied in the technical context, and can only be studied in the form of a dynamic linkage. The contribution by Mica Endsley should also be noted [48, 49], which follows the same track and proposes testing the adaptation of the understanding of the world within which the operator develops, using his concept of *Situation Awareness*.

One example of an unacceptable surprise associated with a design [50]. In the late 1980s, the first Boeing 737 s had an ALT HOLD function which made it possible to stabilise the altitude of the aircraft in the vertical plane (stop it from climbing or descending) by simply pressing a specific button; pressing this button had no effect during the final phase before landing, however, because the aircraft builder wanted to safeguard against accidental pressing of the ALT HOLD button, which could have seriously disturbed the automatic landing process under conditions of poor visibility (CATegory 3 procedure). It should be noted that this protection system had been installed to meet the safety requirements imposed by the regulators to obtain the CAT 3 qualification. The result was telling: during this final phase, to achieve the same result (stabilise altitude) a complex sequence of actions had to be carried out: first disconnect the automatic pilot, then disconnect the two flight controllers at the controls on the right and on the left, then reconnect the automatic pilot, and finally press ALT HOLD. A number of accidents and severe incidents were caused by this ergonomic inconvenience before it was corrected: pilots, not realising that they were in these approach conditions in which direct use of ALT HOLD

[14] deMontmollin M. L'ergonomie de la tâche, Peter Lang, Berne, 1986.

was prohibited, pressed the ALT HOLD button and nothing happened, and this was followed by a moment of surprise and inappropriate action, which caused incidents/accidents.

- **Plan C. The third family of safety solutions involves tolerating exposure to non-standard conditions, while accepting that operators do improvise or deviate from procedural behaviour. We call this plan C.** In many professions, life does not consist of procedural repetition; quite the opposite. This capacity for adaptation has led to a great deal of debate in the small community within the humanities in which the idea of resilience is studied.

The Metaphor of the climber and the rock face. One can consider **hazards** as rock faces. They are an inevitable part of nature. In industry, such rock faces may represent sick patients, the chemical properties of compounds, solar radiation etc. Risks depend on the willingness to deal with these rock faces and the way in which this is done. One can refuse to climb them (plan A), one can limit oneself to climbing only known rock faces and follow all the required procedures (plan B), or one can attempt rock faces in non-standard situations (without equipment, without training, under poor or changing conditions), or worse still, climb unknown rock faces (plan C). The more stable and supervised it is, the more it relies on avoidance, plans A and B, and the less stable it is, the more it will have to rely on its adaptability to deal with changing conditions (plan C).

In short, these solutions cannot be transferred from one environment to another; they have different aims, use distinct models of safety, require different types of experience and training and follow different organisational approaches.

Outside a small number of ultra-safe industries, the majority of human professional activities rely heavily on plan C. Strangely enough, however, all the literature on the quality and safety of systems offers prescriptions only for plans A and B.

It is not because those relying on plan C do not follow all the procedures and result in improvisation that it is not possible to make their practices safe. The problem is that the solutions that would make these practices safe while accepting their reality do not consist in developing procedures. (If they did, one would change to a plan B approach. Instead, the response is ad-hoc and does not cover all the situations that arise during the work, whose very economic rationale often demands that it rely on plan C.) Plan C solutions are found in quite resilient models: becoming more expert, becoming able to judge the difficulty of the task according to one's own skills, learning to learn, drawing from experience, acquiring generic knowledge schemas which allow adaptation to borderline circumstances.

Three generic plans to manage risk:

- PLAN A: refuse to do it or wait for ideal conditions;
- PLAN B: do the work under ideal conditions according to recommended procedures;
- PLAN C: accept that you have to take action without having ideal conditions, including improvisation and working outside procedures.

In aviation, the ratio is 40 % plan A, 55 % plan B and 2 % to 5 % plan C. In medicine, the ratio is 5 % to 10 % plan A, 40 % plan B and 55 % plan C. What is the ratio in your own work between plan A, plan B and plan C? If your plan C ratio is higher than 5 %, what is the value of the procedures that you have in place to cover plan B in the cases when you are working according to plan C?

Two professional contexts and two diametrically opposed safety strategies = helping to survive the risk versus protecting operators against exposure to risk.

Systems that have a relatively modest level of safety (lower than 10^{-4}) have considerable exposure to risk because they literally make a living from that exposure. This is true of fighter pilots, sea fishing skippers and professional mountaineers. In these professions, accepting exposure to risk and even seeking out risk forms the essence of their work. These professions do, however, still want to improve safety. A number of studies carried out among fighter pilots [51] and sea fishing skippers [8, 52] show a real need for safety. Fishing skippers, for example, would like to have an intelligent anti-collision system to offer them better protection in high seas with poor visibility and with the mobility required for trawling (Automatic Radar Plotting Aid). Fighter pilots would like an electronic safety net to offer them better protection when they are undertaking manoeuvres that are likely to make them lose consciousness (Electronic Safety Net). Moving on to the example of civil aviation, everyone also wishes to improve safety in this area. Here, however, the solution is radically different and most commonly involves not exposing crews to the surprising conditions or risks that are thought to be the cause of accidents. For example, the eruption of the Eyjafjallajökull volcano in Iceland in 2010 led to all aircraft immediately being grounded, based on a simple approach: no exposure to risk. These different examples highlight two completely opposite strategies to dealing with risk: one, which is supported by small-scale systems involving skilled trades or highly competitive activities, involves relying on the intelligence of operators and giving them aids to deal with risk; the other involves relying on the organisation and supervision and ensuring that operators are not exposed to risks. It is easy to understand that both of these models have their own approach, but in that case it is also necessary to accept that the safety solutions are not identical in both cases.

Three Authentic Models of Safety Rather Than Only One

Taking into account the risk exposure strategies already mentioned, it make sense to take the view that each one has given rise to an authentic way of organising safety which is original, with its own approach and its own possibilities for improvement [53].

- **The resilient model** involves professions in which seeking exposure to risk is inherent in the economic model of that profession. Skilled trade professions in particular sell their services on the basis of their expertise which allows them to deal with new risks, or even deal with the unknown, innovating, mastering new contexts, coping, winning through and reaping benefits where others fail or are afraid to go. This is the culture of champions, winners… and losers (the losers are part of the context, but they are not perceived as failures of the system but rather as a reflection of the knowledge and skill of the champions). Sea fishing skippers, for example, are capable of seeking out the riskiest conditions in order to prioritise catching the most profitable fish at the best times (sales economy); experts in oil exploration have to find oil almost at all costs once the procedure has been initiated; only success makes sense at that point. Traders constantly have to maximise their profits and military fighter pilots[15] always have to win… All these professions have objective accident statistics which are more or less disastrous. They are not, however, insensitive to their professional risks, and they deal with these through safety and training strategies which are very well thought-out, but of course within a different culture.

In these professions, the individuals' autonomy and expertise take precedence over the hierarchical organisation of the group. In many cases the group is very small (consisting of two to eight individuals) and works in a highly competitive setting. The boss is recognised for his technical ability, his past performance and

[15] The case of fighter pilots is a special and interesting case of a dual context: in peacetime, their administration (the Air Force) operates essentially on an ultra-safe model, but once the aircraft are deployed on active service, the operating model suddenly changes and returns to its fundamentals of resilience. These very contrasting contexts do generate surprises in terms of safety in both directions: persistence of resilient, deviant behaviour (as compared with the model that would be desired in peacetime) after returning from military campaigns, and important opportunities that are missed during the first few days of engagement due to lack of practice in the resilient model, when pilots are suddenly thrust from peacetime into operational theatres. A military air force can also shift the crew of an AWACS surveillance aircraft from peacetime into wartime during the space of a single mission: they may leave a French base in France, having dropped their children off at school in the morning… fly a 12 h mission that involves working in and overflying an operational theatre with a very high risk of aerial engagement and requiring particularly high resilience and return the following night to their air force base in France and also to their homes, which are completely organised around social routines and the challenges of peacetime.

his charisma more than for his official status. Every operator is constantly invited to use a very wide margin of initiative. A correct assessment of his own skill, courage and accumulated experience are the keys to recognition as "a good professional and a winner"; safety is mostly about winning, surviving, and only winners have a chance to communicate their safety expertise in the form of champions' stories. To summarise, there are a small number of procedures, a very high level of autonomy and a very large number of accidents. It is still possible to make progress in terms of local safety, however, by becoming better trained through contact with the best masters, learning from their experiences and adding to one's own mental capacity to adapt to even the most difficult situations. The differences between the least safe and the safest operators within a single resilient, skilled trade are of the order of a factor of ten,[16] which proves that it is possible to make progress through safety interventions, even while remaining within the "micro-Gaussian" distribution of professionals engaged in these hazardous types of work.

- **The HRO model** (*High Reliability Organizations*) uses the same idea of resilience, since it also promotes adaptation, but this is a kind of adaptation which is more local and controlled, involving human activities which are clearly better organised, with less of a tendency to seek out daring exploits (which is more characteristic of the pure resilient model). The HRO model is in fact relatively averse to individual exploits that are not controlled by the group.

HROs typically apply to professions in which risk management is a daily affair, even if the aim is still to keep risk under control and avoid unnecessary exposure to it: firefighters, merchant navy and naval armed forces, professionals in the operating theatre, oil exploration, those operating chemical factories.

HROs rely on the leader and the professional group, which incorporates several different roles and types of expertise in order to maintain a constant perspective on progress being made towards the goal (while avoiding the risks of a local focus), where all the members of the group play a part in detecting abnormalities in a contextual setting (sense making), bringing them to the attention of the group, adapting the procedure to these changes in the context. This includes deviations from procedures when necessary (but only when this makes sense within the group and is communicated to everyone). All members of the group show solidarity in terms of this safety objective.

Combating adversity is an integral part of the HRO approach but the high level of collective regulation (not necessarily only by the leader) imposes considerable limitations on isolated individual initiatives and promotes prudent collective decision-making.

The HRO model analyses its own failures and seeks to understand the reasons behind them. The lessons drawn from these accident analyses, however, are

[16] The rate of fatal accidents in professional deep-sea fishing varies by a factor of 4 between ship owners in France and by a factor of 9 at the global level, source: Morel et al. [52], op. cit.

primarily about ways in which the situation has been managed and could be managed better in future.

This is therefore a model which relies firstly on improving the barriers to detection and recovery, and secondly on barriers to prevention (which involve not accepting exposure to these difficult situations). Training is based on collective acquisition of experience. Once again, the differences between the best operators and those that are less good within a single trade are of the order of a factor of ten.[17]

- **The ultra-safe systems model** no longer makes it a priority to rely on the exceptional expertise of these front-line operators to escape from difficult situations; instead it requires operators to be identical and interchangeable within their respective roles, and in this case requires them to work at a standard level. The model, on the other hand, relies upon the quality of external supervision, making it possible to avoid situations where these operators are exposed to the most exceptional risks; by limiting the exposure of operators to a finite list of breakdowns and difficult situations, the model can become completely procedural, both when working under normal conditions and under abnormal conditions. Airlines, the nuclear industry, medical biology and radiotherapy are all excellent examples of this category. Accidents are analysed to find and eliminate the causes so that exposure to these risky conditions can be reduced or eliminated in the future. This model relies on prevention first. Training of front-line operators is focused on respect for their various roles, the way they work together to implement procedures and how they resopnd to abnormal situations in order to initiate ad-hoc procedures. Once again, the best and the least good operators within a single profession differ by about a factor of ten.[18]

Four lessons can be drawn from this:

- **the three models of safety are radically different**. They represent responses to different economic conditions, each one has its own approach to optimisation, its own approach to training, its own advantages and its own limitations. They can be plotted along a curve in which there is a trade-of between flexibility and adaptability on the one hand, and safety on the other. All three, however, have the same capacity for internal self-improvement, and safety can be improved by a factor of 10 (making them 10 times safer);
- **the three models cannot be mixed**. Mixing the features of one with those of another leads to a failure to improve safety and may even be counterproductive.

[17] The rate of fatal industrial accidents in the gas and oil extraction industry varies from 130 deaths per 100,000 workers in some African countries to 12 deaths per 100,000 workers for the best oil wells; the global average is 30.5 deaths per 100,000 workers, source: http://nextbigfuture.com/2011/03/oil-and-gas-extraction-accidents-and.html.

[18] The rate of aviation accidents ranges from 0.63 per million departures in Western countries to 7.41 per million departures in African countries. These therefore differ by a factor of 12, source: IATA statistics, 23 February 2011, http://www.iata.org/pressroom/pr/pages/2011-02-23-01.aspx.

For example, there is no certainty that reintroducing training in deviating from procedures and dealing with unknown situations in civil aviation will not erode safety rather than improving it (this is why global regulatory authorities have refused to go down this route). On the other hand, introducing restrictive procedures in combat aviation or deep-sea fishing would perhaps result in fewer deaths… but it would kill off the profession itself;

- **local interventions cannot change the model**. If intervention takes place locally to improve the safety of an enterprise or a particular working unit within a profession that belongs to a specific model (resilient, HRO or ultra-safe), there is no opportunity to encourage the adoption of characteristics of a different safety model (for example, if intervention takes place in fisheries, it would represent an illusion to advise them to adopt an ultra-safe system strategy). Instead, it is necessary to rely on the capacity for progress that is available within the model in that specific professional setting (the resilient model, in the case of fisheries) by using the strategies that are specific to that field and have been seen to offer significant opportunities to improve safety, since this can be improved by a factor of 10;
- **it is possible to switch from one model to another, but this requires a changeover event** that will affect the entire profession and its economy. The industrial chemical industry, for example, which in some cases is still based on resilient models dating from the 1960s and 1970s, made a definitive switch to an HRO model after the events that occurred in Seveso in Italy in 1976 and the European Directive that followed in 1982. It is often the regulatory mechanisms that impose such a transition to a new system. It will be noticed that in this case the system migrates gradually, loses the benefits of the previous model (a higher level of adaptation and inclusion of situations that are considered to be manageable within that profession), but gains the advantages of the new model (mainly in terms of safety).

The benefits of each of these models can be assessed on the basis of different beliefs and different approaches.

One might think that applying the resilient model to deep-sea fishing skippers does not represent a major problem, since their accidents have no major consequences outside their own profession. In the end, this is a choice that must be respected. On the other hand, the use of the resilient model in medicine raises complex ethical questions in relation to these two contradictory approaches: providing access, hope and care to everyone in all circumstances (which the resilient model does better than the ultra-safe model) and at the same time doing nothing that could harm or injure the patient (first do no harm) (which the ultra-safe model or even the HRO model do much better than the resilient model).

One may also consider that the nuclear model is not safe enough and call for even more standards and protocols to be introduced to cover situations that are considered to be more and more improbable. This is being done, for example,

after Fukushima; having tested all the power stations in the entire world for the risk resulting from voluntary crashing of a passenger aircraft after the terrorist attack on the Twin Towers on 11 September 2001, the same power stations now have to be tested for their seismic risk and their risks of flooding and dedicated reinforcement measures must be taken. To some extent it will be necessary to bring what had been thought impossible into the realm of the possible and apply the demands of ultra-safe systems to this new possible situation. This strategy, which is typical of ultra-safe systems, is the exact opposite of a resilient solution: it reinforces the idea that the system is able to defend itself properly against known risks, but by doing this one refuses to learn to improvise to deal with a new exceptional surprise which will certainly occur one day (tomorrow? in 5, 10 or 20 years?) in one of the 500 or so nuclear power stations that are in operation throughout the world.

But would it be realistic for the nuclear industry to adopt a different strategy? A truly resilient strategy? Imagine that following Fukushima, the global nuclear industry decided to train its operators to give them greater resilience and recommended training them to deal with unexpected situations,[19] including conditions that have never been seen. It would then no doubt be necessary to train teams of operators to improvise and depart from procedures. It would also be necessary to be coherent and accept that there must be a two-speed system, one ultra-procedural one similar to the one that exists today, which maintains the existing exceptionally high level of safety ($1*10^{-6}$), thanks to strict compliance with procedures (a safety style which applies to 10^6 working hours[20] i.e. 45 years) and the other based on training a few expert operators, who are present in each power station, duplicating the standard operators, who are capable of improvisation, who would be deployed once in every two generations, or three times a century… and who would have to be banned from accessing the controls during the 27 years when no exceptional surprises arise, to avoid dangerous, unnecessary improvisation. Once the terms of this equation are set out, the answer is already clear: it is not feasible.

Another current example: many people consider that the financial crisis involving sub-prime mortgages and European debt that has afflicted the world since 2009 had the same roots: an excessively resilient model, driven by a minority of actors and in which profits are maximised by taking unreasonable risks, while intentionally masking and complicating transactions and financial products to

[19] The unexpected being referred to here is not the surprise of a logged breakdown occurring for which a procedure exists. Of course breakdowns and problems are not reported in advance, but they form an integral part of the ultra-safe model, with operators who are trained to respond. We are talking here about situations that have never been encountered before and for which no written procedures exist. It would therefore be necessary to improvise.

[20] This safety level of 10^6 is the guaranteed level for risk analyses at the design stage for the aviation and nuclear industries.

thwart all the supervisory procedures. In short, many ordinary people now consider that this system should change its approach and adopt more circumscribed, procedural safety rules, and depending on the political convictions of those involved, this would result in it at least adopting the rules of an HRO system, and in some cases becoming a totally supervised ultra-safe type system in which the actors involved have no autonomy at all. The fundamental nature of this system, however, is precisely its capacity to generate money competitively (freely) in order to finance the market; in short, this requires conditions that run completely counter to those of a controlled system.

It is no surprise that the proposals to improve moral standards in the global market economy and the role of the banks, to regulate and circumscribe their activities and even to (re-)nationalise them in the most extreme cases, although it was regularly discussed at G7 and then G10 summits, never went beyond the discussion of good intentions. The reason for this is simple: the models that would have to be called into question are above all societal models, with values and beliefs built on the successes of the past. No-one is really ready to abandon these for the sake of a hypothetical improvement in safety (whose results would not be seen until the entire process of transformation had taken place, and which would give rise to difficulties experienced immediately).

It is not possible to impose a completely new model of safety against the will of local actors and contrary to values that are considered essential to this system.

These underlying values must be shifted first, before making any claim to make people adopt a different safety model.

The lesson from this is simple: changing the safety model means changing the system. If the conditions are not met, and sometimes it is necessary to accept this fact, it is no good tilting at windmills or inventing solutions that have no chance of success.

The properties of three models of safety: the resilience model, the HRO model and the ultra-safe system model.

	Resilience model	HRO model High Reliability Systems	Ultra-safe model
Examples	Himalayan mountaineering professional fishing combat aviation international finance hospital emergencies	Merchant navy, Air freight, Naval armed forces Fire service oil industry Operating theatre	Civil aviation, Trains and metro services, Nuclear industry, Medical biology, Radiotherapy
Rationale	Taking risks is the essence of the profession. People take risks in order to survive in the profession.	Risk is not sought out, but it is inherent in the profession.	Risk is excluded as far as possible.
Key cultural trait	Fighter spirit, cult of champions and heroes.	Cult of group intelligence and adaptation to changing situations.	Cult of applying procedures and safety organised by an effective supervisory organisation.
Characteristics of accidents and lessons learned in terms of safety	Frequent, occur in various places, not many victims at a time, little media interest. Do not affect the profession. Only successes are really analysed (hard luck for the losers); People learn through adversity by analysing the stories of heroes who have survived in exceptional conditions.	Frequent, variable numbers of victims, media pressure sometimes considerable but accidents always result in enquiries and analyses. Learning from past failures mainly involves managing the same suboptimal situation in future (improving detection and recovery, actions focused on managing consequences).	Rare, large numbers of victims and extensive collateral damage, powerful media pressure. What is learned from past failures is essentially never to allow exposure to similar conditions again (better prevention, actions focused on eliminating causes).

(continued)

(continued)

	Resilience model	HRO model High Reliability Systems	Ultra-safe model
Principal feature of the model Who creates the safety	The expert skills of actors and their accumulated experience in the technical domain.	The ability of the group to organise itself (roles), to provide mutual protection to its members, to apply procedures, to be suspicious of excessively routine simplification of the situation, to adapt, perceive changes in the context and make sense of it.	The ability of the regulators of the system to avoid exposing front-line actors to unnecessary risks. The ability of front-line actors to follow protocols.
Standard types of operator training to improve safety	Learning through shadowing, acquiring professional experience, "training for zebra", working on knowing one's own limitations.	Training in teamwork to gain knowledge of the capacity of the group and adaptability in terms of applying procedures to suit the context.	Training in teamwork to apply procedures and apportion the work even if abnormal events occur.
Objective safety level of these systems	Between 10-2 and 10-4	Between 10-4 and 10-6	Better than 10-6
Capacity for progress within the model	Factor of 10	Factor of 10	Factor of 10

Each model represents a response to a type of environment, has its own rules for optimisation and offers little scope for mixing with any other model. It will be noted that safety can be improved within of these models by a factor of 10; a mixed form that incorporates some features of one safety system and other features of another will generally offer no results in terms of improvements in safety (it will be unsuitable or even counterproductive).

A Few Additional Rules When it Comes to Taking Action

We have just seen that building safety is not simple. It is necessary to recognise the risks, to build defences and above all to be realistic and choose the correct safety model. That, however, is not enough. Safety conceived in this way, however relevant it may be, needs to be introduced and maintained in a way that always stays in touch with those working on the ground. The management has a vital role to play in this, both in influencing behaviour (not just directing it) and in understanding clearly what the safety plan does not cover. Finally, it is also necessary to win the battle of the future and to avoid simply designing a system that corrects the errors of a past that no longer exists.

The role of management: to do well what it has been decided to do, and to have a clear knowledge of what it has been decided not to do

It makes no sense to create a safety strategy, however good it may be, unless it is understood and communicated. The management needs to have understood the aims (what must be achieved) and also the impasse situations that have been accepted and the reasons for these (what it has been decided not to do, as a trade-off for other commercial or business benefits that one wishes to retain). Education and training of middle management[21] and front-line managers in these two challenges are fundamental to a successful process. There is a plethora of studies in the literature that address this area [56, 57]. A very good summary is presented in the industrial safety manuals of the Institut pour la culture de sécurité Industrielle; these represent an important source of inspiration for this section.

Doing well what it has been decided to do: *the key role of middle management*

The traditional role of a manager is to manage, which means to carry out his duties as well as possible, to plan the activities or directing others. One should also add: a willingness to influence and offer guidance or orientation for his workers. This is the skill that makes a manager a leader. It is vital in order to make improvements in safety, since this is another area where collective mobilisation absolutely requires the management to show leadership, defined as the manager's ability to influence behaviour so that it becomes safer.

On the one hand, each person manages his own priorities in accordance with his own working context and the messages that he receives. People are generally attentive to the concerns of their own line managers, even if they are not explicitly held to account: in other words, if the line manager is not interested in a particular area, it is not very likely that his workers will be interested in it

[21] Middle managers are managers who work at an intermediate level in the hierarchy, between the executive level and front-line managers; they typically run a functional unit, source Uyterhoeven [54], Thakur [55].

either! On the other hand, the fact that personal safety naturally involves every individual, his integrity and his health does not necessarily result in spontaneous mobilisation in this area. For that to happen, each person needs to be aware of the challenges, convinced of the aims and their actions need to be coordinated. In brief, the management's own behaviour in relation to safety constitutes a message that carries much more weight than all the slogans posted on the company premises. These things demonstrate the true value of safety to the enterprise and they strongly influence the extent to which employees are motivated to behave safely.

The management has a key role in translating and monitoring the safety plan. This skill can be defined according to seven fundamental principles which are described in the box below.

BOX: Seven principles for leadership in industrial safety (source: ICSI, 2011)[22]

Principle 1	**Create the safety vision (which must be coherent with the management's values and principles)**
	• Embrace and promulgate the safety policy of the enterprise • Give safety the ranking that it deserves in relation to other challenges • Imagine the future situation that one wants to see, based on the diagnosis of safety • Set targets that are specific, measurable and achievable • Build the vision collectively • On the basis of the vision, define the principles of responsibility and the expectations in terms of behaviour
Principle 2	**Give safety the place it deserves in the organisation and in the management, and guide it in everyday practice**
	• Integrate safety at every level in the organisation • Clarify everyone's roles and responsibilities • Define a progress plan that sets out the vision • Systematically catalogue the obstacles that exist • Make sure appropriate resources are available
Principle 3	**Ensure the safety vision is shared: influence, persuade and promote information feedback**
	• Regularly reiterate the aims and expectations in terms of behaviour • Reiterate these messages • Communicate clearly • Organise and promote the observation and classification of situations involving risk, including detection of weak signals • Create a climate of trust and promote transparency • Encourage the emergence of good practice; encourage and guide initiatives • Remind people that safety is everyone's business

(continued)

[22] ICSI, Leadership en sécurité, Cahiers de la sécurité industrielle, accessed on 29 December at http://www.icsi-eu.org/francais/dev_cs/cahiers/CSI-LIS-pratiques-industrielles.pdf.

(continued)

Principle 4	**Be credible: exemplary behaviour and coherence** • Ensure that all actors are sufficiently competent to allow them to take the safety aims on board • Be competent, fair and consistent in judgements on safety • Be an example in terms of compliance with safety requirements and commitments, even in situations where things have gone wrong • Get personally involved in rolling out the Action Plan • Safety • Be capable of questioning yourself and questioning attitudes, including those of your superiors • Justify your decisions
Principle 5	**Promote team spirit and cooperation across the organisation** • Develop exchanges of views to resolve safety problems • Ensure that coordination resources exist to allow an overview of risks • Promote sharing of tools and methodologies • Bring safety officials together with operational employees on the ground • Make sure everyone feels integrated and has a sense of solidarity • Create connections where aims appear contradictory • Ensure that the traditional practices of the group are not in conflict with transparency or collective progress
Principle 6	**Maintain a presence on the ground to observe, listen and communicate effectively** • Organise visits on the ground • Organise regular meetings with the various professionals • Involve service provider enterprises in site visits, encourage and promote front-line access for managers of service provider enterprises • Emphasise what is going well and reiterate the lessons learned from past accidents • Catalogue problems that people experience in carrying out instructions and look for solutions together • Provide feedback to the actors involved about observations on the ground • Meet the victims of accidents
Principle 7	**Recognise good practices and apply sanctions justly** • Highlight good safety initiatives • Choose key moments for rewards and raising awareness • Communicate non-emotively about what is unacceptable and the related rules on penalties (possibly on a sliding scale) • Carefully analyse the context (technical and organisational environment, training) before applying any penalties and take care always to be fair and just • Be able to justify the penalty completely transparently

Have a good understanding of the safety trade-offs that are accepted in the action plan that is chosen

The management should also understand the measures that have been removed from the safety plan and the reasons for making these sacrifices (often strategic, financial or trade-offs against other priorities).

These decisions make safety more fragile and they need to be taken into account according to a dedicated strategy which involves reducing exposure, where possible, to those risks for which neither procedures nor training exist, and discussed at meetings to the point where operators are at least capable of detecting the situations, as well as providing keys to avoiding them.

For example, modern automated aircraft are protected from stalling by electronic safeguards that take back control from the crew when the flight characteristics breach threshold tolerance values (due to the airspeed or unusual attitude). Under very rare conditions, however, these aircraft may lose these electronic safeguards and the crew may face a loss of control and a stall. The aviation industry has chosen not to fully train crews to deal with these exceptional cases (for reasons of time, cost and technology—simulators are unable to mimic these conditions). In this case it is necessary for everyone to be fully aware of this gap in their competence and to learn to minimise exposure to these exceptional conditions, either through strategic anticipation to avoid conditions in which such stalls can occur (particularly a total loss of airspeed information), or by responding immediately when the alarm signals that these conditions are approaching.

Another example: most health care institutions have risk maps that do not include the risk of care being incorrectly managed by juniors in training posts when they are left alone on call in the hospital on the wards or in emergency departments, particularly at night, on public holidays, in August or at weekends. Hospitals do usually integrate the risk of errors by juniors in their maps, but they do this by stating that protection exists due to their supervision by seniors. In this situation, however, there are no seniors. The hospitals do not include this risk in their analysis because it would force them to avow an attitude which is virtually illegal and unjustifiable, since according to the official protocols it is unthinkable to leave trainees alone [58–60]. In these cases, executives should be aware that no protection exists against this risk and should organise the work as best they can to take this exposure into account: clearer instructions for calling seniors, words actually used when discussing this subject with juniors, teaching generic measures to safeguard patients etc.

Final example: the official risk map for professional sea fishing skippers on trawlers in the North Sea places considerable emphasis on the risks of collisions with tankers or ferries [6, 56]. The resulting response to their safety plan involves protecting authorised fishing zones by keeping them separate and avoiding overlaps with major shipping routes and Motorways of the Sea, and involves fitting these vessels with systems to detect other vessels (ARPA-Automatic-Radar Plotting Aid). This risk map, however, fails to take into account two factors relating to the fundamental economics of a highly competitive profession which is subject to quotas (first come, first served, best paid): “the fish are not subscribers” to the zones reserved for fishing, and the anti-collision systems broadcasting your own position (as intended) may therefore attract other fishing skippers to your fishing ground. It is not surprising that fishing skippers often work outside the permitted zones and intentionally turn off their anti-collision systems so that they cannot be tracked, but this situation is not covered by the risk analysis. Once again it is important to ensure that these practices are open to discussion among skippers

(rather than leaving them under a "code of silence") to ensure that the risk matrix being used is not based on naïve assumptions.

These three examples set out quite well the four commonest reasons for excluding an (identified) risk from the safety plan: excessive interference with the business model, too rare, too expensive to manage, too inappropriate to be admitted.

Practical rule: it is always useful, at the end of the process of building a safety plan, to go through all the areas that have been sacrificed, using the above categories to catalogue and classify them. The risks that are identified as a result should be treated in a specific way, which is more about sharing information and awareness about the fact that they exist, that they must be identifiable and that every effort must be made to avoid encountering them.

Thinking in the Future Rather Than the Past

Risk management techniques are essentially built using the rear-view mirror. They read the past to generate warnings for the future, and they aim to stabilise the world to safeguard the usefulness of lessons drawn from the mistakes of the past. The process is integrated in the form of an evolution rather than a revolution, and the inferences in terms of risks tend to be linear.

Unfortunately, the world is not quite so linear: it evolves through ruptures and sudden alterations, sometimes after two or three decades of stability. It is technical innovations that most commonly lead to professions being suddenly displaced. In a very short time, a technological system can become completely obsolete, together with its safety rules. In less than 15 years, for example, this change has occurred with the replacement of silver-based photography with digital photography, the end of cathode ray tube televisions and the explosion of mobile technologies. These revolutions have also affected a number of major industrial systems: the ongoing transition from oil prospecting and operations to large-scale oil shale recovery, the ongoing transition from essentially human-based air traffic control to automated regulation (datalink), the arrival of new and lighter building technologies and materials permitting structures to be built twice as quickly with characteristics that had never been achieved before in terms of height and on fragile ground, the arrival of new automobile engines, the switch from invasive surgery to non-invasive surgery by percutaneous or natural routes, the expected end of blood transfusions, as it is replaced by the production of blood from stem cell cultures, etc.

The essential point to remember is that the technology never changes by itself. It also changes the way the system is organised, its business model and its actors (new arrivals benefit from the technological leap to replace the old ones, using the example of digital photography), in short the whole model is replaced, together with its entire risk map and the defences that need to be built.

As a result it is essential to keep one eye on the horizon and constantly question the safety model that has been built on the foundations of the past. As technologies

move forward more quickly, prospective methods may prove to be more effective than retrospective methods at avoiding the accidents of tomorrow.

Safety thinking in a system undergoing rapid change: the example of medicine. Medicine has entered into a period of rapid change, and a major crisis is coming in the long term, at least over the next 10 years (by 2020 or 2025). This situation is true in many industrial and service sectors: whether these are large sectors such as the nuclear industry in the post-Fukushima age, international banking and finance and the search for a new model, the aviation industry which is undergoing global restructuring, the oil industry which is facing the exhaustion of natural resources… or smaller high-risk sectors such as professional fishing, whose survival is threatened on a daily basis.

There are four types of forces that are simultaneously acting on all these sectors:

- the advent of radical innovations affecting both the substance and organisation of the work: in medicine: minimally invasive surgery by natural routes, genomic and personalised medicine, plus a series of other discoveries (oral chemotherapy etc.) which are spectacularly reducing the length of stays in hospital and thus creating the need for a different model of a short-term hospital (fewer technical hospitals, fewer beds, supplemented by a "hospital at home"). The equivalent in industry could be the change in the traditional oil exploration industry to intensive exploitation of oil shales…
- a sociological transformation of what is on offer and the professions involved; a drastic reduction in the number of surgeons, in favour of interventional professionals using minimally invasive surgical techniques (cardiologists, radiologists, gastroenterologists, ophthalmologists) resulting in a challenge to the historical position of operating theatres and the possible relocation of some of these procedures away from hospitals and into local primary care practices. We are also seeing a huge feminisation of health care actors and doctors in primary care, with a tendency to create joint practices in small towns, where everyone works part time and as a result rural areas become complete deserts with no medical provision at all. This leads to the introduction of telemedicine and the delegation of medical work to local nursing staff, and then from nursing staff to patients and their carers at home. In industry, one example of the equivalent change is the growing arrival of professions related to new forms of energy, whether it is solar, wind or fuel cells for the forms of transport of the future (to what extent will we still need engineers specialising in traditional combustion engines by around 2025?);
- an unlimited demand for safety with powerful pressures in the direction of transparency and external supervision;
- and clearly an unprecedented financial crisis, which is bringing the safety model face to face with the economic model more than ever before.

And Where is the Safety Culture in All This?

The reader of a book on risk management and safety management will expect such a book to address the safety culture and may even expect it to be quite a central subject. You will have noticed that this one does not. The aim is not to deny or reject the value of this concept, but to give the concept its true value in terms of the scale of its contribution towards safety. The first thing that is noticeable when doing this is the huge variability in the way the concept of the safety culture is used in the literature and the meanings that are given to it, so that one might conclude that the culture is closely linked to the safety model, but is rarely a concept that permits direct, primary action to improve safety. The time required to bring about changes is long, very long, and the process of enhancing the culture requires real perseverance in order to reap the rewards.

The themes of the safety culture and the safety climate are among the most popular subjects for publications in scientific journals specialising in industrial risks (and risks in public services, transportation and medicine). Most of the articles and books propose tools to evaluate the culture, especially questionnaires.

What, however, can truly be learned from these concepts in order to improve the safety of the system? This question deserves to be asked, because the answer is rather uncertain.

Cultures and Climates (of Change, Effectiveness and Safety), Multiple Areas of Ambiguity and Confusion

There are seven characteristics that dominate the literature on cultures; almost all of them raise fundamental questions about the usefulness of this concept for the purpose of improving safety.

1. Cultures are about values (significant ideas) and norms (expected behaviours) which are (1) moral, shared by all the individuals in a given community (social mores, relationships between men and women, relationships with the truth), (2) ethical (unacceptable conditions for success or failure in the community), and (3) social (definition of success, hierarchical distances, relationship with uncertainty, roles and expertise).
2. The concept of culture was first used to characterise national communities or enterprises long before it was extended to the specific context of the safety culture. In the context of national cultures, the work of psychosociologist Geert Hofstede [61] is a well-known reference point. It identifies five dimensions, and by combining these national cultures can be classified in relation to each other (without making value-judgments). The five dimensions are (1) the degree of hierarchical distance, (2) the need to reduce uncertainty, the degree of tolerance that a culture is able to accept in the face of concerns over future events (3) individualism versus collectivism, (4) the masculine macho versus feminine

dimension and (5) orientation towards the short or long term; links to traditions if the orientation is short-term, values of economy and perseverance if the orientation is long-term. Other major contributions have been put forward to characterise enterprise cultures (or organisation cultures or corporate cultures), particularly by O'Reilly, Chatman and Caldwell [62] who identify seven dimensions: innovation, stability, respect for people, focus on results, attention to detail, group solidarity, competitiveness and desire to win (aggressiveness). Finally, it would be difficult not to quote the important contribution made by sociologist Edgar Schein [63] who identifies three levels in a business culture: a visible level (artifacts) which shows the observable behaviours and rituals (this is typically the level that reflects the concept of the climate of an enterprise), the level of conscious values (values), which comprises the shared beliefs about the enterprise, its strong points, its weak points, its enemies, its friends, and finally a third level (the organisation's tacit assumptions) consisting of the tacit values, unconscious aspects or taboos that are shared but must not be named by the actors (unspoken rules), for example: "in this hospital we practice euthanasia for patients reaching the end of their lives in order to control the workload for staff".

At the outset none of these approaches explicitly referred to a value-classification of nations or enterprises, but they were all quickly used by other authors to describe "good cultures" and "bad cultures". From this time all kinds of difficulties began to arise.

Clearly the first problem in seeking to classify one culture in relation to another is to state what result is desired: in fact nations or enterprises can be classified according to their commercial performance, their capacity for change, their safety and many other criteria. It is no surprise that the classification systems in relation to a "good culture" will therefore differ depending on which criterion is used. Worse still, a good culture according to one scale (capacity for change, or efficiency), may turn out to be a culture that performs less well according to another scale (for example safety). So the first level of difficulty is: a culture is never good according to all the advantages that one might value according to every dimension. If enterprises choose to talk in terms of a "good culture" in the area of safety, this may lead them to adopt cultural traits that may be unhelpful or may even handicap them in relation to other key aspects of the challenges that they face in order to survive.

3. The concept of a safety culture is not homogeneous in any "genetic" sense. The same term is used to refer to very different theoretical approaches.

- Helmreich [64], Flin [65], Guldenmund [66] and many other authors (most of them following the other approaches set out below) have addressed the subject of the safety culture by looking through the prism of psychosociological theories on small groups and the roles of leaders and front-line managers, while prioritising the way in which front-line operators view their working environment. Many questionnaires have been developed on this basis, both on assessing cultures and on assessing the safety climate. It is no doubt due to these questionnaires, which

are readily available, that these approaches have become so popular in the industrial world when "diagnosing the safety culture and diagnosing human and organisational factors".[23] The points that are identified in relation to a good safety culture by such "questionnaire diagnostics" are: a democratic leadership style, respect for everyone's role (in the hierarchy), respect for procedures, absence of a blame culture, ability to report errors/events/incidents without punishment, a sense that the hierarchy is listening, a high level of solidarity and mutual assistance within the group, low numbers of industrial accidents etc.

- Other authors have focused their definitions of a good safety culture on the way in which the management (middle and top management) deals with incidents and accidents (Westrum [67], Reason [68]) by insisting on the need for an in-depth analysis; some have gone even further by insisting on the sanctions that should be linked to these undesirable events, while pointing out the absolute need to maintain the system's ability to avoid judicial consequences when human error is involved—since this is necessarily involuntary—(concept of just culture [69, 70]).
- A very wide-ranging theoretical framework of organisational theories has provided the inspiration for approaches that address governance cultures and the macro-scale organisation of the system, and ultimately these are quite distant from perspectives centred on small groups and operators. This work has fed into the development of knowledge on enterprise cultures, for example by linking the quality of production, the capacity for innovation (climate of creativity [71]) and different families of cultures: tribal cultures, change cultures, cultures reliant on hierarchies and rational cultures[24] and by evaluating the specific characteristics and the progress that can be achieved in each type of culture.
- Still in the context of organisational theories centered on risk, the HRO (High Reliability Organizations) approach should definitely be mentioned. This specifically describes a good safety culture as consisting primarily of the ability of the group to adapt to non-standard situations, stressing the importance of leadership, expertise and everyone playing their role, and above all resilience or even improvisation, two ideas that were little discussed (or were even contradicted) in earlier streams. The HRO diagnosis of a culture is carried out not by using questionnaires but rather by auditing organisations.
- Ultimately, others have very strongly (exclusively?) equated the safety culture with the quality culture, from the perspective of improving the productivity and performance of the system; examples of this are the Toyota way [73] and Lean management. Once again, we are quite a long way from the earlier theories, with a culture that prioritises an organisation centred on the flow, gives the front-line management a key role in reducing the errors that cause falling performance, and manages quality in the production line, while considering the issue of serious accidents only to a very limited extent.

[23] Daniellou F, Simart M, Boissière I. Human and organizational factors of safety: state of the art, ICSI, http://www.foncsi.org/media/PDF/CSI-HOFS.pdf.

[24] A good summary of this whole approach in Braithwaite et al. [72].

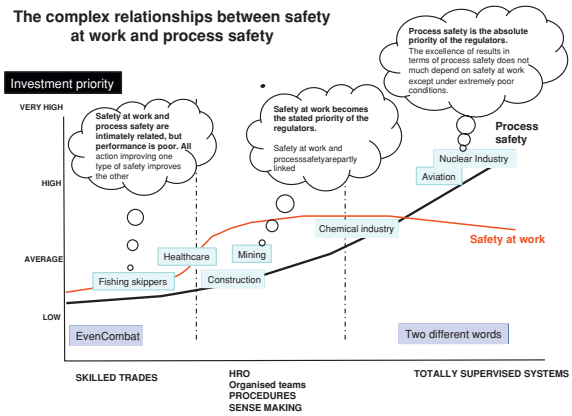
This list is a long way from reflecting all the contributions and theoretical currents that exist in the area of the safety culture: different theories, different messages and a different focus. Most of the time, occasional readers and users cannot be aware of this range in its entirety and find themselves locked into a single point of view or approach. As a result they do not understand the contradictions that may arise if they mix an approach centred on the production line around a Toyota Way or Lean Management type culture, while simultaneously stating that they aim to be an HRO organisation and at the same time maintaining—in different circles—that the priority of the enterprise is to develop the culture and adopt a climate that promotes change so that it can address the coming challenges posed by new socioeconomic conditions.

In short, the phrase "changing the saefty culture" can easily hide major ambiguities and can lead to huge disappointment if it is actually rolled out at the operator level with no precautions. In many cases, fortunately—or perhaps unfortunately—the use of this terminology is no more than a fine-sounding form of words intended for external consumption which has no major consequences internally and no real usefulness. Having said that, every enterprise does have a culture, and it is perhaps more important when intervening in safety to understand all the contradictory aspects of that culture.

Cultures and accident rates in civil aviation. Helmreich [74] showed, building on the work done by Hofstede in the early 2000s, that crews from highly collectivist countries in which considerable hierarchical distances are maintained (Central America, Central Asia) had accident rates in passenger aircraft which are two and a half times higher than those in Western countries, which are characterised by more individualistic cultures without any hierarchical distance. Initially the interpretation naturally tended towards the idea of making value-judgements when classifying cultures and towards putting forward the idea that everyone should adopt the characteristics of Western cultures since these are associated with the best safety results. This hypothesis, however, very soon ceased to be used in such explicit terms. Ethical grounds played an important part in this (the need to avoid insulting these countries or their national values), but the real reason was more trivial: the study had simply revealed something quite obvious: the design of a complex system (like an aircraft) is profoundly marked by culture; it is much easier for users who share this culture to use the product correctly: modern, automated aircraft require very direct collaboration between crew members and expect subordinates to be able to constantly express surprise and question the actions and decisions of their boss; countries that maintain more distant hierarchical relationships inevitably have difficulty working within this model. To some extent there is no such thing as a "good culture", but there are bad "marriages".

4. **A poorly understood link between two drivers that produce the safety culture: worker safety and site and product safety.** Quite strangely, the literature has developed two parallel frameworks in which safety culture diagnostics are applied within enterprises: safety at work and site and product safety. Three observations can be made about this ambiguity:

   – the priority given to each of these two areas depends on the maturity and the public priorities of regulators in various types of industries and public services. The two priority curves intersect. For the most immature and least safe activities (skilled trades, medicine), the public priority mostly concerns production safety (for example patient safety); in industries which are more labour-intensive and more mature than skilled trades and have powerful regulators, a higher priority is given to safety at work (reducing the number of industrial accidents). Paradoxically, for the safest industries (in terms of industrial accidents), the priority switches back to process safety; these industries, which are the safest (in terms of the risk of accidents) often have no more than average performance in terms of safety at work (and certainly lower than the industries in the previous group);
   – other than general ideas, we are not very aware of the theoretical links that exist between these two areas of safety. The link is clearly a complex one, since excellence in one of these two areas is rarely associated with excellence in the other;
   – this ambiguity continues in the use of many tools to measure cultures and safety climates, particularly questionnaires, which have often been validated for only one of these two areas but are still used without any precautions to assess the other.



5. **It is possible to change a safety climate quickly, but a safety culture cannot be changed quickly.** The concept of a safety climate, which was first inspired by Schein (op. cit.) refers to objective aspects (facts), while the concept of the safety culture refers to subjective aspects (values). It may

be possible to change certain elements of the safety climate quite quickly through dedicated actions by the management or organisation, but the values that characterise a culture cannot be changed as quickly. Having said that, it is possible to significantly influence the culture by making authoritative changes to the fundamentals of the technical system and introducing major changes to the economics of the system, but of course this is beyond the limitations of an ad-hoc intervention in an enterprise in an industrial sector or service (such as a hospital or a bank). In short, the market economy dictates the culture rather than vice versa. The levers for change are systemic, not local.

Changing the culture in civil aviation: a systemic lever far ahead of the human factors lever. The passenger aviation sector was long the domain of heroes, in which captains were in command under God alone, deciding which route to take and making exceptions to procedures whenever they considered it appropriate. The introduction of air traffic control after the Second World War represented the first limitation on this autonomy, but it was above all the tremendous global standardisation of the supervision of the flight system, the arrival of automated aircraft that erased handling differences between pilots and the recording of all actions taken by the crew in the cockpit (systematic flight analysis) that definitively tipped the culture of civil aviation in the 1980s towards the ultra-safe model. The introduction of crew training during the 1990s and the initiatives towards voluntary no-blame reporting, which was given a high profile in the media, accompanied these changes rather than being the real reason for changes in the culture (which is now characterised by equality among actors, a high level of transparency surrounding incidents and teamwork centred on coordination and monitoring of procedures).

6. **There is no ideal culture, but there are cultures that are suitable for every situation.** This perspective has gradually become established as the only one that can cope with the real situation. All normative approaches to this area have turned out to be counterproductive. We have seen in the paragraphs above that there are several different models of safety rather than just one. It makes sense that these different models of safety, which reflect different trade-offs between flexibility, competitiveness, adaptability and safety performance, should be based on different ways of managing the safety culture.
7. **The development of the characteristic values of a culture takes a very long time.** Some people speak of a generational lever. None of the standard risk matrices or safety action plans cover such long periods of time.

In the end, assessing the safety culture of a production unit is a useful activity and forms an integral part of a diagnostic process. It requires quite an in-depth knowledge of the theories behind the measurement tools, to avoid the occurrence of contradictory effects. It is not, however, sufficient in itself. The interpretation of such assessments is always relative, because it depends on local challenges (which need to be properly analysed and understood). Finally, this measure makes it possible to ascertain the margin available for progress in the domain of safety that is available to the enterprise.

> If a local safety intervention has to be undertaken in an enterprise within a specific period of time, rather than expecting to change its culture, the opposite approach should be taken: deducing (from an assessment of the culture) what margin exists for real progress to be achieved by the enterprise, in view of its culture.

This diagnostic process will promote the identification of the safety model that best characterises the environment (and the needs) within the enterprise being audited. In short, the culture of an enterprise cannot be changed by a single intervention motivated by a demand for safety. There is no action at all that will achieve this. It is possible, however, to understand and identify the culture that does exist, in order to assess what margin exists within that culture to improve the results (reverse approach).

If a more ambitious approach is adopted that claims to be able to change the culture within a profession, one must have systemic levers, change the demands of the business model at the level of the whole profession, take action at least at the regional level if not nationally or internationally and sustain this action over the long term (long-term intervention is vital, with regulatory systems designed for the purpose).

To the extent that the safety culture is a consequence of the economics of the profession and its safety model rather than a cause of that model, it is legitimate that this paragraph should come at the end of this text rather than at the beginning.

References

1. Reason J (1990) Human error. Cambridge University Press, Cambridge
2. Reason J (1997) Managing the risks of organizational accidents. Ashgate, Aldershot
3. Heinrich HW (1931) Industrial accident prevention. McGraw-Hill, New York
4. Hollnagel E, Woods D, Levison N (2006) Resilience engineering: concepts and precepts. Ashgate, Aldershot
5. Dekker S (2004) Ten questions about human error. A new view of human factors and system safety. Ashgate, Aldershot
6. Morel G, Amalberti R, Chauvin C (2008) Articulating the differences between safety and resilience: the decision-making of professional sea fishing skippers. Hum Factors 1:1–16

7. Roussel P, Moll MC, Guez P (2007) Étape 2: Identifier les risques a priori. Risques & Qualité en milieu de soins 4:239–247
8. Roussel P, Moll MC, Guez P (2008) Étape 3: Identifier les risques a posteriori. Risques & Qualité en milieu de soins V-1:46–58
9. Tengs T, Adams M, Pliskin J et al (1995) Five hundred life-saving interventions and their cost-effectiveness. Risk anal 15(3):369–390
10. Johnson C (2003) Failure in safety-critical systems: a handbook of accident and incident reporting. University of Glasgow Press, Glasgow
11. Reason J (1997) Managing the risks of organizational accidents. Ashgate, Aldershot
12. Marx D (2001) Patient Safety and the "Just Culture": a primer for health care executives. Columbia University, New York
13. Dekker S (2007) Just culture, balancing safety and accountability. Ashgate, Aldershot
14. Guldenmund F (2007) The use of questionnaires in safety culture research—an evaluation. Saf Sci 45:723–743
15. Guldenmund F (2000) The nature of safety culture: a review of theory and research. Saf Sci 34(1–3):215–257
16. Flin R (2007) Measuring safety culture in healthcare: a case for accurate diagnosis. Saf Sci 45:653–667
17. Lawton R, Parker D (2002) Barriers to incident reporting in a healthcare system. Qual Saf Health Care 11:15–18
18. Cullen D, Bates D, Small S et al (1995) The incident reporting system does not detect adverse drug events: a problem for quality improvement. Jt Comm J Qual Improv 1:541–548
19. Jha A, Hupeerman G, Teich J et al (1998) Identifying adverse drug events. JAMIA 5:305–314
20. Goldman RM, de Leval AP, Cohen MR et al (2004) Pitfalls of adverse event reporting in paediatric cardiac intensive care. Arch Dis Child 89:856–885
21. Vincent C, Stanhope N, Crowley-Murphy M (1999) Reasons for not reporting adverse incidents: an empirical study. J Eval Clin Pract 5:13
22. Evans SM, Berry JG, Smith BJ et al (2006) Attitudes and barriers to incident reporting: a collaborative hospital study. Qual Saf Health Care 15:39–43
23. Ricci M, Goldman AP, de Leval MR, Cohen GA, Devaney F, Carthey J (2004) Pitfalls of adverse event reporting in paediatric cardiac intensive care. Arch Dis Child 89:856–885
24. Dekker S (2007) Just culture, balancing safety and accountability. Ashgate, Aldershot
25. Classen D, Resar R, Griffin F et al (2011) Global trigger tool shows that adverse events in hospitals may be in times greater than previously measured. Health aff 30:581–589
26. Resar R, Rozich J, Classen D (2003) Methodology and rationale for the measurement of harm with trigger tools. Qual Saf Health Care 12:39–45
27. Rozich JD, Haraden CR, Resar RK (2003) Adverse drug event trigger tool: a practical methodology for measuring medication related harm. Qual Saf Health Care 12:194–200
28. Sjohania K, Sampson M, Ansari M et al (2007) How quickly do systematic reviews go out of date? A survival analysis. Ann Int Med 147:224–233
29. Hollnagel E (2004) Barriers and accident prevention. Ashgate, Aldershot
30. Gerstein M, Ellsberg M, Ellsberg D (2008) Flirting with disaster: why accidents are rarely accidental. Sterling Publishing, New York
31. Ostberg G (2009) Some intangibles in human handling of risks. Lund University, Sweden
32. Chateauraynaud F, Torny D (1999) Les sombres précurseurs: une sociologie pragmatique de l'alerte et du risque. EHESS, Paris
33. Amalberti R (2006) Optimum system safety and optimum system resilience: agonist or antagonist concepts? In: Hollnagel E, Woods D, Levison N (eds) Resilience engineering: concepts and precepts. Ashgate, Aldershot, pp 238–256
34. Woods DD (2005) Creating foresight: lessons for resilience from Columbia. In: Starbuck WH, Farjoun M (eds) Organization at the Limit: NASA and the Columbia Disaster. Blackwell, Hoboken, pp 289–308
35. Aslanides M, Valot C, Nyssen AS, Amalberti R (2007) Evolution of error and violation description in French air force accident reports: impacts of human factors education. Hum Factors Aerosp Saf 6:51–70

36. Amalberti R, Vincent C, Auroy Y, de Saint Maurice G (2006) Framework models of migrations and violations: a consumer guide. Qual Saf Healthc 15(suppl 1):i66–i71
37. Amalberti R (2001) La maîtrise des situations dynamiques. Psychologie Française 46–2:105–117
38. Rasmussen J (1997) Risk management in a dynamic society. Saf Sci 27:183–214
39. De Saint Maurice G, Auroy Y, Vincent C, Amalberti R (2010) The natural life span of a safety policy: violations and migration in anaesthesia. Qual Saf Health Care 19:327–331
40. Cooper D (2009) Behavioral safety interventions. Professional Safety: 37. http://www.behavioural-safety.com/articles/behavioral_safety_interventions_a_review_of_process_design_factors.pdf
41. Amalberti R, Auroy Y, Berwick D, Barach P (2005) Five system barriers to achieving ultrasafe health care. Ann Intern Med 142(9):756–764
42. Morel G, Chauvin C (2007) A socio-technical approach of risk management applied to collisions involving fishing vessels. Saf sci 44:599–619
43. Wilson R (1979) Analyzing the daily risks of life. Technol Rev 81:40–46
44. Amalberti R, Auroy Y, Berwick D, Barach P (2005) Five system barriers to achieving ultrasafe health care. Ann Intern Med 42:756–764
45. Perrow C (1984) Normal accidents: living with high-risk technologies. Basic Books, NY
46. Bainbridge L (1987) Ironies of automation. In: Rasmussen J, Duncan K, Leplat J (eds) New technology and human errors. Hoboken, Wiley publishing, pp 271–286
47. Woods DD, Hollnagel E (2006) Joint cognitive systems: patterns in cognitive systems engineering. Taylor & Francis, Boca Raton
48. Endsley MR, Garland DJ (2000) Situation awareness analysis and measurement. Lawrence Erlbaum, Mahwah
49. Endsley M (1995) Toward a theory of situation awareness in dynamic systems. Hum Factors 37:32–64
50. Sarter N, Woods D (1992) Pilot interaction with cockpit automation: operational experiences with the flight management system. Int J Aviat Psychol 2(4):303–321
51. Amalberti R, Deblon F (1992) Cognitive modelling of fighter aircraft's control process: a step towards intelligent onboard assistance system. Int J Man-Mach stud 36:639–671
52. Morel G, Amalberti R, Chauvin C (2009) How good micro/macro ergonomics may improve resilience, but not necessarily safety. Saf Sci 47:285–294
53. Grote G (2012) Safety management in different high-risk domains—all the same? Saf Sci 50(10):1983–1992
54. Uyterhoeven HE (1972) General managers in the middle. Harvard Bus Rev 50:75–85
55. Thakur M (1998) Involving middle managers in strategymaking. Long Range Plan 31:732–741
56. Hopkins A (2005) Safety, culture and risk, 1st edn. CCH Australia Ltd, Australia
57. Hopkins A (2007) Holding corporate leaders responsible. Keeping Good Co 59:340–344
58. Bell CM, Redelmeier DA (2001) Mortality among patients admitted to hospitals on weekends as compared with weekdays. N Engl J Med 345:663–668
59. Aylin P, Yunus A, Bottle A et al (2010) Weekend mortality for emergency admissions: a large multicentre study. Qual Saf Health Care 19:213–217
60. Young J, Ranji S, Wachter R et al (2011) July effect: impact of the academic year-end change over on patient outcomes. Ann Intern Med 155:309–315
61. Hofstede G (1983) Culture's consequences: international differences in work-related values. Adm Sci Q (Johnson Graduate School of Management, Cornell University) 28:625–629
62. O'Reilly C, Chatman A, Caldwell D (1991) People and organizational culture: a profile comparisons approach to assessing person-organization fit. Acad Manag J 34:487–516
63. Schein E (1985) Organizational culture and leadership. John Wiley & sons, Hoboken Ed 2010
64. Helmreich RL, Merritt AC (1998) Culture at work: national, organizational, and professional influences. Ashgate, Aldershot
65. Flin R, O'Connor P, Crichton M (2008) Safety at the sharp end: a guide to non-technical skills. Ashgate, Aldershot

66. Guldenmund F (2007) The use of questionnaires in safety culture research—an evaluation. Saf Sci 45:723–743
67. Westrum R (2004) A typology of organisational cultures. Qual Saf Health Care 13:22–27
68. Reason JT, Carthey J, de Leval MR (2001) Diagnosing vulnerable system syndrome: an essential prerequisite to effective risk management. Qual Health Care 10(suppl 2):i21–i25
69. Marx D (2001) Patient safety and the just culture, a primer for health care executives. MERS-TM. Columbia University, New York
70. Dekker S (2008) Just culture, balancing safety and accountability. Ashgate, Aldershot
71. Ekvall G (1991) The organizational culture of idea-management: a creative climate for the management of ideas. In: Henry J, Walker D (eds) Managing Innovation. Sage, London, pp 73–79
72. Braithwaite JJ, Hyde P, Pope C (2010) Culture and climate in health care organizations. Palgrave MacMillan, Basingstoke
73. Liker J (2003) The Toyota Way: 14 management principles from the world's greatest manufacturer, First edn. McGraw-Hill, New York
74. Helmreich R (1993) Attitudes towards automation across five cultures—NASA report/ University of Texas/FAA Aerospace Crew Research