Accepted for publication in *Cognition, Technology & Work*. The final publication is available at springerlink.com

I want to believe: Some myths about the management of industrial safety

Denis Besnard & Erik Hollnagel

Mines ParisTech
Centre for research on Risks and Crises
Rue Claude Daunesse
BP 207
06904 Sophia Antipolis
France

Abstract. There are many definitions of safety, but most of them are variations on the theme that safety can be measured by the number of adverse outcomes and must be enforced by using more of the same safety measures (e.g. more procedures). What the industry thinks safety is, and how it can be achieved are both questionable. This article looks at six safety-related assumptions, or safety myths, which can be found across most industry practices. These myths are related to human error, procedure compliance, protection and safety, root causes, accident investigation, and 'safety first'. We argue that while relying on such myths makes safety management easier, they also make it flawed and ineffective. Therefore, we propose a critical discussion leading to a set of managerial messages. In this paper, we present the six myths in sequence, and then discard them through the analysis of various industrial examples. We then replace the myths with six clearer statements in order to build an alternative view of safety.

Keywords. Industrial safety; safety management; human error; procedure compliance; protection and safety; root causes; accident investigation; safety first

1 ON MYTHS AND SAFETY

In the best of all possible worlds, safety is managed by highly trained and rational people using carefully chosen indicators and effective methods. In reality, safety management is usually a collection of best practices based on a number of assumptions that are taken for granted, hence rarely discussed. Examples include the traditional dictum of 'safety first', the belief that increasing protection will increase safety, or the notion that most accidents are caused by human error. These, and other, assumptions are common to many (if not all) industrial sectors and determine both individual attitudes, corporate policies and regulatory practices. Since these assumptions express common beliefs rather than facts, they are not verifiable and can therefore be considered as myths.

This paper will consider six major safety myths and try to challenge them on the basis of alternative views. The origin of these myths is first-hand interaction with the industry during consultancy work. More precisely, they originate from direct exchanges that took place with (or material that was gathered from) large European companies. The selection of the myths was done on the basis of the perceived frequency of occurrence and the potential impact on safety. Strictly speaking, we cannot claim that these myths are representative of the beliefs held by the industry. However, because they are implicit indicators of a safety culture, these myths might be regarded as the building blocks of an industrial safety policy. For these reasons, the discussion of the myths that follows is not a rhetorical exercise but aims to disseminate scientifically-informed managerial messages.

1.1 Myths do matter

For the scope of this paper, an assumption is something that is taken for granted rather than verified. Assumptions, whether as hunches, guesses, or hypotheses, are an important and essential part of human activity since we rarely have sufficient time to make sure that what we assume is actually true (Hollnagel, 2009a). While assumptions are usually considered in relation to what individuals think and do, they may also be shared among social or professional groups. This is illustrated by the common definition of safety culture as *"that assembly of characteristics and attitudes in organizations and individuals which establishes that, as an overriding priority, nuclear plant safety issues receive the attention warranted by their significance"* (INSAG, 1986). Schein's (1992) definition of organisational culture is also relevant here: it is *"a pattern of shared basic assumptions that the group learned as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems."* Many people happily accept definitions such as these without reflecting on what they actually mean or imply. It is taken for granted that they are sensible, simply because everybody else uses them.

The step from an assumption to a myth is not very large. A myth is an idea or story that many people believe, but which is not true. This definition emphasises both the fictional aspect of myths and their social nature. And because they express simple 'truths', they are also excellent vehicles for communication.

In the industrial world everyone, from the sharp end to the blunt end, seems to share a number of safety-related myths. Risk, in particular, is a social object that is affected by many biases (Bohnenblust & Slovic, 1998). For instance, a manager might believe that in their company, safety comes first. However, the reality of field operations is almost always that people implement workarounds and safety trade-offs so that their task can be completed given the available resources and the constraints at play. Therefore, our objective in discussing myths is

to encourage reflection among the stakeholders of industrial safety, and to propose some alternatives from which new safety practices may be derived.

1.2. Safety

The definition of safety usually refers to the absence of unwanted outcomes, either simply as the freedom from unacceptable risk, or tongue-in-cheek as *"a dynamic non-event"* (Weick, 2001, p. 335). For the discussion in this paper, we will adopt the following working definition: *Safety is the system property that is necessary and sufficient to ensure that the number of events that could be harmful to workers, the public or the environment is acceptably low*. This definition of safety emphasises the relative nature of the concept: safe systems produce *acceptably* low numbers of unwanted events. Alternatively, safety can be seen as producing an *affordable* number of unwanted events, as captured by the As-Low-As-Reasonably-Praticable principle (Woodruff, 2005). This is the starting point for looking at safety assumptions as well as some aspects of the industrial safety culture. In doing so, we will rely on a human factors perspective (see for instance Norman, 1988; Vicente, 2003), even though it has not yet been fully integrated into all industrial safety practices.

In this paper, it would be impossible to provide an extensive discussion of all existing safety-related myths. Instead, we will look at six of them (see for instance Allinson, 2007 for more), provide a short description and attempt to analyse the underlying assumptions. We will do so by providing examples and by referring to evidence that supports reconsidering the assumptions. We end the discussion of each myth with a short alternative formulation.

2 HUMAN ERROR

2.1 The myth

Human error is the largest single cause of accidents and incidents.

2.2 Description and criticism

The web announcement for the Intersec Trade Fair and Conference (held in Dubai, January 2010), included the following topic under the heading *Latest News*: *"Human error is involved in over 90% of all accidents and injuries in a workplace"* (Intersec, 2009).

Numerous books and papers have been written about human error. An increasing number of them openly question the simple-minded use of the term (e.g. Dekker, 2005; Hollnagel & Amalberti, 2001; Woods *et al*., 1994). Yet as the above announcement shows, the myth of human error as the cause of most accidents prevails. Human error is also a fundamental focus of many accident investigation methods and, of course, the very foundation of human reliability assessment. The tradition is old; one of the early candidates for a theory to explain industrial accidents was a single-factor model of accident proneness (Greenwood & Woods, 1919). In methods such as root cause analysis, for instance, the 'human error' level often marks the maximum depth of analysis, in the sense that a human error is accepted as the root cause. In human reliability assessment the focus is still the Human Error Probability (HEP), despite numerous criticisms (e.g. Dougherty, 1990; Hollnagel, 2000). The concept of human error became part of safety lore when Heinrich (1931, p. 43) noted that as improved equipment and methods were introduced, *"accidents from purely mechanical or physical causes decreased, and man failure became the predominating cause of injury."* This assumption became the second of the five dominoes in the famous domino model, described as *"Fault of person – proximate reason for committing unsafe act, or for existence of mechanical or physical hazard"* (Heinrich, 1934, p. 1). Observers of the human mind, philosophers and psychologists alike, have studied human error at least since the days of

David Hume (see Hollnagel, 2000), and have generally treated human error as an individual characteristic or a personality trait. A good example of this is the zero-risk hypothesis of driving (Summala, 1985; 1988), which proposes that drivers aim to keep their subjectively perceived risk at zero-level.

Our daily lives are littered with instances of the expression 'human error'. They can be found in the news, in accident reports, in public statements, etc. Recent examples include a news item reported by the BBC (2009) about a software problem with Google's search services where a 'human error' by a Google employee caused all search results to be unduly flagged as malevolent. Also, the French radio station France Info (Colombain, 2009) announced that a (programming) 'human error' in a piece of software, handling operations on bank accounts at the French BNP bank, caused almost 600,000 debits or credits to be performed two or three times.

The futility of using human error as a cause of accidents can be demonstrated by the following argument. If we consider a safe system to be one where the probability of failure is low, e.g., 10^{-5}, then there will be at least 99.999 cases of acceptable performance for every case of unacceptable performance. In other words, accidents will be quite rare. If the so-called 'human error' is the cause of the event that goes wrong, what the cause is of all the other events that go right? In our opinion, the only possible answer is: humans. Humans try to make sure that their actions produce the intended effect. However, they behave in the same manner regardless of whether the outcomes of their actions turn out to be positive or negative, simply because they cannot know that at the time of acting. It follows that 'human error' should not be used to explain adverse outcomes since it invokes an *ad hoc* 'mechanism'. Instead, a more productive view is to try to understand how performance varies, and determine why the behaviour that usually makes things go right occasionally makes things go wrong.

2.3 A possible revision of the myth

'Human error' is an artefact of a traditional engineering view, which treats humans as if they were (fallible) machines and overlooks how performance adjustments are used to match activities to the working conditions.

The expression 'human error' contains assumptions that are counter-productive for the effective understanding of things that have gone wrong. To start with, it is a judgement; a loaded term that implies some form of wrongdoing and asks for a culprit to be found. What is more, it is a judgement made after the outcome of an action has become known, and is therefore heavily influenced by the hindsight bias (Fischhoff, 1975; Woods *et al*., 1994; 2010). It is also in practice limited to people at the sharp end, i.e., the operators who are directly involved with the process. There are two main reasons for it. First, the consequences of actions are seen almost immediately at the sharp end. It is therefore deceptively easy to associate adverse outcomes with the preceding actions, and therefore also easy to blame the people at the sharp end. Second, in a safety culture that focuses on mistakes and sanctions, blaming someone establishes or maintains a power or authority gradient, which makes management prone to follow the hierarchy to its bottom where the sharp end is conventionally found (Hollnagel, 2004). However, actual work is subject to constraints that are imposed by managers and the higher strata of organisations. This must be taken into account in trying to understand how and when performance varies. Finally, 'human error' focuses on hypothetical psychological or cognitive mechanisms and pays little attention to the context of work. This narrows the cause of mishaps to people's actions, without including how and why operators adjust their performance, i.e. how to *"bridge the gap between what must be done and what can be done"* (Runte, 2010, p. 3).

3 PROCEDURE COMPLIANCE

3.1 The myth

Systems will be safe if people comply with the procedures they have been given.

3.2 Description and criticism

Generally speaking, procedures are essential references for how to carry out a given task. They act, for example, as memory aids, or as guides for decisions in situations where people have little or no experience. Procedures vary in nature, size and complexity and may range from a six-line cooking recipe to entire bookshelves of ring binders in control rooms of nuclear power plants. The safety myth is that safety can be ensured by procedure compliance and conversely that safety is jeopardised by non-compliance.

There is an entrenched belief in the correctness of engineering design, as well as in e.g. the design of interfaces, work specifications and procedures. When any of these fail, the explanation is typically found as to be 'human error' or as a case of violation or non-compliance. The frequent use of non-compliance as an explanation shows that there is a strong belief that everything would go well if only people would follow procedures, rules, and regulations. The belief in procedures is demonstrated by the fact that their number usually keeps growing, even after it has reached a level where procedures cease to be useful. The standard recommendation from a large number of accidents is to reinforce procedure compliance, or even to propose new procedures.

The compliance bias implies that humans, as fallible machines, are a source of uncontrollable variability that contributes to the occurrence of unsafe events. The assumption is that complying with the procedure will not only get the job done, but will also get it done well, i.e., safely. In relation to safety, the idea that safe and effective performance requires procedure compliance reflects the principles of Scientific Management (Taylor, 1911) and also the assumption that people can be considered as machines – possibly complex ones, but machines nonetheless.

Looking at procedures from a human factors standpoint tells a somewhat different story. Procedure compliance will not always guarantee safety. One reason is that procedures are inherently underspecified both in scope and in depth: a procedure cannot cover *all* possible configurations a worker might face when performing a task, nor can it precisely and exhaustively describe *what* a worker has to do, *how*, and *when*. Numerous studies tell us that humans overcome these limitations by interpreting the procedure vis-a-vis the situation and by adjusting it when necessary (Schulman *et al*., 2004). Humans also rely on their experience to interpret procedures when actions are described in fuzzy terms such as ‘enough’, ‘quickly’, ‘slowly’, etc.

An example of operators reacting to an exception by adapting a procedure is the near-loss and recovery of the Snorre A offshore platform (Wackers, 2006; Figure 1).

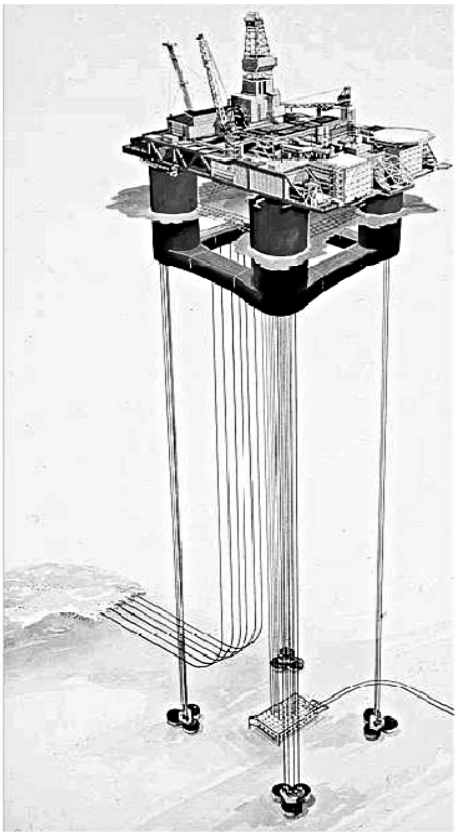

Figure 1: Graphical representation of the Snorre A platform showing the wells on the seabed between thefour tension legs (from Wackers, 2006)

On 28th November, 2004, one of the gas wells attached to the platform started to leak from the seabed. The leak was serious enough for a large gas cloud to build up around the platform, exposing the entire installation to a high risk of explosion. Under such circumstances, safety procedures stipulated that the platform must be evacuated. However, applying this procedure meant that the leak would be left unplugged, and that the unattended platform would be exposed to potential destruction. Should this happen, the platform would sink and crash onto the seabed, obliterating the wells themselves and making the plugging of the leak impossible for a considerable time. With such a catastrophic scenario in mind, and taking into account the long-term consequences, the platform manager decided to remain aboard the platform with a small team in order to plug the well with concrete. They succeeded, stopped the leak, and were able to put the platform back into service; it still is in service today.

In this case, the mandatory compliance with a safety procedure was deliberately disregarded in order to respond proactively to the potential evolution of a catastrophic situation. Although the objection can be made that this analysis benefits from hindsight, the reaction to the event at the time is a clear demonstration of a ‘safe violation’ (Besnard & Greathead, 2003). In the analysis of this case, we wish to highlight that the mere departure from a procedure cannot a priori be interpreted as an impediment to safety. The conditions under which such departures from procedures lead to expected consequences is a crucial question, but one that falls outside of the scope of this research.

3.3 A possible revision of the myth

Actual working situations usually differ from what the procedures assume and strict compliance may be detrimental to both safety and efficiency. Procedures should be used carefully and intelligently.

Safe operations cannot be ensured by rigid and blind compliance. Instead they require that operators assess the adequacy of, and adapt, procedures to operational conditions (Dien, 1998; Besnard, 2006). Humans constantly perform this assessment and fill in the gaps between the assumed and actual conditions of their task. This is why there is always a difference between work as imagined and work as done. Given that it is impossible to anticipate all the possible configurations of a work situation and prescribe each and every step of an activity, industrial operations depend on these adjustments. Therefore, strict procedure compliance may actually have detrimental consequences since it will limit the beneficial effects of human adaptation in response to underspecification of the work situation. A good example of this is the situation where people 'work to rule' to express their dissatisfaction with working conditions, or Vicente's (1999) description of malicious compliance.

4 PROTECTION AND SAFETY

4.1 The myth

Safety can be improved by barriers and protection; increasing the layers of protection leads to higher safety.

4.2 Description and criticism

Practically all definitions of safety (see for instance ICAO[1] or WHO[2]) define safety as the absence of accidents or as the freedom from unacceptable risks. It follows from this that safety can be achieved either by eliminating risks or by protecting against their effects. At first glance it seems reasonable to expect that safety is higher the more protection there is, and the better designed it is. It is the philosophy behind safety in many systems such as motor vehicles, where multiple active and passive safety systems (Anti Blocking System, crumple zones, safety belts, airbags, etc.) protect drivers from physical injury. It is also the philosophy behind the concept of defence-in-depth (INSAG, 1996). This notion has been institutionalised by the nuclear industry (IAEA[3]) and is seen in practice in most workplaces. The physical structure of industrial plants demonstrate the steadfast adherence to the principle of multiple barriers. However, multiple barriers only lead to higher levels of safety if they always function as intended. The latter is rarely the case.

.

As far as individuals are concerned,the relation between risk exposure and protection is not simple. In the case of protective equipment, one factor is the expected utility of being protected versus being unprotected. A rule of thumb is that protection will be used when the feedback from not being protected is negative, immediate and tangible. In the absence of negative feedback, using a protection might be disregarded, even if life is at stake. An example is the safety belt in cars. It is obviously possible, although risky, to drive for an entire lifetime without using the safety belt and without having an accident. As a matter of fact, this was the norm until a few decades ago. The message here is that the perceived consequence of risk exposure is a strong determinant of human behaviour, which may defeat the purpose of

1 International Civil Aviation Organisation

2 World Health Organisation

3 International Atomic Energy Agency

barriers and protection. The weaker the link between risk exposure and consequences, the less likely it is that a protection will be used. The psychological explanation is that the potential negative outcomes are less salient (or available; Tversky & Kahneman, 1974).

There are two main reasons why more protection is not necessarily better. One is psychological and has to do with risk homeostasis (Wilde, 1994) whereby people adjust their risk exposure to the perceived level of protection. The classical example is the introduction of the ABS braking system in the automotive industry. A large-scale study conducted by Aschenbrenner and Biehl (1994; quoted by Wilde, 1994) showed that taxi drivers whose cars were equipped with ABS tended to drive more aggressively. The other interesting result was that the drivers of ABS-equipped vehicles had an accident rate that was slightly *higher* than that of the other taxi drivers. This clearly demonstrates the counter-intuitive nature of the human response to increased protection. Another example is the consequence of equipping winding Finnish country roads with reflector posts: people drove faster and closer to the edge of the road, thereby vastly increasing the number of accidents at night (Hamer, 1991).

The second reason why more protection is not necessarily better is technical. Adding protection invariably increases the complexity of the system, regardless of how that is measured (more components, more couplings, etc.). The added components or functions may not only fail themselves, but will also significantly increase the number of combinations that can lead to unwanted and undesired outcomes. In addition to making common mode failures more likely, this also makes understanding the system more difficult.

4.3 A possible revision of the myth

Technology is not value neutral. Additional protection changes behaviour so that the intended safety improvements might not be obtained.

Expected increases in safety from additional barriers and protection can be defeated by psychological reactions. The introduction of new and better (safer) technology should not be treated as the simple replacement of one function by another. Any change will affect the established equilibrium, and people will usually respond by changing their behaviour. The risk homeostasis hypothesis described above is one illustration of that. A more general expression of this principle is found in the *Law of stretched systems,* originally proposed by Lawrence Hirschhorn. The law states that:

"Under resource pressure, the benefits of change are taken in increased productivity, pushing the system back to the edge of the performance envelope" (quoted in Woods & Cook, 2002, p. 141).

This means that every system is stretched to operate at its full capacity and that any (technological) improvement will be exploited to achieve a new intensity and tempo of activity. Rather than simply enabling humans to manage existing risks better, additional barriers and protection may lead people to take greater risks in order to improve efficiency. Of course, this does not mean that less protection is the way to improve safety or that increased protection never works. It only means that one should carefully consider both the intended and the unintended effects of implementing protection in socio-technical systems.

5 MISHAPS AND ROOT CAUSES

5.1 The myth

Root cause analysis can identify why mishaps happen in complex socio-technical systems.

5.2 Description and criticism

Root Cause Analysis (RCA) is a generic term referring to a family of methods used to find the various causal factors that can explain a specific adverse outcome such as an accident (for a review, see Livingston, Jackson & Priestley, 2001). This approach is widely applied in many industries. The basic steps of RCA are: determine *what* happened, determine *why* it happened (by going stepwise backwards from an effect to the causes) and finally, find ways to *reduce* the possibility that it will happen again. RCA assumes that the parts of a system are causally related so that effects propagate in an orderly fashion. This assumption justifies the idea that cause-effect links are followed in reverse order to discover where the problem started. In this way, the various steps are mapped out in a tree-like diagram of causes (Figure 2).

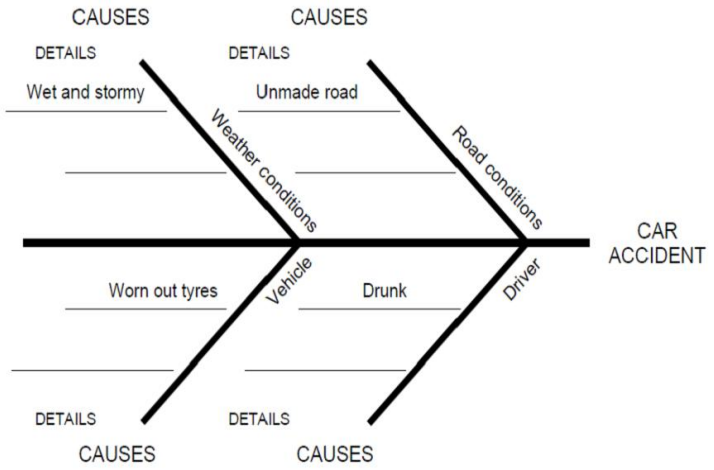

Figure 2: An example of a method belonging to root cause analysis: a fish-bone model of a car accident and causes

RCA is deeply embedded within safety-related practices in the industry. It is also widely taught, to the extent that there are certificates for people who have completed courses. In the field of healthcare (patient safety) it is the most commonly used method. This also applies to other sectors of the industry. In a survey of risk analysis practices in the Norwegian oil and gas industry, Andersen and Mostue (2011) report that 80% of surveyed companies use Fault Tree Analysis (a analysis method belonging to the RCA family). The most cited method that includes human, technical and organisational dimensions is the Norwegian CRIOP, which is cited by less than 60% of respondents. These figures are in sharp contrast with the use of recent systemic methods such as FRAM or STAMP which are reported as being in use in less than 10% of the companies surveyed.

One of the reasons why root cause analysis is attractive to the industry might be its decompositional nature. However, the validity of its methods depends on the critical assumption that outcomes of specific events are bimodal; i.e., outcomes are either correct or incorrect (Hollnagel, 2009a). This view is not tenable for the performance of all technical

systems (Manion, 2007), let alone complex socio-technical systems. Indeed, in the latter case, individual and collective human performance normally varies considerably but rarely fails completely. And even when performance for some reason fails, humans can individually or collectively recover from failure and resume normal operation. In short, the very flexibility and adaptability of human performance is a unique contribution to safety (Reason, 2009). Yet this flexibility is disregarded when a root cause analysis points to a human as the origin of an unwanted event. The analysis only sees the failure (see Section) and fails to recognise that things go right and wrong for the same reasons. The possible elimination of the human activity that was deemed the cause of the outcome will therefore also eliminate the far more frequent and far more probable positive contribution.

The preference for clear and simple methods has both psychological and pragmatic explanations. The psychological explanation is what the philosopher Friedrich Wilhelm Nietzsche (2007; org. 1895, p. 33) called the error of imaginary causes:

"To trace something unfamiliar back to something familiar is at once a relief, a comfort and a satisfaction, while it also produces a feeling of power. The unfamiliar involves danger, anxiety and care – the fundamental instinct is to get rid of these painful circumstances. First principle – any explanation is better than none at all."

5.3 A possible revision of the myth

Human performance cannot be described as if it was bimodal. In socio-technical systems, things that go wrong happen in the same way as things that go right.

This means that there are many cases where root cause analysis cannot – and should not – be used. Fortunately, there are several alternatives that are more appropriate. One is the well-established MTO approach that considers huMan, Technical and Organisational factors either alone or in combination. This approach has been used by both nuclear and off-shore industries for more than twenty years (Rollenhagen, 1995). Another is the Swiss cheese model (Reason, 1990), which offers a high-level view of how latent conditions can combine with active failures and thereby lead to unexpected and unwanted outcomes. A more recent proposal is STAMP (Systems-Theoretic Accident Model and Processes; Leveson, 2004). STAMP is a causal analysis method based on a systems theory model that makes a number of assumptions about how the general system is structured. On a different tack, the Functional Resonance Analysis Method (FRAM) replaces the cause-effect relation by the concept of functional resonance (Hollnagel, 2004; Woltjer & Hollnagel, 2007). This approach provides a way to describe unexpected events as emerging from the low-amplitude variability of everyday performance.

6 ACCIDENT INVESTIGATION

6.1 The myth

Accident investigation is the logical and rational identification of causes based on facts.

6.2 Description and criticism

The purpose of accident investigation is to discover the causes of unexpected and adverse outcomes. However, the number of serious unwanted outcomes is so large that it is impossible to investigate them all. Furthermore, when unwanted events do get investigated, it is often necessary that the results are ready by a certain deadline, for judicial reasons, for communication purposes, or because of a lack of resources. Because of that, the depth or extent of analysis, the methods deployed, or the choice of data that are scrutinised are not

simply determined by the particulars of the case at hand (its complexity, severity, potential for learning, etc.). In addition, resources and demands dictate *what* will be done and *how* it should be done. The management of the investigation then becomes a trade-off between what can be done and what should be done: a trade-off between efficiency and thoroughness (Hollnagel, 2009a).

In practice, accident investigations always imply some assumptions about how accidents happen and what one should do to prevent them (Lundberg *et al*., 2009). For example, Benner (1985) evaluated the merits of seventeen investigation methodologies from the USA, and found considerable differences in their effectiveness. Accident investigation guidelines embody a set of assumptions about how accidents happen, what the important factors are, and how events can best be prevented in the future. They also define an implicit (and sometimes explicit) norm for what a satisfactory investigation is.

Another bias is the need to establish responsibilities. This can turn the identification of causes of an event into a secondary issue. This approach is paramount within judicial enquiries. A recent example is the crash of a Rafale air fighter in December 2007 in France, causing the death of the pilot. Shortly after the accident, a representative of the French Air Force declared on a national radio station that all of those responsible would be identified by the investigation. This can be a major obstacle to safety because it confuses responsibility and cause. At best, it makes it easier to find who to blame next time a similar event occurs. But it also means that the investigation fails to meet any common criteria for rationality. As Woods *et al*. (1994, p. xvii) put it, "*attributing error to the actions of some person, team, or organisation is fundamentally a social and psychological process and not an objective, technical one."*

6.3 A possible revision of the myth

Accident investigation is a social process, where causes are constructed rather than found.

An accident investigation must be systematic, hence follow a method or a procedure. The purpose is, however, not to *find* causes but to *build* explanations. There are many different methods available (Benner, 1985; Sklet, 2002), both between and within domains, and these may differ with respect to how well formulated and how well founded they are. The method will direct the investigation to look at certain things and not at others. Indeed, it is simply not possible to begin an investigation with a completely open mind, just as it is not possible passively to 'see' what is there. Accident investigations can aptly be characterised as conforming to the What-You-Look-For-Is-What-You-Find (WYLFIWYF) principle (Hollnagel, 2008b; Lundberg *et al*., 2009).

7 SAFETY FIRST

7.1 The myth

Safety always has the highest priority and will never be compromised.

7.2 Description and criticism

This is by far the most commonly heard myth in the realm of safety management. Here, the assumption is that safety is an absolute priority in the sense that it cannot be compromised. An instance of this myth appeared in a statement by the Chief Executive Officer of Air France, in the aftermath of the AF 447 accident (Amedeo & Ducros, 2009):

"The company is safe. It was safe yesterday, it is safe now but it will be even safer tomorrow. Because everything will be scrutinized: the mechanical parts, human factors, the weather. Every possible accident scenario will be analysed. Everything will be looked at and we will improve the elements that may be related to the accident as well as others that are not. There is no contradiction between safety and economy. When safety improves, it improves the image of the company and therefore also improves its economic performance. There has never been any trade-off between these two areas. For example, it is clearly written that pilots should fly around thunderstorms. There is no question of saving on fuel. Pilots are free to choose their route."

Statements like these are often used because they are concise and suited for communication purposes. They basically express a value that is both clear and noble. If one looks at the safety management of an industry such as commercial aviation in western countries, there are clear examples of 'safety first'. One is scheduled maintenance for aircraft. At regular intervals, every wide-body commercial aircraft goes through total dismantling, down to the very last piece of wire, and is then rebuilt with the necessary upgrades or replacement parts. It is hard to be more dedicated to safety than that. To our knowledge, aviation is the only industry to have adopted such a radical practice, at least in principle. In practice however, economic considerations may sometimes lead to compromises with the schedule (Woltjer & Hollnagel, 2007).

One can also find a discrepancy between policy statements and reality. One example is provided by an assessment of safety behaviour and culture at the BP Texas City refinery that was carried out between 8th and 30th November 2004. In 2004, BP Texas City had the lowest injury rate in its history, nearly one-third the average of the oil refinery sector. In the following year, on 23rd March 2005, a major explosion occurred in an isomerisation unit at the site, killing fifteen workers and injuring more than 170 others. This was the worst American industrial accident in over a decade. The 2004 study interviewed 112 individuals to solicit their views on a number of issues (Telos, 2005). They were asked to rank their perception of the priorities at the Texas City site, using a set of given options. For the purposes of this discussion, the most interesting finding was that the first three choices were *Making money*, *Cost/budget* and *Production*, respectively. *Major Incident* and *Security* only came in fifth and seventh position.

Safety has financial implications that cannot be ignored and it is understandable that costs do have an influence on the choice and feasibility of safety measures. It is all the more understandable because the costs are real and immediate whereas the benefits are potential and distant in time. A further complication is that safety performance is often measured by the relative reduction in the number of situations where things go wrong rather than as an increase in the number of situations where things go right. This means that there is less and less to measure as safety improves. The lack of information can then be (mis)interpreted to mean that the process is under control, when in actual fact the opposite might be the case.

7.3 A possible revision of the myth

Safety will be as high as affordable – from a financial and ethical perspective.

An illustration that 'safety first' is a relative rather than an absolute statement is provided by Negroni (2009):

"In October, the agency, the Federal Aviation Administration, issued an operations bulletin for 'ultra-long-range flights' that doubled the amount of time that pilots and flight attendants must remain at their overseas destination. The change to 48 hours from 24 was intended to

ensure that flight crews got two full periods of sleep before making the return flight. But seven airlines have asked the U.S. Court of Appeals for the Federal Circuit in Washington to set aside the new requirements, arguing that they would impose 'substantial burdens and costs."

In other words, safety comes first if the organisation can afford it. If not, safety is traded off against economy. It is not realistic to expect that all possible safety measures, however good they may be, can be implemented without prioritisation, or without considering feasibility and consequences. This was clearly illustrated in the reactions to the eruption of the Eyjafjällajökull volcano in April 2010. The safety-centred initial response was to close most of the European airspace. But after some time other concerns became more important. On 19th April for example, the 27 European Union transport ministers, in a meeting with representatives of the air transport industry, agreed to (partly) resume flying after taking both safety and economic considerations into account. It is inevitable that air safety is traded-off against other concerns. Indeed, if safety really was the overriding priority, the few aviation accidents that occur annually in e.g. western Europe would be a sufficient reason to ground all flights.

8 DISCUSSION

In this paper, we have looked at six commonly held safety assumptions and compared them to actual practices, policies, and scientific knowledge. The assumptions were seen as myths for the following reasons:

- They are shared by large groups of people inside and outside of companies, including managers, politicians, and sometimes the public, and can be found in various industrial sectors and social contexts.
- They express a set of attitudes and values that determine decisions and actions related to safety.
- They are not usually noticed or questioned.
- They resist change.
- They cannot be verified.

Safety involves all layers of organisations, from operators to CEOs, as well as society in the form of investigation boards, regulators, and the courts. Since the myths permeate every layer, the practice of safety is affected everywhere. Taken together these myths, and potentially others, are part of the common safety culture, i.e., the pattern of shared assumptions that affect how we perceive, think, and respond to adverse outcomes. Because they are myths and therefore rarely questioned, it will take more than just facts and reason to undo them and alleviate their effects. An alternative approach may be to consider the object of safety myths, namely safety itself.

First, instead of defining safety solely as a system property, we argue that safety should also be seen as a process. Safety is not something a system *has*, it is something a system *does*. Safety is under constant negotiation, and the way that safety is managed varies continuously in response to, and in anticipation of, changes in operating conditions – as well as changes in demands and resources. Therefore, from an operational point of view, it is crucial to understand how an organisation *produces* safety.

Second, we would like to emphasise that the goal of safety should be to increase what goes right rather than to reduce what goes wrong. Consequently, safety indicators should measure

what goes right rather than what goes wrong. Today we have many methods that focus on unsafe functioning but few, if any, that focus on safe functioning. Yet the aim of safety should not only be to reduce the number of adverse events. It should also improve the ability to succeed under varying conditions. This is consistent with the principles of resilience engineering (e.g. Hollnagel, Woods & Leveson, 2006; Hollnagel *et al*., 2011), which defines safety as the ability of an organisation to succeed under varying conditions. Therefore, the preoccupation with what goes wrong overlooks what allows an organisation to sustain acceptable everyday performance.

In light of this, it seems odd that safety is measured by simple, context-free performance indicators such as fatality rates or accident tallies. Safety should rather be tied to indicators that account for the way an organisation maintains its stability through changing conditions. Such indicators can be the effectiveness of control of the process at hand, the amount of available resources, the degree of social acceptance, the level of sustained effectiveness, the quality of outcomes, and so on. These indicators are compatible with resilience engineering, where safety comprises the abilities to respond, to monitor, to anticipate, and to learn (Hollnagel, 2009b). This view is consistent with Slovic's (2001) plea that traditional safety measurements (fatality rates or accident frequencies) should be combined with others, in order to measure the efficiency of the (safety) process rather than the number of outcomes. As disturbing as it might seem, this would recognise the way that safety is actually managed: as a complicated trade-off.

9 CONCLUSION

Together, the myths discussed here represent a conviction that it is possible to achieve safety by properly engineering systems, including the people who work in them. This is congruent with the view that safety is simply the absence of failures. More precisely, it is assumed that systems work because: (1) systems are well designed and scrupulously maintained, (2) procedures are complete and correct; (3) people do what they have been taught or trained to do; and (4) system designers are able to foresee and anticipate every contingency (Hollnagel, 2008a). Taken together the myths describe well-tested and well-behaved systems where human performance variability clearly is a liability and where the human inability to perform in an expected manner is a risk.

While the above view may have been reasonable fifty years ago (and even that can be disputed), it is clearly not reasonable today. One question is therefore: why do these myths still exist? In our opinion, one reason is that they simply do not get questioned. In turn, this might be because managing safety on a day-to-day basis is more a matter of producing and managing conventional indicators (Chaplin, 2009; Hopkins, 2009) or complying with safety checks (Hodkingson, 2009) than understanding and anticipating, e.g. the causes of mishaps. Therefore, we need to discard the existing myths and instead replace them with clearer statements. For the six myths considered in this paper, the following revisions were proposed:

- *'Human error' is an artefact of a traditional engineering view, which treats humans as if they were (fallible) machines and overlooks how performance adjustments are used to match the working conditions.*
- *Actual working situations usually differ from what the procedures assume and strict compliance may be detrimental to both safety and efficiency. Procedures should be used carefully and intelligently.*
- *Technology is not value neutral. Additional protection changes behaviour so that the intended safety improvements might not be obtained.*

- *Human performance cannot be described as if it was bimodal. In socio-technical systems, things that go wrong happen in the same way as things that go right.*
- *Accident investigation is a social process, where causes are constructed rather than found.*
- *Safety will be as high as affordable – from a financial and ethical perspective.*

According to this revised view, complex socio-technical systems work because: (1) people learn to identify and overcome design flaws and functional glitches; (2) people can adjust their performance to the current conditions (resources and demands); (3) people can interpret and apply procedures to match the situation; and (4) people can detect when something is about to go wrong, and intervene before the situation becomes seriously worsened. This means that systems work because people are flexible and adaptive, rather than because the systems have been perfectly thought out and designed.

We live in a complex world, where work takes place in conditions of multiple interacting technical, financial, cultural and political constraints. Doing things perfectly under such conditions is hardly a feasible option. But a view of safety management that involves complicated trade-offs does not blend well with the ideal of a well thought-through endeavour, driven by scientific knowledge and practices, and conducted by rational people. The safety myths described in this paper all derive from this ideal. As myths, they are counter-productive because they lead to unrealistic safety management attitudes, policies and targets. In order to have any chance of successfully operating increasingly complex socio-technical systems, we need to abandon the myths and the idealised approach to safety that they imply.

10 ACKNOWLEDGEMENTS

This article was written thanks to the sponsorship of the industrial partners of the Industrial Safety Chair at Mines-ParisTech (Afnor, Allianz, Apave, Arcelor-Mittal, GDF-Suez, Ineris, SNCF, Total and the alumni association of Mines-ParisTech). Also, the authors are grateful to the anonymous reviewer who helped improve the quality of this paper.

11 REFERENCES

Allinson, R. E. (2007). Risk management: demythologising its belief foundations. *International Journal of Risk Assessment and Management*, 7, 299-311.

Amedeo, F. & Ducros, C. (2009). AF 447: Tous les scénarios du drame vont être analysés. *Le Figaro*, July 9th. On-line at http://www.lefigaro.fr/actualite-france/2009/07/08/01016-20090708ARTFIG00504-af-447-tous-les-scenarios-du-drame-vont-etre-analyses-.php (last accessed on 12 Oct, 2011).

Andersen, S. & Mostue, B. (2011). Risk analysis and risk management approaches applied to the petroleum industry and their applicability to IO concepts. *Safety Science* (in press).

Aschenbrenner, M. & Biehl, B. (1994). Improved safety through improved technical measures? Empirical studies regarding risk compensation processes in relation to anti-lock braking systems. In R. M. Trimpop & G. J. S. Wilde, *Challenges to accident prevention: The issue of risk compensation behaviour*. Groningen, The Netherlands: Styx Publications.

BBC (2009). '*Human error' hits Google search*. On-line at http://news.bbc.co.uk/2/hi/technology/7862840.stm (last accessed on 12 Oct, 2011).

Benner, L. (1985). Rating accident models and investigation methodologies. *Journal of Safety Research*, 16, 116-126.

Besnard, D. (2006). Procedures, programs and their impact on dependability. In Besnard, D., Gacek, C. & Jones, C.B. (Eds) *Structure for Dependability: Computer-Based Systems from an Interdisciplinary Perspective.* London, Springer.

Besnard, D. & Greathead, D. (2003). A cognitive approach to safe violations. *Cognition, Technology & Work*, 5, 272-282.

Bohnenblust, H. & Slovic, P. (1998). Integrating technical analysis and public values in risk-based decision-making. *Reliability Engineering and System Safety*, 59, 151-159.

Chaplin, R. (2009). Process safety indicators: Response to Andrew Hopkins. *Safety Science*, 47, 467.

Colombain, J. (2009). Quand le cybermonde s'emballe. *France Info*, 02 march, 2009. On-line at http://www.france-info.com/spip.php?article259661&theme=34&sous_theme=35 (last accessed on 12 Oct, 2011).

Dekker, S. (2005). *Ten questions about human error*. Mahwah, NJ: Lawrence Erlbaum.

Dien, Y. (1998). Safety and application of procedures, or how do 'they' have to use operating procedures in nuclear power plants? *Safety Science*, 29, 179-187.

Dougherty, E. M. Jr. (1990). Human reliability analysis - Where shouldst thou turn? *Reliability Engineering and System Safety*, 29, 283-299.

Fischhoff, B. (1975). Hindsight – foresight: the effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance,* 1(3), 288–299.

Greenwood, M. & Woods, H. M. (1919). A report on the incidence of industrial accidents upon individuals with special reference to multiple accidents. *Reports of the Industrial Fatigue Research Board*, 4, 3-28.

Hamer, M. (1991). Safety posts make roads more dangerous. *New Scientist*, 1786.

Heinrich, H. W. (1931). *Industrial accident prevention*. McGraw-Hill.

Heinrich, H. W. (1934). *The accident sequence*. Presentation given to the Detroit Industrial Safety Council, November 30, 1934.

Hodkingson, M. (2009). Process safety indicators: Response to Andrew Hopkins. *Safety Science*, 47, 469.

Hollnagel, E. (2000). Looking for errors of omission and commission or the hunting of the Snark revisited. *Reliability Engineering and System Safety*, 68, 135-145.

Hollnagel, E. (2004). *Barriers and accident prevention*. Aldershot, Ashgate.

Hollnagel, E. (2008a). From protection to resilience: Changing views on how to achieve safety. Proceedings of the *8th International Symposium of the Australian Aviation Psychology Association*, April 8-11, Sydney, Australia.

Hollnagel, E. (2008b). Investigation as an impediment to learning. In Hollnagel, E., Nemeth, C. & Dekker, S. (Eds.) *Remaining sensitive to the possibility of failure* (Resilience engineering series). Aldershot, UK: Ashgate.

Hollnagel, E. (2009a). *The ETTO principle: Efficiency-Thoroughness Trade-Off: Why things that go right sometimes go wrong*. Aldershot: Ashgate.

Hollnagel, E. (2009b). Extending the scope of the human factor. In E. Hollnagel (Ed.), *Safer complex industrial environments*. Boca Raton, FL: Taylor & Francis.

Hollnagel, E. & Amalberti, R. (2001). The Emperor's New Clothes, or whatever happened to "human error"? *4th International Workshop on Human Error, Safety and System Development*. Linköping, Sweden.

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

Hollnagel, E., Paries, J., Woods, D. D. & Wreathall, J. (Eds.) (2011). *Resilience engineering in practice: A guidebook*. Farnham, UK: Ashgate.

Hopkins, A. (2009). Thinking about process safety indicators. *Safety Science*, 47, 460-465.

INSAG (1986). *International Nuclear Safety Advisory Group. Summary Report on the Post-Accident Review Meeting on the Chernobyl Accident*. INSAG-l. Vienna: IAEA.

INSAG (1996). *International Nuclear Safety Advisory Group. Defence in depth in nuclear safety*. INSAG-10. Vienna, IAEA.

Intersec (2009). On-line at http://www.intersecmiddleeast.com/index.asp?url_id=1&pgName=Home (last accessed on 12 Oct, 2011)

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

Livingston, A. D., Jackson, G. & Priestley, K. (2001). *Root causes analysis: Literature review*. Research Report 325/2001. Norwich, UK: Health & Safety Executive.

Lundberg, J., Rollenhagen, C. & Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find - The consequences of underlying accident models in eight accident investigation manuals. *Safety Science*, 47, 1297-1311.

Manion, M. (2007). The epistemology of fault-tree analysis: an ethical critique. *International Journal of Risk Assessment and Management*, 7, 382-430.

Negroni, C. (2009). Air safety debate focuses on rest. *International Herald Tribune*, March 4.

Nietzsche, F. (2007; org. 1895). *Twilight of the Idols*. Ware, Hertfordshire: Wordsworth Editions Limited.

Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.

Reason, J. (1990). *Human error*. Cambridge, Cambridge University Press.

Reason, J. (2009). *The human contribution*. Aldershot: Ashgate.

Rollenhagen, C. (1995). *MTO—En Introduktion, Sambandet Människa, Teknik och Organisation*. Lund, Sweden: Studentlitteratur.

Runte, E. (2010). *Productivity and safety: adjustments at work in socio-technical systems*. Doctoral dissertation, Mines-ParisTech, France.

Schein, E. H. (1992). *Organizational culture and leadership*. Jossey-Bass, San Francisco.

Schulman, P., Roe, E., Eeten, M. v., & Bruijne, M. d. (2004). High Reliability and the Management of Critical Infrastructures. *Journal of Contingencies and Crisis Management,* 12(1), 14-28.

Sklet, S. (2002). *Methods for accident investigations*. Report N° ROSS (NTNU) 200208, Norwegian University of Science and Technology, Norway.

Slovic, P. (2001). The risk game. *Journal of Hazardous Materials*, 86, 17-24.

Summala, H. (1985). Modeling driver behavior: A pessimistic prediction? In Evans, L. & Schwing, R. C. (Eds), *Human behavior and trafic safety* (pp. 43-65). New York: Plenum.

Summala, H. (1988). Risk control is not risk adjustment: The zero-risk theory of driver behaviour and its implications. *Ergonomics, 31*(4), 491-506.

Taylor, F. W. (1911). *The principles of scientific management*. On-line at http://www.gutenberg.org/dirs/etext04/pscmg10.txt (last accessed on 12 Oct, 2011).

Telos (2005). *BP Texas City site report of findings. Texas City's protection performance, behaviors, culture, management, and leadership* (BPISOM00122320). The Telos Group.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.

Vicente, K. (1999). *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Mahwah, NJ: Lawrence Erlbaum Associates.

Vicente, K. (2003). *The human factor: revolutionizing the way we live with technology*. Toronto, Knopf.

Wackers, G. (2006). *Vulnerability and robustness in a complex technological system: Loss of control and recovery in the 2004 Snorre A gas blow-out*. University of Maastricht, Research report 42/2006.

Weick, K. E. (2001). Making sense of the organization. Oxford, UK: Blackwell Publishing.

Wilde, G. J. S. (1994). *Target risk. Dealing with the danger of death, disease and damage in everyday decisions*. On-line at http://psyc.queensu.ca/target/index.html#contents (last accessed on 12 Oct, 2011).

Woltjer, R. & Hollnagel, E. (2007). The Alaska Airlines flight 261 accident: a systemic analysis of functional resonance. Proceedings of the *14th International Symposium on Aviation Psychology*, Dayton, OH.

Woodruff, J. M. (2005). Consequence and likelihood in risk estimation: A matter of balance in UK health and safety risk assessment practice. *Safety Science*, 43, 345-353.

Woods, D. D. & Cook, R. I. (2002). Nine steps to move forward from error. *Cognition, Technology & Work*, 4, 137–144.

Woods, D. D., Johannesen, L. J., Cook, R. I. & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight*. Columbus, OH: CSERIAC.

Woods, D. D., Dekker, S., Cook, R., Johannesen, L. & Sarter, N. (2010). *Behind human error (second edition)*. Farnham, UK: Ashgate.