The ‘Wheel of Misfortune’: a taxonomic approach to human factors in accident investigation and analysis in aviation and other complex systems

D. O’HARE*

Department of Psychology, University of Otago, Dunedin, New Zealand

Keywords: Accident investigation; Human error.

The analysis and reporting of the human factors aspects of accidents in aviation and other complex systems continues to present difficulties for investigators and analysts alike. Reason’s ‘latent conditions’ model has had a major impact on the way accidents are conceptualized but it has proven difficult to apply as a practical tool. Recent attempts to overcome these difficulties are discussed and an alternative conceptualization is proposed. This conceptualization is based on a blend of several well-supported theoretical models in cognitive engineering and can be used to formulate a parsimonious analysis system for the investigation and reporting of the human factors aspects of accidents. Two well-known examples of transportation disasters are briefly described and related to the proposed conceptual framework. The proposed framework serves three important functions in accident investigation and analysis: a heuristic function, an investigative function, and an integrative function.

1. Introduction

Now that it is widely accepted that human factors are the most significant source of failures in aviation and other complex systems, the attention of accident investigators and safety analysts has been focussed on finding the appropriate theoretical and empirical tools to describe, analyse and communicate the sources of human performance failures.

In terms of theoretical analysis, the theory of latent failures described by Reason (1990) and presented at many major aviation gatherings since (e.g. Reason 1996, 1997a) has had by far the greatest impact. At least one accident investigation authority, the Bureau of Air Safety Investigation in Australia (BASI), has attempted to formulate its investigative efforts around this theoretical model. While the ‘Reason model’ has expanded one’s conceptual understanding of the range of factors that need to be taken into account in investigating and analysing the causes of accidents in aviation and other complex systems, there is a need more effectively to bridge the gap between theory and practice in accident investigation and analysis.

Despite its complexity and breadth, the ‘Reason’ model is generally represented in terms of a linear sequence originating with ‘fallible decisions’ of high-level decision-makers, passing through ‘line management deficiencies’, ‘precursors of unsafe acts’, ‘unsafe acts’ and ‘defences’ (Reason 1990). This representation of a linear sequence of ‘planes’ obscures the fact that accident causation is better thought

*Author for correspondence. e-mail: ohare@psy.otago.ac.nz
of in terms of intersecting influences spreading outward from various points. As Zotov (1996: 8) pointed out, it is in fact ‘the accumulation of factors, rather than just their chaining together, which can affect the crew’.

Zotov used an expanded version of Helmreich’s (1990) ‘Flightcrew environments’ model to illustrate the advantages of a ‘causal web’ rather than a linear sequence perspective on accident investigation. Zotov also notes that lawyers representing organizational interests can easily challenge the validity of a lengthy causal chain by arguing that the accident might still have occurred without one of the early steps in the chain, thus nullifying the early step as a causal factor. Attempts to prosecute high-level managers in the wake of transportation disasters (e.g. after the Herald of Free Enterprise sinking in 1987, described below) have been unsuccessful for this reason.

The simplicity of Helmreich’s representation of the factors affecting flightcrew and its use of concentric spheres of influence provides the necessary basis for the proposed model. The model proposed in this paper uses this form of representation to capture the three levels of analysis contained in the Reason (1990) model, namely: (1) the acts performed by the front line personnel, (2) the local conditions that may have directly precipitated those acts, and (3) the more global organizational processes that govern the work activity. Each of these levels is discussed in detail using concepts derived from a variety of theoretical frameworks. Earlier attempts at model synthesis (e.g. Hale and Glendon 1987) have focussed exclusively on the level of individual behaviour.

It will be argued that the proposed model provides an organizing conceptual framework for addressing human factors issues in accident investigation and analysis, and that it can be used to generate a simple taxonomic system for use in classifying the human factors of complex systems accidents at each of the levels described above. Such a taxonomy would greatly aid the communication and dissemination of human factors information, both within a domain such as aviation, and between domains, which is presently disguised by the use of mutually incompatible frameworks. Section 2 deals with the problems of analysing and classifying human performance in one particular domain, that of aviation.

2. Taxonomies of human error in aviation

Information about human error in aviation can be found in a variety of places. Civil and military investigators generate post-accident investigation reports and summary information is often available from national regulatory bodies such as the US FAA or UK CAA. Increasingly, these reports are available on the World Wide Web. There are also now a number of national confidential reporting schemes, the largest and most well known being the ASRS operated in the USA by NASA. Other countries such as New Zealand (‘ICARUS’) and the UK (‘CHIRP’) have similar schemes. In addition, many industry organizations maintain their own error monitoring schemes (e.g. British Airways ‘BASIS’). The International Civil Aviation Organization (ICAO) has run an accident and incident reporting scheme (ADREP) that provides data on world-wide accidents and incidents involving aircraft over 2250 kg maximum certificated take-off mass. The ICAO (e.g. ICAO 1987) publishes an annual summary of these data.

Unfortunately, each source of information uses different schemes for coding and analysing human errors making comparisons between different schemes virtually impossible. In recognition of this problem, NASA recently (September 1998) held a
workshop on human factors taxonomy. The lack of common definitions and criteria for coding human errors is a significant problem in the global sharing and dissemination of information about human error in aviation. This may be one of the reasons why the proportion of accidents considered due to human error has remained stubbornly high for several decades (O’Hare 1999). As Reason (1997b: 117, emphases added) observes: ‘To learn the right lessons from the past, it is best to analyse several domain-related events using a common classificatory framework’.

A classificatory framework derived from Rasmussen’s (1982) analysis of errors made by nuclear power plant operators was successfully applied to aviation by O’Hare et al. (1994). This taxonomy has subsequently been applied by Zotov (1997) and Wiegmann and Shappell (1997) and has been shown to provide a reliable and comprehensive framework for the classification of the information processing aspects of human performance which occur in accidents. Wiegmann and Shappell reported a much higher level of coder reliability (k = 0.935) for the Rasmussen taxonomy than for one based on Reason’s (1990) framework (k = 0.777). In subsequent work (O’Hare et al. 1994, Zotov 1997), every action examined has been codable by the system. While this is one important part of solving the accident, as Reason (1997b) and others have pointed out, it is also necessary to identify the more general, higher-level conditions which generate specific lower-level failures.

The view that it is higher-level organizational conditions that primarily determine the safety of a high-risk system has been well received in aviation. An attempt to incorporate elements of this perspective into a classificatory framework for the analysis of aviation accidents has recently been reported by Shappell and Wiegmann (1997a). Their ‘Taxonomy of Unsafe Operations’ (TOS) has been adopted by the US Navy and Marine Corps for accident investigation and analysis.

The TOS taxonomy contains three levels. At the lowest level are the unsafe acts of individual operators. These are described using the distinctions provided by Reason (1990). At the next level are unsafe conditions that may lead to accidents. These include such things as adverse physiological states such as hypoxia, adverse mental conditions, such as overconfidence, and ‘crew resource mismanagement’. At the uppermost level are unsafe supervisory practices that include inadequate design, not adhering to rules and regulations, failing to note the potential adverse effects of life-changes and so forth.

The TOS represents a valuable attempt to develop a structured framework for the classification of human factors in accident investigation and analysis. By including a wide range of factors beyond the unsafe acts of the individual operator, the taxonomy is more in tune with current theoretical developments in safety analysis (e.g. Glendon and McKenna 1995, Reason 1997b) than earlier schemes. However, as Shappell and Wiegmann (1997a) acknowledge, much remains to be done to incorporate higher-level organizational conditions into the taxonomy. Since the taxonomy already has 26 subcategories, there is a danger that the eventual number of classificatory categories will become too unwieldy for practical use. This problem has been addressed in the most recent version of the taxonomy of unsafe operations known as HFACS (Shappell and Wiegmann 2000) which reorganizes the organizational factors into three groups.

The purpose of the present paper is to propose a revised theoretical model and associated classificatory framework for use in guiding the accident investigation process (‘solving the accident’) and in communicating the results of accident investigations. Such an effort is required to provide the accident investigator with
more practical tools and to avoid the present difficulties in pooling accident information from different sources due to the many different frameworks used to classify and organize the information.

3. The ‘Wheel of Misfortune’
The basic structure of the proposed model is based on Helmreich’s concentric spheres (figure 1) with the innermost sphere representing the actions of the front line personnel, the middle sphere representing local precipitating conditions, and the outermost sphere representing the global conditions generated by organizations. A detailed description of each sphere is provided below. Since a sphere is difficult to represent in two dimensions an alternative representation is proposed. A section through a sphere can be represented as a disc or wheel showing the concentric spheres of influence as described. Labelling this as a wheel preserves something of the three-dimensional nature of the original while facilitating presentation in two dimensions.

3.1. Local actions of the crew
As mentioned above, a system for describing the actions of the individual operator has been outlined by O’Hare et al. (1994). This ‘internal function’ taxonomy was originally developed by Rasmussen (1982) and is part of a broader taxonomic effort which includes the ‘Skill–Rule–Knowledge’ framework (Rasmussen 1983), which has become very well known in human factors and ergonomics in recent years (Sanderson and Harwood 1988). The ‘internal function’ taxonomy and level of control hierarchy are combined in the ‘decision ladder’ scheme as described by Rasmussen (1980).

The revised ‘internal function’ taxonomy, consisting of six steps between ‘Information’ and ‘Action’ provides a logical and complete description of the necessary steps that may be performed by the operator of any complex system. An algorithm for coding errors in terms of the six possible stages of ‘internal malfunction’ (figure 2) has been described by O’Hare et al. (1994).

Whether an operator will consciously consider all six steps depends on whether the level of control is skill-, rule- or knowledge-based. If the activity is a familiar and well-practised one, the operator moves directly between ‘Information’ and ‘Action’. The errors found at this level will be associated with failing to take in the necessary information, or with mis-executing the action in some way.

Any of the remaining four steps (Diagnosis, Goal Setting, Strategy Selection, Choice of Procedure) may be involved in either rule- or knowledge-based behaviours. *Diagnosis* errors (believing the system is in one state when in fact it is in another) and *Choice of Procedure* errors (having correctly perceived the system state, choosing an incorrect procedure) are most likely to be associated with rule-based control which is generally conceived of in ‘IF (X) THEN (Y)’ form, where X refers to system state and Y to a procedure. Such errors could also arise if the operator is functioning at a knowledge-based level and limited working memory resources curtail the gathering of information for diagnosis.

*Choice of an incorrect goal or strategy is most likely to arise when the operator is functioning in a controlled, effortful, knowledge-based manner. This would occur when conditions arise that are unfamiliar and for which no skill- or rule-based routines exist. This might be the case for most circumstances confronting a novice or neophyte, or for rather unusual circumstances confronting an expert. As Hart and Bortolussi (1984) noted, operator errors could be a cause as well as a consequence of
increased effort and workload, so the choice of an incorrect goal or strategy can lead to rapidly compounding effects. Designers are advised to try and prevent situations arising which may require the operator to function at a knowledge-based level (Hale et al. 1990) since behaviours are more likely to become idiosyncratic and unpredictable.

Given that errors at a knowledge-based level involving goal selection in particular, are likely to have more serious consequences than other kinds of errors (O’Hare et al. 1994), considerable effort is placed in training to develop skill- or rule-based routines for dealing with most of the conceivable problems an operator is likely to confront. Unfortunately, in any complex system it is impossible to predict every eventuality (Perrow 1984) so that operators must be prepared to operate at a knowledge-based level at some point.
The ‘internal malfunction’ taxonomy has been successfully applied to the analysis of New Zealand aviation accidents (O’Hare et al. 1994, Zotov 1997). The system has been found to provide a simple but comprehensive framework for describing the flight crew actions in accidents. It has been adopted as part of the New Zealand Civil Aviation Authority’s accident analysis system. Wiegmann and Shappell (1997) compared this taxonomy with a traditional four-stage model of information processing and with a model of unsafe acts (active failures) based on Reason (1990). These three taxonomies were used to code 289 types of pilot-causal factor cited in the US Naval Safety Center database. The ‘internal malfunction’ taxonomy was found to account for a slightly greater percentage of the reports (88.4%) than either the Reason model (84.3%) or the traditional information-processing model (80.6%).

A simplified ‘decision ladder’ approach is proposed as the central core of the present system. This is a simplified version of the original decision ladder described by Rasmussen (1976). Each ‘cognitive failure’ exhibited by the flight crew can be assigned to one of the six steps of internal malfunction and one of the three levels of cognitive control (figure 3). Thus, an event in which a reasonably experienced pilot slightly over-corrects for a crosswind resulting in scraping a wingtip on the ground would probably be coded as an action error at a skill-based level. A pilot whose only engine stops while the aircraft is 1000 feet above inhospitable terrain and who does not immediately set the aircraft up for a forced landing is committing a goal error at either a rule- or knowledge-based level, depending on the pilot’s expertise and training.

Reason’s (1990) distinctions between mistakes, slips, lapses, and violations are all encompassed by the above scheme (table 1). For example, violations are either rule-based (for ‘normal’ violations) or knowledge-based (for ‘exceptional’ violations).
errors usually involving choice of goal, but also including strategy and procedure errors. Slips are invariably action errors at a skill-based level. Coding them in this fashion thus provides more detail than simply coding them as slips or violations. Furthermore, this coding has greater practical utility than the Reason scheme used in the TOS. For example, finding that many accidents are codable as memory lapses does not immediately suggest an appropriate corrective strategy, other than practising memory improvement. By contrast, knowing that many fatal accidents arise because of faulty goal selection suggests that improved training in this particular aspect of aeronautical decision-making is urgently required.

The actions of the flightcrew are figuratively at the centre of the accident investigation and hence it is appropriate that they are literally at the centre of the proposed model. A description of the flightcrew actions in terms of the proposed model provides additional information about the flightcrew actions in non-evaluative terms. Furthermore, the Rasmussen model has been widely applied in
other systems (Goodstein et al. 1988) and hence the present system provides commonality with the analysis of errors in other contexts. Empirical research cited above, confirms that the model provides a complete, and slightly more comprehensive description than any other framework at this level. However, a description of the flightcrew actions is simply a description, and provides information as to ‘what’ happened, not ‘why’ it happened. To deal with this question, contributing influences at both the local and global levels need to be described.

3.2. Local conditions

There are numerous factors that might be critical in the breakdown of human performance in aviation and other complex systems. Most commonly these include extremes of weather (e.g. snow, icing, etc.) or other external conditions. Equally, extremes of internal states of the flightcrew (e.g. excessive fatigue, alcohol consumption, distraction, etc.) might precipitate performance breakdown. In most cases, research has shown that the relationship between these local conditions and performance breakdown is not one to one, but highly probabilistic. Thus, in Reason’s (1990) terms, the ‘type-token’ distinction becomes particularly important. That is to say, that any given type of condition (e.g. fatigue) can manifest itself in a variety of specific failures (tokens).

How should these local conditions best be categorized? The most obvious system would seem to be to differentiate those intrinsic to the flightcrew (such as fatigue or memory failure) from those extrinsic (weather, ATC, etc.). However, this becomes problematic, in that an ‘external’ factor such as flightcrew rostering inevitably has ‘internal’ consequences, such as fatigue. Shappell and Wiegmann (1997a) dichotomize the ‘preconditions for unsafe acts’ into ‘substandard conditions of the operator’ such as fatigue, illness, overconfidence, etc.’ and ‘substandard practices of the operator’, which include misjudgements such as over-exercising before a mission, poor aircrew coordination, and disregarding crew rest requirements. It is not particularly clear what criteria govern the allocation of factors to one or other of these groups. For example, would using a hay fever medication before a flight be categorized as a substandard practice or would it be categorized under the resulting substandard condition? The distinction in this and other cases appears somewhat arbitrary.

The theoretical approach adopted here is to view successful task performance as achievable so long as the resources available to complete the task are equal to, or

<table>
<thead>
<tr>
<th>Reason’s unsafe act</th>
<th>Rasmussen’s internal function</th>
<th>Rasmussen’s level of control</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLIP</td>
<td>action</td>
<td>skill-based</td>
</tr>
<tr>
<td>LAPSE</td>
<td>action</td>
<td>skill-based</td>
</tr>
<tr>
<td>MISTAKE</td>
<td>goal (strategy, procedure)</td>
<td>rule-based, knowledge-based</td>
</tr>
<tr>
<td>VIOLATION Normal</td>
<td>strategy/procedure</td>
<td>rule-based, knowledge-based</td>
</tr>
<tr>
<td>VIOLATION Exception</td>
<td>goal</td>
<td>rule-based, knowledge-based</td>
</tr>
</tbody>
</table>
exceed the demands of the task. This process of matching task demands to resource supply has been used to account for traditionally difficult human factors concepts such as stress and workload. For example, Cox (1978) proposed that stress is the consequence of an imbalance between perceived demands and the perceived capability to meet those demands. Wickens (1992: 390) set out a similar model for workload which ‘is fundamentally defined by this relationship between resource supply and task demand’. More generally ‘performance breakdowns are a function of mismatches between cognitive demands and the cognitive resources available to be applied to the task’ (Roth and Woods 1988: 55).

The classificatory framework proposed here is based on the ‘cognitive triad’ outlined by Roth and Woods (1988). Task performance in real-world environments is constrained by three sets of factors. The first is the inherent demands of the task environment. The demands in most of the domains of interest can be characterized at an abstract level by complexity, dynamism, tight coupling (Perrow 1984) and uncertainty and risk. The second set of factors consists of the resources supplied by the operators of the system. These include physiological and psychological capacities as well as skills and expertise. The third set of factors involves the representation of the task environment through which the operators act on the world. In concrete terms this essentially involves the interface between operators and the system comprising displays, controls and communication.

The pertinent local level factors can be classified into one of the three parts of the ‘cognitive triad’. Inherent task demands are represented by factors that increase the dynamism, complexity, coupling, or uncertainty of the task. In aviation or maritime environments, extremes of weather provide a good example of such a factor. All systems are designed to certain operational limits so ‘extreme’ can have a defined meaning in terms of closeness to these predetermined limits. A 20-knot wind may make the task of handling a microlight aircraft highly demanding but have no effect on the task of handling a large passenger airliner. Mission requirements (e.g. low-level flight) can easily increase the inherent demands of a task. Other factors may also increase the inherent difficulty of a task. For example, scheduling requirements to turn around a ship or plane in an unrealistically short period affects several of the factors that determine inherent task demand.

The resources which front line personnel have available to perform their tasks can be categorized under two headings (figure 4): capacities, both physical and psychological, and skills. Physical resources include physiological capacities and states of readiness as well as basic sensory capabilities and limitations. Psychological resources include attentional and decision-making abilities, motivation, and attitudes. Some of these resources are the subject of selection procedures (e.g. sensory and psychological abilities) and others are the subject of training programmes (skills). These include the coordinated use of individual resources as suggested by Crew Resource Management (CRM) principles and procedures.

The operator’s resources may be affected by many conditions both stable and unstable. Both sensory and psychological capacities may be affected by duty times and scheduling of rest periods. Motivational and attitudinal difficulties might be generated in the wake of company mergers and the handling of seniority issues.

The manner in which the task is represented to the operator (i.e. the human – machine interface) can significantly affect performance. The propensity for misjudgements induced by the three-pointer altimeter is one well-known example in aviation (Rolfe 1969). Poorly designed checklists or flight manuals might be
considered 'error-inducing', as might the design of certain cockpit systems and displays. For example, one of the local precipitating conditions in the crash of a B737-400 at Kegworth, UK, in 1989 was the newly designed electronic engine vibration indicators that replaced the relatively large inside-scale pointer with a small electronic pointer outside the scale (Air Accidents Investigation Branch 1990).

In light aircraft the design and placement of fuel selectors has been an error-inducing factor in numerous crashes (O'Hare and Roscoe 1990). Similarly, the design of conventional flight displays such as the artificial horizon has been shown to lead to significantly higher error rates than alternative designs (Roscoe 1980). A significant factor in weather-related crashes in general aviation may be the absence of direct representations of weather systems. The use of graphical aids for weather display in the cockpit is now under investigation (Lind et al. 1994). Failures of information transfer are a particularly common problem in complex technological systems (Billings and Reynard 1984). These include communication amongst front-line operators or between operators and other personnel (e.g. pilots and air traffic controllers).

The distinction between the demands placed on personnel (both those inherent to the task and those due to the representation of the task world) and the resources they bring to bear can be mapped onto the fundamental distinction between the activities of design (involving both equipment and procedures) on the one hand, and selection and training on the other: 'the former serves to reduce the need for the latter; the latter completes the job left undone by the former' (Roscoe 1980: 3). Design and training are the fundamental cornerstones of any systems operation.

As noted above, deficiencies in design manifest themselves in terms of excessive or misplaced demands placed on the human operator. For example, there has been much discussion about the demands placed on aircrew by modern automated flight
management systems (e.g. Sarter and Woods 1997). Deficiencies in selection and training manifest themselves in the behaviour of operators who are unable to supply the necessary resources for successful system performance. It is somewhat misleading to describe either of these activities as ‘defences’ since they are an integral part of normal system operation.

3.3. Global contexts
The outermost sphere of the model represents the global context within which the task activity takes place. These organizational processes have received much attention since the ‘Reason model’ became a popular framework in thinking about hazardous technologies. This organizational approach to safety has been expressed as follows: ‘human error does not take place in a vacuum, but within the context of organizations which either resist or foster it... no matter how humans excel in their individual or small team’s performance, they can never be better than the system which bounds them’ (Maurino et al. 1995: xi).

The initial emphasis of the model was on the ‘fallible decisions’ made by high level decision-makers within organizations and the potential for the consequences of such decisions to lie dormant within systems. The metaphor of the organic pathogen provided a vivid reminder of the potential power of ‘latent failures’ to wreak havoc on the system in the future. More recently, to avoid the concern that the focus of blame was simply being shifted from flightcrew to boardroom, the pejorative term ‘latent failures’ has been replaced by the more neutral ‘latent conditions’. The emphasis is now (Maurino et al. 1995: 9) more clearly placed on organizational systems which can make the potentially negative consequences of high-level decisions ‘visible to those who manage and operate the system’.

Reason (1990) earlier pointed out that even the best run organization cannot be expected to eliminate all potential workplace hazards, ‘But they can anticipate the possibility, if not the particular form of occurrence and provide adequate defences against their unsafe consequences’ (Reason 1990: 206). More recently (Reason 1997b: 113, emphases added) explicitly stated that the most important force driving safety efforts at the organizational level is cognisance: ‘Neither commitment nor competence will suffice unless the organization has a correct awareness—or cognisance—of the dangers that threaten its operations’.

The most important distinction to be made at the global/organizational level of analysis is therefore whether the hazards and risks were correctly recognized, or whether there was a lack of knowledge that the hazards existed, and/or a failure to appreciate the risks involved. In the first case, efforts to control the hazards and risks may have been inadequate, or the operator may simply have made an unwise decision to accept a known risk. In the second case, the causes of the failures could range from the breakdown of specific internal feedback mechanisms (e.g. hazard reporting schemes) to a pervasive company philosophy that denies the possibility or significance of organizational practices on accidents. Perhaps the most glaring recent example of this (see below for further detail) was the company that operated the Herald of Free Enterprise that sank on leaving port in 1987 (Department of Transport 1987).

This distinction is very similar to that made by Shappell and Wiegmann (1997a) who distinguish between ‘unforeseen unsafe supervision’ and ‘known unsafe supervision’ at the ‘line management deficiencies’ level of Reason’s (1990) model.
In the ‘unforeseen unsafe supervision’ category Shappell and Wiegmann include the somewhat nebulous category ‘loss of supervisory situational awareness’ along with inadequate documentation and procedures. The ‘known unsafe supervision’ includes lack of guidance, failing to correct known problems, intentional violations, and improper work scheduling. While these are all important issues, the groupings have something of an *ad hoc* quality.

The present scheme attempts to simplify this burgeoning complexity by proposing just three taxonomic categories under each main heading (table 2). The proposed categories are based on the analysis of aviation organizations described by Degani and Wiener (1994). At the head of Degani and Wiener’s framework is the organization’s philosophy: ‘By philosophy we mean that the airline management determines an over-arching view of how they will conduct the business of the airline’ (1994: 49). This can be determined in part by the top-level decision-makers, but is also shaped by, and reflected in, the corporate culture of an organization. Since, as noted by Zotov (1996), flight crew operate within the influence of many organizations (e.g. airline, manufacturer, regulator, and air traffic control provider), the philosophies of any of these organizations might be relevant.

At the next level are policies, which ‘are broad specifications of the manner in which management expects operations to be performed’ (Degani and Wiener 1994: 50). Policies lead to the development of procedures, which are specifications for a task or series of tasks to accomplish a predetermined goal. Standard Operating Procedures (SOP) are the cornerstone of commercial aviation, leading to a high degree of consistency and uniformity in performance.

These three distinctions (Philosophy, Policy and Procedures) at the organizational level of analysis can be equated with the concepts of Goal, Strategy and Procedure, used at the individual level. Thus, to return to the central metaphor of the title, the conceptual scheme comes full-circle. Section 4 shows how the proposed conceptual scheme can be used as an over-arching classificatory framework for representing the outcomes of accident investigation and analysis into human performance breakdowns in complex systems. While the primary function of the present scheme is to

<table>
<thead>
<tr>
<th>Philosophy</th>
<th>Policy</th>
<th>Procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazards unrecognized (risks not known or correctly appraised)</td>
<td>Hazards forseen (risks anticipated but response not adequate)</td>
<td>Hazards unrecognized (risks not known or correctly appraised)</td>
</tr>
<tr>
<td>Philosophy</td>
<td>Policy</td>
<td>Procedures</td>
</tr>
<tr>
<td>Safety not source of corporate pride</td>
<td>Safety seen as source of corporate pride</td>
<td>Documentation inadequate</td>
</tr>
<tr>
<td>Regulatory standards seen as maxima</td>
<td>Regulatory standards seen as minima</td>
<td>Loop-hole in defences</td>
</tr>
<tr>
<td>Policy</td>
<td>Policy</td>
<td>Procedures conflict with one another or with organizational policy</td>
</tr>
<tr>
<td>Internal monitoring schemes inadequate (e.g. employee concerns not</td>
<td>Known deficiencies (e.g. equipment, maintenance) not addressed</td>
<td>Procedures</td>
</tr>
<tr>
<td>communicated upwards)</td>
<td>Defences not adequately monitored</td>
<td>Procedures</td>
</tr>
<tr>
<td>Insufficient resources allocated to safety</td>
<td>Defences compromised by other policies (e.g. adversarial employee</td>
<td>Procedures conflict with one another or with organizational policy</td>
</tr>
<tr>
<td>Safety managers insufficiently trained or equipped</td>
<td>relations, incentive systems, performance monitoring)</td>
<td>Procedures</td>
</tr>
<tr>
<td>Reliance on other organization’s criteria (e.g. equipment manufacturer)</td>
<td></td>
<td>Procedures</td>
</tr>
<tr>
<td>Procedures</td>
<td></td>
<td>Procedures</td>
</tr>
<tr>
<td>No written procedures</td>
<td></td>
<td>Procedures</td>
</tr>
</tbody>
</table>
provide a means of integrating the results of analyses carried out within different frameworks, it can also serve as a stimulus to original investigation by suggesting the appropriate questions to address at each of the three conceptual levels.

4. Using the taxonomy to represent the findings of accident investigations
The purpose of this discussion is to illustrate the value of the Wheel of Misfortune taxonomy as a general classificatory scheme for summarizing the outcomes of accident investigations. The two examples below are taken from different domains (marine and air transport). The analyses presented here are not intended to be exhaustive but are illustrative of the value that may be obtained by representing accident reports in a common classificatory framework.

The first example used to illustrate the application of the taxonomy is the shipping disaster involving the sinking of a cross-channel ferry, The Herald of Free Enterprise off the coast of Belgium in 1987 (Department of Transport 1987). The basic facts of the disaster are well known and have never been disputed. The immediate cause of the disaster was the failure of one of the crewmembers to close the bow door before the ferry left dock. He had in fact gone to his bunk and fallen asleep. When the ferry reached open sea and accelerated, water entered through the open doors, flooding the open car deck, causing the ferry to capsize.

At a local level, the critical actions in this disaster were those of the crewmember in charge of the bow doors, his immediate supervisor and the captain of the ship. In Reason’s (1990) terms, the crew member, in choosing to catch up with sleep rather than to perform his allocated tasks, committed an exceptional violation, which would be coded as a goal error at the knowledge-based level of reasoning (table 2). The same could be said for the immediate supervisor who was apparently aware that there was no one present to close the doors but chose not to do so himself as it was not part of his prescribed duties. The captain carried out his duties as normal and thus committed no error at this level of analysis.

Some precipitating factors can be detected in this tale. The necessary human resources were not forthcoming to achieve satisfactory performance. The failure of the crewmember at the bow was associated with tiredness and fatigue (he was soundly asleep before the ship even left port). One may wonder about his motivation and attitude towards the job and whether his training had equipped him with the appropriate skills to carry out his tasks safely and efficiently. Although awake, his immediate supervisor displayed little interest in ensuring that the tasks essential for the safe operation of the vessel were being performed. There was a glaring omission on the design side that contributed to this disaster. This was the absence of any warning indication on the bridge to indicate that the bow doors were open. The Captain’s ability to ensure that the ship was safely configured before proceeding to sea was thus completely undermined.

The inquiry into the disaster focussed attention on the higher-level management of the company that operated the ferry. The critical question is whether the shipping company were fully cognisant of the danger posed by a ferry leaving port with its bow doors open to the sea. Almost unbelievably, the answer seems to have been that they were not. This is in spite of the fact that the company had received at least five reports of this happening since 1983.

There were significant deficiencies at each of the three levels of corporate philosophy, policy and procedures. The corporate philosophy was based entirely on commercial performance. Individuals were to be pressurized into performing their
tasks as rapidly as possible. Suggestions from the ships’ captains regarding the installation of a bow door-warning signal were treated with sarcasm and contempt. The most explicit statement of company policy (the ‘Ship’s Standing Orders’) made no reference to the closing of bow doors before sailing. There was therefore no policy covering the closure of the doors, or any system for ensuring that this critical task was performed regardless of the behaviour of any particular individual. The task of preparing the vessels was significantly different from one port to another, yet there was no policy regarding this matter. Company policy regarding failures was that the absence of any report would be taken to indicate that there were no deficiencies. These widespread failures of company policy were reflected in the absence of specific procedures to deal with the bow doors. There were no written procedures to ensure the task was done, or that it was checked.

Even from this abbreviated discussion of the disaster, it is clear that there is strong support for the thesis proposed by Reason (1990, 1997b) that the safety critical element in any complex system is the performance of higher level management. The application of the taxonomy derived from the Wheel of Misfortune model summarizes the essential deficiencies at each of three organizing levels. Other accidents can be represented in the same way, thus allowing the commonalities between apparently disparate events in different domains to emerge and be compared. Table 4 summarizes the factors discussed above from the Herald of Free Enterprise disaster and compares them to a brief summary of factors present in the Kegworth air disaster (Air Accidents Investigation Branch 1990).

The details of the Kegworth disaster will not be described here. Suffice it to say that the initiating event was the failure of a fan blade in one of the aircraft’s two engines. Following this, the flight crew shut down the undamaged (no. 1/port) engine. The aircraft was unable to reach the diversion airfield, crashing just short of the threshold. Details of this crash are available in a number of primary (Air Accidents Investigation Branch 1990) and secondary sources (e.g. Reason 1990, Job 1992, Smith 1992, Johnson et al. 1995). Table 3 shows the critical actions of the crew as a rule-based misdiagnosis by the captain (incorrectly attributing the observed symptoms to a failure of the no. 2/ starboard engine), and a skill-based information error (misreading the electronic engine vibration display) by the first officer. The failure of the cabin crew to communicate information, which was seen in the cabin concerning the state of the engines, to the flight deck was a goal error. This was probably due to the cabin crew following the common belief that ‘the pilots must know what they are doing’.

The task of the flightdeck crew was undoubtedly made more difficult by the design of the electronic engine vibration display which used a small cursor on the outside of the display instead of the large central pointer used on the old electromechanical displays. There was nothing on the display to emphasize the fact that the display was reading at its maximum. The location of the display also contributed to the ease of misreading the indications. The resources brought to bear by both the flightdeck and the cabin crew were limited by deficiencies in their training. The airline provided a 1-day audio-visual training course for conversion from the B737-300 with its electromechanical flight displays to the B737-400 with an electronic ‘glass-cockpit’. The airline did not have a ‘glass-cockpit’ flight simulator, so the crews had no opportunity to acquire ‘hands-on’ experience with the systems used on the B737-400. The airline did not provide joint CRM (crew resource management) training to enable the cabin and flightdeck crews to operate together in the most effective manner.
Table 3. Comparison between the *Herald of Free Enterprise* and Kegworth disasters.

<table>
<thead>
<tr>
<th>Level</th>
<th>Herald of Free Enterprise</th>
<th>Kegworth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local actions</td>
<td>(1) Bow-Door Operator</td>
<td>(1) Captain</td>
</tr>
<tr>
<td></td>
<td>Goal error/knowledge based</td>
<td>Diagnosis error/rule based</td>
</tr>
<tr>
<td></td>
<td>(2) Immediate Supervisor</td>
<td>(2) First Officer</td>
</tr>
<tr>
<td></td>
<td>Goal error/knowledge based</td>
<td>Information error/skill based</td>
</tr>
<tr>
<td></td>
<td>(3) Captain</td>
<td>(3) Cabin Crew</td>
</tr>
<tr>
<td></td>
<td>No error</td>
<td>Goal error/rule based</td>
</tr>
<tr>
<td>Local conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task demands</td>
<td>Crew scheduling/rostering</td>
<td>Engine failure</td>
</tr>
<tr>
<td></td>
<td>rapid turn around times</td>
<td></td>
</tr>
<tr>
<td>Interface</td>
<td>No representation of bow door status</td>
<td>Design of electronic engine vibration indicator</td>
</tr>
<tr>
<td></td>
<td>(absence of door open warning on bridge)</td>
<td>Alignment of gauges with power levers</td>
</tr>
<tr>
<td>Resources/training</td>
<td>Lack of hazard awareness</td>
<td>One day audio visual conversion course</td>
</tr>
<tr>
<td></td>
<td>Narrow views of duties and responsibilities</td>
<td>No B737-400 simulator</td>
</tr>
<tr>
<td></td>
<td>Fatigue</td>
<td>No training in high engine vibration</td>
</tr>
<tr>
<td>Global context (operator)–hazards not recognized</td>
<td>Philosophy</td>
<td>No coordinated CRM involving cabin crew</td>
</tr>
<tr>
<td></td>
<td>Entirely operations centred</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Safety indicated by absence of warnings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Procedures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not explicit</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No redundancy</td>
<td></td>
</tr>
</tbody>
</table>

The views expressed here are those of the author and do not represent the views of any other person or organization.
As was typical of accident investigation at the time, the official report provides little insight into the organizational factors that may have generated the failures in design and training noted above. Smith (1992) reports the application of two broader conceptual frameworks to the events that occurred in the Kegworth crash. Both the DEPOSE framework (Perrow 1984) and his own ‘6C’ system provide re-descriptions of the disaster without moving far from the obvious locus of events in the actions of the flight crew. One of Smith’s 6Cs is ‘culture’, which in this case is related to both the aeroplane and engine manufacturers as well as the airline’s ‘communication and training practices’.

5. Summary and conclusions

There has been considerable progress in the past decade in the field of human error, particularly in terms of conceptualizing the broader constraints that shape individual behaviours. It is now commonly recognized that there is much more to accident investigation than describing the actions of the nearest individuals in the last moments before the event. Despite a number of attempts to bridge the gulf between theory and practice (e.g. Reason 1997b), including the recently developed TOS (Shappell and Wiegmann 1997a), there remains a need for a simple, but comprehensive, analytical framework for integrating the findings from accident investigations in various domains and for guiding the accident investigation process towards the common constraints on performance in any system.

The Wheel of Misfortune framework serves three potentially valuable functions in this regard. First, as a heuristic model, the concentric spheres-within-spheres representation is a better approximation of the reality of accident causation than any linear sequence of factors. Instead of focussing entirely on one concept such as ‘defences’ (Reason 1997b), the present model represents the safety health of a system in terms of the over-arching consideration of whether the organization was ‘cognisant’ or aware of a specific hazard and how this awareness was reflected in the philosophy, policies and procedural practices of the organization.

An alternative to Reason’s (1990) ‘Swiss Cheese’ metaphor for system failure is illustrated in figure 5. In the Wheel of Misfortune model, the outer shell of the model primarily determines the strength of a system. A system that has clear and accurate hazard awareness and which has developed a set of appropriate philosophy, policies and procedures, will be prepared to deflect many of the slings and arrows of potential misfortune. A system that does not have proper cognisance of the hazards it faces will be somewhat porous to these outside influences. Lacking an appropriate set of philosophy, policies and procedures, the delicate balance between demands and resources will eventually fail, precipitating inappropriate actions by front line personnel. Some failures may of course arise from the inner levels—poorly designed interfaces, inadequately trained or fatigued crew—or exceptional operating conditions may trigger the active failures observed at the inner level. More often than not, however, it is the same front-line personnel who end up averting disaster by their actions. The deflected arrows in figure 5 represent these influences.

Second, as a practical investigative tool, the Wheel of Misfortune directs the attention of the investigator to specific questions within the three layers of concern: the local actions of directly involved operators; the immediate realities of the operational environment; and the general influences of organizational functioning.
Had the investigators of the Kegworth crash followed this model, for example, we would be much better informed as to the precise manner in which the organizational practices of the airline, aircraft manufacturer, engine manufacturer, and government regulator, came to contribute to this accident. The present framework is also consistent with recently developed airline industry approaches such as Boeing’s procedural event analysis tool (PEAT) described by Graeber (1999) and the HFACS classification system (Shappell and Wiegmann 2000). A practical application of the framework for coding incidents reported to the New Zealand confidential aviation reporting system has been developed.

Finally, as an over-arching framework, the Wheel of Misfortune is expressed in terms of general processes that are quite independent of functioning within any
specific domain. It should therefore be possible to integrate the information represented in other models within this framework. The conceptual basis for the framework is firmly grounded in recent theoretical developments in cognitive engineering and cognitive ergonomics. Almost all previous models have been tied to the particularities of a specific domain thus making it impossible to see commonalities between events involving, say, aircraft, trains, and drilling platforms. The ‘TOS’/HFACS (Shappell and Wiegmann 1997a, 2000) is one example which avoids this particular problem. The present independently developed framework is fully consistent with this taxonomy, but involves a representation of the necessary processes at a slightly higher level of abstraction, thus achieving both greater comprehensiveness and greater parsimony.

Acknowledgements

The author is indebted to Dmitri Zotov for many helpful observations on the manuscript.

References


