

J. PARIÈS

Dédale SA & CNRS
Paris, France

Sunday June 6, 2004

08:30-09:15

Room 5B

A GROWING SAFETY CHALLENGE

All risk sensitive industries, including nuclear electricity, chemistry, trains, aviation and the like, but also hospitals and the health care system, are now facing a real safety challenge. Public demand for safety improvements towards “zero risk” is growing. The feeling that any fatal accident is unacceptable is gaining popularity among the customers, the media, or the political people. In France for example, disasters like the Concorde accident, the AZF factory explosion in Toulouse, the BSE disease or the blood transfusion contamination by AIDS, triggered a large and intense public debate and highly advertised judicial trials.

But current safety levels are already very good, and as a matter of fact, very difficult to further improve. To make things worse, there is also an increased social pressure *against* safety: Virtually no activity, even considered of prime importance, can nowadays escape economical competition, and cost reduction programs, while the pressure to reduce duty time, to have more leisure time, to retire earlier, is increasing. Cost reduction is not necessarily synonymous with safety reduction. A better process, a better organisation, a better design can be at the same time more efficient and safer. However, the result of the conflict between economical and well-being strain is that our systems are more and more stressed. Additionally, they are technically and socially larger, more complex, and more tightly coupled to their environmental constraints. This means that they are more opaque from a safety perspective, that their behaviour is more difficult to anticipate, that their failure modes are more difficult to identify.

All this looks like a real challenge, and it is a real challenge. And the next question is: how to cope with it? Can we keep the same strategy, and just do more of what we are already doing, or do we need a real change of perspective? According to R. Amalberti (1996), when a safety strategy is implemented with more and more intensity, its efficiency is first growing, then the return on investment is decreasing and the strategy reaches an “apogee”, before starting to lose efficiency due to negative return. This suggests that when the apogee of the current safety strategy has been reached in a specific domain, what is needed is a change of strategy. Because safety strategies are based on people’s understanding of safety, such a change cannot be obtained without a deep modification of the way people understand safety.

THE AVIATION EXAMPLE

The recent history of aviation provides a good example of that. Aviation training has been oriented towards technical and operational skills since the very first flights. Most accidents were understood as a result of individual pilot error, mainly attributed to a lack of basic flying skills. But it so happened that the 1970s were marked by the occurrence of several accidents “instigated” by good crews (experienced, properly qualified, well considered) flying “good condition” aircraft (without any failure or only minor ones)...

On the 27th of march 1977, on the foggy runway of Tenerife Airport, one of the most dreadful aviation events, the collision of two Boeing 747 “jumbo jets”, was about to become reality. It was day light and the visibility was fluctuating from 300 to 1000 m. Many flights had been diverted to Tenerife airport due to a bomb alert in a neighbouring airport, and the taxiway entry out of the jam-packed apron was blocked, so that the two B747s had to taxi on the runway, the Pan Am following about 1000 m after the KLM, to reach the take-off position. Once at the runway end, the KLM captain turned 180° and initiated the take off. Thirty seconds later, the KLM 747 collided with the Pan Am jumbo, still on the runway. Though the first officer and the flight engineer voiced concerns that they were not really cleared to take off, they were unable to convey this message properly to the captain’s mind.

This accident was the most devastating, but certainly not the only one of its kind. The 70's were marked by a series of air tragedies sharing these highly challenging features: nothing serious was going wrong with the aircraft, the weather was not an issue, and the crew members were highly trained staff employed by major airlines and individually known as good professionals. The aviation community overcame the temptation to consider these disasters as aberrant occurrences, and accepted the challenge of questioning how experienced, well trained, and incidentally well paid professionals could have deceived the average expectations by such a

scale. A milestone towards the answer was provided at an international conference held in 1978 under the aegis of NASA: the Cockpit Resource Management (CRM) concept was raised for the first time. It was a dramatic acceptance that the then prevailing safety equation – *individual pilot proficiency plus aircraft reliability equal safe flights* - had been proven wrong. Evidence was provided regarding the connection between crew resource management deficiencies and the above mentioned accidents.

CRM corresponds to a revolution in accident causation perception. Roughly speaking, the years up to the early 80's in aviation can be seen as the 'Lonesome Cow-Boy' years, when individual pilot errors or violations were considered the key safety issue. Then came the 'Crew years', with the stress being put on the role of team performance. It was recognised that a crew is not an addition of individuals, but an interaction between individuals, so that personalities, attitudes, lack of communication skills, poor leadership, can lead to poor interaction.

And the shift continued. While the rapid expansion of cockpit automation actually delivered most of the intended benefits (higher accuracy, higher efficiency, higher reliability, automated protections), it also led to new difficulties, and even to new kinds of accidents, in which the core of the tragedy scenario was a wrong or limited understanding of the automated systems behaviour (e.g. Nagoya, 1996; Cali 1996). These difficulties highlighted the role of crew-cockpit interactions, and particularly the role of cockpit design. It also raised questions like: "would a better crew knowledge about the systems design improve their ability to maintain situation awareness?".

So the focus has shifted from the crew to a larger system, to include the aircraft, the training system, the ground staff, the procedures, and so on. Accidents were then understood as resulting from a loss of control by the crew, and beyond the crew, by the whole company, on the error management process. The accident model was extended to include all the dimensions of human action, (including cognitive ones), all the interactions (including the man/machine one) and all the people (company resource management). Front line operators behaviour was acknowledged to be strongly determined by organisational influences (selection, training, procedures, cultures, work conditions, organisation structures). So every accident appeared as a failure of the organisation.

A SAFETY PARADIGM SHIFT

What the recent history of aviation shows is actually *a shift in the "safety paradigm"*. A "safety paradigm" is the set of fundamental rules and principles that people believe to be both the definition of and the conditions for safety. The safety paradigm drives the principles underlying accident & incident investigation, the design philosophy & concepts, the training philosophies, the perceived role of procedures, the perceived role of punishment, and the like. Safety paradigms are not stable in time. They come within the scope of a historical move through a series of crises, generally triggered by a series of major accidents, or major changes in the socio-technical background.

TWO MAIN EVOLUTIONS

And this is not specific to aviation. Across most industrial domains, two elements have particularly changed in the comprehension of safety during the past fifty years,:

- The *size of the system* which is considered relevant in order to identify the causality of a safety event has expanded from individuals (one operator, one workplace), to local teams (one team, one workshop), then to global interactions (a whole company). In other words, the "system" considered has grown from individuals to organisations. Meanwhile, the causality *time scale* has been considerably extended, from real time, « hands on », front line operators actions, to longer term, senior managers decisions, possibly made years before the consequential event. The tendency has been a fundamental shift in the notion of cause: from direct causality to indirect and remote, or "root" causality.
- The *nature of the explanation* through which people satisfy themselves that they "understand" the "cause" of a safety event. The tendency has been to dig deeper and deeper, with a second fundamental shift in the notion of cause: from failures, errors and deviations, to a loss of control on the system's dynamics.

TOWARDS AN ORGANISATIONAL PERSPECTIVE ON SAFETY

The former trend introduces the notion of “organisational accidents”. The latter refers to a “systemic approach” to accidents and safety. These two perspectives are very often mixed up. Nevertheless, they are not really equivalent. One can look at what happens on a single work place as a complex interaction between a human operator, a machine, a procedure, and a specific environment. One can also look at a large organisation with a simple cause-and-effect model in mind. Actually, the “organisational” point of view addresses a specific system: the structured and organised collective activity of a group of humans within the framework of a company. The “systemic” point of view attempts to consider all the elements of a socio-technical system, and mainly the interactions between the humans, in different roles, the technical equipment, the “software” resources (procedures, knowledge, data bases...), and the environment.

TWO PERSPECTIVES ON ORGANISATIONS

There are two main perspectives on organisations.

The first is interested in the forming, the evolution, the transformation, the life duration of organisations. It mainly tackles the way organisations interact with, and adapt to, their environment, match their internal structures and features to their external constraints (e.g. “Contingency Theory”). A reliable organisation is seen as an organisation which is able to respond adequately to its environmental demands and changes. Rochlin and his colleagues (1987) from the “Berkeley School” found that “High Reliability Organisations (HRO)” shared the following features: i) total adherence of staff to objectives ii) redundant decision and monitoring channels iii) continuous learning process iv) both centralized (strategic) and distributed (operational) decisions. Symmetrically, Perrow (1984) made the point that a tight coupling of organisations to their environmental constraints, added to a tight and rigid coupling of their internal components, would inevitably lead to accidents that he called “Normal Accidents”.

The second perspective is more interested in the internal game. It focuses on the relationships between the different actors within the organisation, and mainly the power relationship. The structures and functioning rules are seen as the result of decisions and arbitration made by the different players considering their objectives, their resources, and their constraints. With his famous ‘Swiss Cheese’ model, Jim Reason (1990) linked managerial decisions with real time safety through the notion of latent failures and active failures: senior management decisions propagate downwards to determine working conditions, work practices and safety defences, hence create the psychological precursors that will determine front line operators reliability.

TOWARDS A MORE SOPHISTICATED APPROACH OF SAFETY

The second thinking shift suggested previously is about the nature of accidental mechanisms. We tend to think that failures, errors, deviations, cause accidents because most of the time, accident scenarios include failures, errors, or deviations. But such events are high frequency events. How could something happening up to several times in every safe flight *really explain* something as exceptional as an accident? As a matter of fact, we tend to confuse the initial deviation process – like the psychological *process* of errors (a deviation from the operator’s intention) - with the unexpected (and in some cases devastating) *effect* of deviations (e.g. erroneous actions) in a specific system. Actually, the ‘resistance’ and /or the ‘tolerance’ of the system to errors is a much more critical and relevant safety parameter than the number and the nature of errors produced. These features are determined by the design of the system, and are a pivotal safety factor.

A second pivotal safety factor is the ability of the system, from the whole organisation (e.g. through incident reporting and analysis systems) to front line operators, to detect and evaluate the potential consequences of errors and deviations, in order to correct or mitigate the ‘risky’ ones without wasting resources in correcting the harmless ones (Wioland & Amalberti 1996). In other words risk management is the key issue. All human operators are risk managers. They permanently arbitrate between conflicting goals. They permanently face and manage ‘deviations’: their own errors and violations, as well as technical failures and all kinds of unexpected events. And one of the key conditions of success for risk management is an adequate risk perception, including ‘feeling the limits’ properly. From this perspective, paradoxically, raw errors serve safety: they ‘flag’ the limits, and the ‘flag’ can be memorised. We mainly learn our risk management skills from error recovery.

This introduces the idea that all philosophies of safety are a particular mixture of two main components:

- The first component, *safety through invariance*, calls on invariant functional properties of the world (the system considered and its environment) to elaborate relevant (safe) operational responses (e.g. procedures). The ideal achievement is to design responses which are perfectly adapted to a situation demand. These solutions are then memorised under the format of a functional program (e.g. procedural knowledge - rules and skills, or automated system software) and implemented according to the situation. The main problem is then to remain within the frame of a known situation, and to stick to the solutions. Safety is impaired when the environment varies (de-adaptation) or when the standard solution is not adhered to (error or violation).
- The second component, *safety through adaptation*, calls on a more generic, more flexible strategy. Operational responses are permanently imperfect, because they include latent, open solutions to different situations. They use procedural knowledge as well as declarative knowledge (more generic and abstract properties of the world). They include a monitoring and a management of random or unforeseen variations both externally in the environment and internally in the operational responses (failures, errors and deviations). Keeping control on these deviations is a process similar to biological immunity principles: defences are mainly based on pathogen recognition (identity paradigm) and ironically they need aggressors to develop.

Table 1 summarises the implications of these two components.

TABLE 1 - TWO SAFETY PHILOSOPHY COMPONENTS

Safety Philosophy Component 1	Safety Philosophy Component 2
Operations can be entirely specified through standardised procedures, programs, schedules, rules, nominal tasks, certification, selection, norms,...;	Operations cannot be entirely specified through standardised procedures, programs, and the like. One reason is it includes Humans.
Safety improvement will result from more specification (more extensive, comprehensive, and detailed procedures, ...) and more discipline from the operators;	Safety improvement will result from a better respect of the «ecology» of the system and a better acknowledgment of its self-protection mechanisms
Deviations from nominal operation are both a cause of lower performance, and the main threat for safety;	Deviations from nominal operation are both a necessity for adaptation to random dimension of real life, and a potential threat for safety;
Human operators are ultimately the only unpredictable and non specifiable components of the system. They are the main source of deviation;	Human operators are up to now the only intelligent, flexible and real time adaptable component of the system. They are a deposit and source of safety;
Automation, whenever feasible and reliable, will decrease deviation rate and therefore improve both performance and safety;	Automation will increase reliability, improve performance, and make the operation more rigid. As long as Humans are kept in the system, automation will also make their environment more complex, and create new problems in man-machine coupling
Errors are non intentional but regrettable deviations from standard actions. Errors are unfortunately inevitable;	Errors are deviations from operator's intentions, but at the same time they are part of the normal process of achieving intentions. Errors are necessary;
Errors are just as negative for safety as any other deviation. Every effort should be made to reduce the number of errors;	Uncorrected errors may be a threat for safety. However, self error awareness is a critical governor of operator's behavior and food for risk management processes (regulator of confidence level)

The cornerstone of safety improvement for the last 50 years in all industrial domains can be seen as a drastic change in the balance between these two components, with a total triumph of normative safety (component number 1). Indeed, the dominant strategy has been based on:

- improved technical reliability (less variation)
- front line operators training, with an expanding use of simulation, to increase skills (more routine based responses)
- development and improvement of procedures (less variation)
- more and more automation (less variation)
- fighting against deviations, through work place design, procedures design, persuasion, discipline, blame and punishment (less deviations)
- incident analysis to understand error and deviations (and hopefully eradicate them)

However, such a normative safety approach has probably reached its apogee in large, highly complex, tightly coupled systems like aviation or nuclear industry. Over the last ten years, most of the accidents in aviation that happened to the last “glass-cockpit” generation are related to a loss of control on the situation by the crew due to a loss of understanding of automation (programmed) behaviour.

In other domains, and certainly in the medical field, there is still a lot to be gained from boosting the normative component. But at the same time, it must be fully acknowledged that, as long as humans are needed in a system, they are both a hazard source and a permanent protection. We need their knowledge and their adaptability, in a word their intelligence. And this means that we need also to boost the adaptive component.

TOWARDS A COMPLEX SYSTEM PERSPECTIVE

The difference between organisational and systemic perspectives on safety has been introduced previously. The “systemic” point of view attempts to consider all the components of a socio-technical system, and mainly the interactions between the humans, in different roles, the technical equipment, the “software” resources (procedures, knowledge, data bases...), and the environment. As a matter of fact, “organisational safety” models and “systemic safety” models refer most of the time to non-complex systems.

A “complex system” is a compound system, including a large number of similar and interacting components, in which specific behaviour, and very often sophisticated ones (like adaptation to environmental changes), *emerge* at the global system level (macro level), as a result of simple interactions between basic components (micro level).

Ants’ nests are an example of complex systems. Ants have no brain, no intention, no understanding of what they do. They react automatically to a few external stimuli, like the presence of chemical substances within their (very short: a few millimetres) perception horizon. For example, oleic acid released by a dead ant body will incite worker ants to seize and carry the body. But ants have no memory and they just drop the body after a random distance. Surprisingly, the global outcome of such simplistic individual behaviour is that ants clear the ants’ nest of dead bodies and pile them up in specific places (this can be reproduced on a computer simulation of virtual ants). As a matter of fact, collectively, ants gather dead ants bodies in “cemeteries”, with the effect of reducing infection risk within the colony. Not only they have no idea of it, but such a collective property of the “ants’ nest system” cannot be derived from the knowledge of individual ant’s behaviour laws only.

To understand that global outcome, one needs to develop a macroscopic model accounting for the interactions between individual ants behaviour. In the “simple” cases (e.g. ants behaviour), it is possible to develop a deterministic model predicting the global behaviour. In the case of the incredibly more complex human systems, that’s another story! Some specific human collective behaviour have been successfully modelled through a complex systems approach (e.g. some crowd behaviour). But the immense majority of collective human behaviour have not been modelled (and stock exchange behaviour still resists any real prediction...). Macroscopic models of socio-technical systems, describing global properties (e.g. HRO model), and microscopic models (e.g. cognitive models), describing local properties do not communicate.

One important consequence of this disconnection between macro and local models of safety is that the ideas we have about safety tend to be overly linear and dependent on direct causality. Accidents are seen as the result of individual decisions or actions (at different levels within the organisation), or as a strike of chance, rather than as an outcome of multiple interacting behaviour. Unfortunately, the functioning and the failure modes of complex systems are counter-intuitive, and challenge traditional, linear models. Even with a case as "simple" as ants' nets, the model turns out to be non linear*. Complex systems are the scene of circular causality, with strong feedback and feed-forward effects that can augment or inhibit the consequences of an action, with very long term and remote effects ("ripple" effect). They have stable states, generally resist disturbance forces with a huge inertia or plasticity, and then shift to another state of balance when the proper parameter is changed (leverage effect). They can also be the scene of divergent processes ("butterfly effect").

Consequently, the next challenge is to drive the safety paradigm shift towards a complex systems approach, to bridge the gap between macroscopic and microscopic models.

REFERENCES

1. Amalberti R. (1996) *La conduite de systèmes à risques*, Paris: PUF
2. Controlled Flight Into Terrain, American Airlines Flight 965, Boeing 757, N651 AA, near Cali, Colombia, December 20, 1995. Aircraft Accident Report, Aeronautica Civil of The Republic of Colombia, Santafe de Bogota, DC, Colombia (1996)
3. e.g. Nagoya (1996) ;
4. Paries, J., Amalberti, R., (1994), "Recent Trends in Aviation Safety", Muster Program, European Commission, H. Andersen (Ed), Riso, D Roskilde, DK.
5. Perrow, C. (1984). *Normal Accidents : Living With High Risk technologies*. Basic Books. New York.
6. Reason J. (1990) *Human error*, Cambridge University Press
7. Rochlin, G. I., La Porte, T.R., Roberts K.H. (1987), *The Self-designing High-Reliability Organization : Aircraft Carrier Flight Operations at Sea*, Naval War College Review, Volume 40, Number 4, Autumn, p 76-90.
8. Wioland L. Amalberti R. (1996) *When errors serve safety : towards a model of ecological safety*, Cognitive Systems Engineering in Process Control, November 12-15, 1996, Kyoto, Japan

*The initial bodies density is a critical factor, with several threshold values leading to totally different final outcomes.