Paper presented at Sixth International Australian Aviation Psychology Symposium
Sydney, December 1-5, 2003

Error 'molecules' and their implications for system safety

ALAN HOBBS
SJSU Foundation/NASA Ames
Mail Stop 262-4
Moffett Field, CA, 94035

Unanticipated human behaviour, in the form of errors and violations, has always been a major threat to complex safety-critical systems (Perrow, 1984). Human error featured significantly in early aircraft accidents (Wilmer, 1979) and 19th century rail accidents (Rolt, 1998) and it continues to present a risk in fields such as medicine, transport and power generation (Redmill & Rajan, 1997). The threat of human error originates not only with system operators such as pilots or drivers. There is an increasing recognition that maintenance errors are a major cause of losses in a wide range of industries (Reason & Hobbs, 2003).

Despite the importance of human reliability to system performance, until recently, the failure modes of engineered devices received more attention than the failure modes of human operators. In 1941 Heinrich, the insurance engineer who produced the influential domino model of accident causation, called for psychologists to study accidents and industrial errors. Yet for much of the 20th century, psychologists left the study of error to engineers or linguists. Spearman (1928) noted that 'Psychologists positively decline to investigate error. They regard it as not being their job' (p. 30). Until the 1970s, the most widely used error models described the observable form of the error, such as omission or commission (Miller and Swain, 1987) but stopped short of considering the mental processes that produced the behaviour. Without a model of error grounded in cognitive psychology, the human actions that lead to accidents can appear bizarre or random, presenting risks that are mysterious and largely unmanageable. We may see that a person pushed a wrong button or pulled a lever at the wrong time, but we are at a loss to explain why they did it, or whether somebody else will do the same in the future.

Table 1. Five types of unsafe act (after Reason, 1990).

Level of cognitive control	Unsafe act
Automatic	Slip
	Lapse
	Rule-based mistake
Conscious	Knowledge-based mistake
	Violation

In the 1970s, there was a renewal of interest in the errors of everyday life and industry. Researchers including Reason (1979), Rasmussen (1982), Rouse and Rouse (1983) and Norman (1988) revealed that operator errors took forms that could be categorized with inferences about the operator's mental processes.

The error model developed by Reason (1990) is perhaps the most widely used error framework in the study of accidents. According to Reason, skill-based, or unintended unsafe acts occur when task

performance is being directed by automatic habit routines. These errors take the form of slips and lapses. Unsafe acts involving intended actions can occur when the person is paying conscious attention to a task. These errors can be categorised into rule-based or knowledge-based mistakes, and violations. The term 'violation' is used here to refer to errors that involve intentional deviations from procedures, such as operational shortcuts, workarounds, or conscious risk-taking, but not sabotage or other malicious acts. More will be said about the varieties of error in later pages.

In most accident scenarios, the unsafe actions of operators are the final events that trigger the accident, although the circumstances surrounding the events may have originated deep within the organization. Reason (1997) distinguishes between *active failures*, and *latent failures*. Active failures are the immediate unsafe acts that directly compromise the safety of a system. Latent failures are pre-existing weaknesses that may have been present in the system long before the accident occurred. While acknowledging the importance of latent failures, this paper focuses on the active failures that immediately precipitate accidents. The message outlined in the following pages is that an understanding of the cognitive origins of error, and the manner in which different forms of error combine in safety-critical settings can provide valuable insights into the safety health of complex technological systems.

Combination patterns of unsafe acts

In most well-defended systems more than one unsafe act must occur before a hazardous situation will result. A useful metaphor is to consider individual errors as *atoms* that pose the greatest dangers when they combine to form *error molecules*. Yet relatively little attention has been given to how errors combine to create hazards or reduce system reliability.

Williamson and Feyer (1990) examined records of fatal workplace accidents and noted that fatalities often resulted from chains of error. For example, a rule-based or knowledge-based error often started a sequence of events that culminated in a skill-based error with fatal consequences.

Violations are a persistent feature of aviation maintenance activities. McDonald, Corrigan, Daly and Cromie (2000) interviewed European airline maintenance personnel about their most recent maintenance tasks and found that 34% of routine maintenance tasks at airlines were performed contrary to procedures. Similarly, eighty per cent of aircraft mechanics surveyed by Hobbs and Williamson (2000) reported that they had deviated from maintenance procedures at least once in the previous year, with nearly 10% reporting that they did so 'often' or 'very often'.

Violations are not necessarily dangerous in isolation, but can lead to accidents by creating opportunities for errors to cause harm (Lawton, 1998; Parker, Reason, Manstead and Stradling, 1995; Reason, Parker and Lawton, 1998). The authors cited above have generally envisioned violations preceding errors (such as a driver first speeding, and then losing control on a wet road), but violations can also be dangerous when they occur after an error. An example is when a mechanic forgets to complete a task and then decides to omit a check that would have uncovered the omission.

Hobbs and Williamson (2003) gathered over 600 accounts of aircraft maintenance incidents, including information on the sequence of unsafe acts that led to the incident. An analysis of these incidents has indicated that when two unsafe acts were reported to have led to an incident, the two acts tended to be drawn from different error categories. A common 'error molecule' was a violation + skill-based error, where the violation preceded a slip or a memory lapse. In the following example, a violation was followed by a memory lapse:

> In order to make a procedure more convenient, a mechanic disconnected an electrical cannon plug inside the aircraft, even though this action was not in accordance with the maintenance manual. The mechanic did not document what he had done, but left a moveable work stand near where he had disconnected the plug as a reminder that he needed to return and reconnect it. During the shift however, someone moved the work stand, and the mechanic only remembered that he had left the wiring disconnected when he arrived home at the end of the shift. A discrete phone call to a colleague rectified the situation (Incident reported to Hobbs and Williamson).

In other cases, the 'error molecules' comprised memory lapse + violation, and memory lapse + failure to perceive. The following example illustrates the memory lapse + violation combination:

> A mechanic forgot to fill an engine oil tank. Due to a lack of available time, he then decided not to test run the engine on the ground before the aircraft departed. As a result, engine turbine bearings failed, requiring extensive repair work (Incident reported to Hobbs and Williamson).

As can be seen from these cases, the risks associated with each unsafe act are amplified when they occur in sequence. In most cases, the violation would have had no ill effects had it not been paired with an error, and the error would have been captured, or would not have eventuated had the violation not occurred.

Implications for system safety

The combination patterns of unsafe acts in safety-critical environments have some important implications for safety management. Most notably, the resilience of a system can be expressed in terms of its tolerance to errors. Figure 1 illustrates the interaction between the number and type of unsafe act required to create a hazard in a safety-critical system. The horizontal dimension of the safety space represents the level of cognitive control associated with the unsafe act. At one extreme are automatic, unintended actions, at the other extreme are actions resulting from conscious decision making. The vertical dimension of the figure represents the number of unsafe acts that must occur before an accident or other critical system failure will occur.

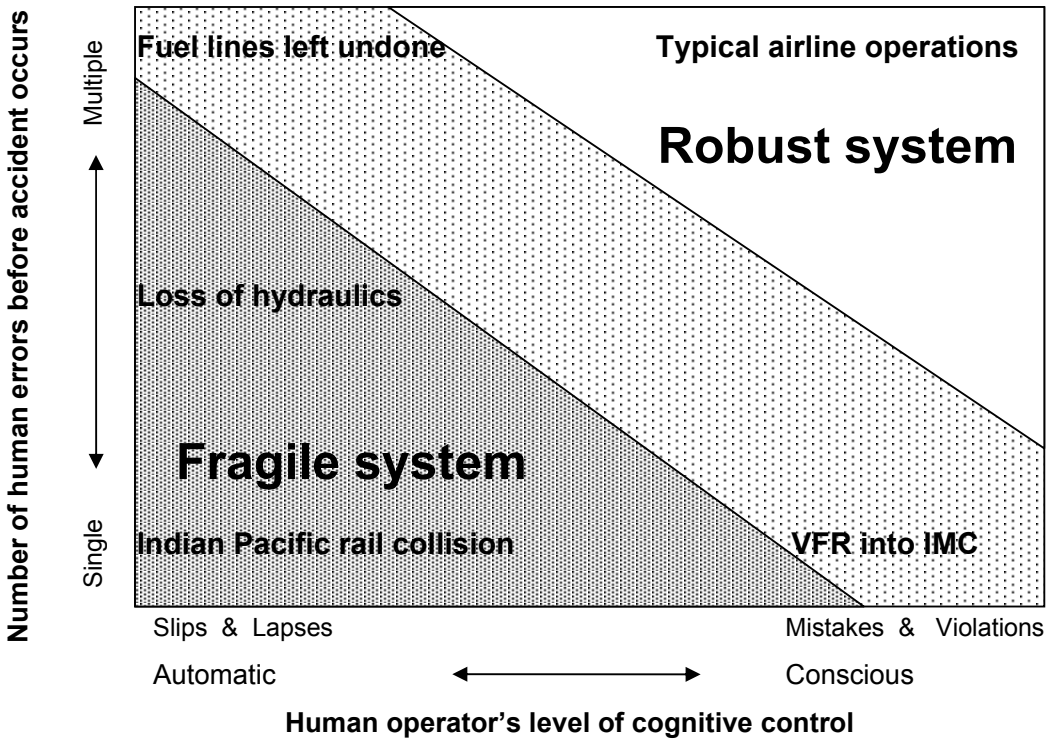

Figure 1. A two-dimensional 'safety space' representing system safety in terms of the number and type of human errors a system can tolerate before hazards are released.

Systems that are vulnerable to single, conscious errors. Systems that can fail due to a single operator mistake or violation are represented at the lower right corner of Figure 1. As noted earlier, mistakes are failures of conscious decision making. The error-maker formulates a plan of action based on the available information, yet the planned actions are inadequate to achieve the goal that the person had in mind. An example of a single-mistake accident is a Visual Flight Rules (VFR) pilot deciding to press-on into deteriorating Instrument Meteorological Conditions (IMC). Systems that can be breached by a single mistake have a higher level of risk than systems that require multiple errors before a hazard is created.

Systems that can fail due to a single violation are also represented in the lower right sector of the safety space. Violations, being conscious decisions, are likely to be preceded by an informal risk assessment, however imperfect. Battmann and Klumb (1993) consider that people evaluate the potential costs and benefits of a course of action before they decide to contravene a procedure. The types of violations that can unleash an accident without pairing with other unsafe acts are often self-evident hazards, much like rocks projecting above the surface of a body of water. Such violations are relatively rare in safety-critical systems, largely because their risks, being available to conscious awareness, are likely to provoke an aversive reaction. Yet in rare cases, single egregious violations lead to accidents, as the following example illustrates.

> In 1994, while practicing for an air show, an experienced pilot attempted to perform a low speed 360-degree turn in a B-52 at around 250 feet above ground level. Approximately three quarters of the way through the turn, the bomber banked past 90 degrees, stalled, clipped a power line with the left wing and crashed. There were no survivors. The captain of the aircraft had a history of reckless flying (Kern, 1999).

The reasons for such reckless violations are beyond the scope of this paper, however Lawton and Parker (1998) present evidence suggesting that unstable extraverts may be most likely to engage in such risk-taking behaviour.

Systems that are vulnerable to multiple errors, including conscious acts. Most aviation systems lie at the top right section of the safety space shown in Figure 1. These robust systems must be confronted with a 'molecule' of multiple unsafe acts, including mistakes or violations, before a human performance-related accident can occur. An analysis of 37 major accidents involving US air carriers found an average of 8.2 errors per accident (Orlady and Orlady, 1999). The world's worst airline disaster, the much-cited collision at Tenerife, resulted from a cluster of errors that included mistakes made by the pilots of the two aircraft and the air traffic controller. Although as previously stated, informed operators will generally avoid violations that bring obvious and immediate risks, they may be less concerned about violations that serve as 'amplifiers' of errors rather than primary hazards in their own right. This is because the risks of such violations, being indirect, require an effort to imagine and hence provoke less concern. The case study presented earlier, in which aircraft electrical wiring was disconnected, illustrates a violation that may have appeared benign to the mechanic involved.

Systems that are vulnerable to multiple 'automatic' errors. The top left region of the safety space represents systems that can be compromised by a combination of skill-based errors. Skill-based slips and lapses are fragments of behaviour that are performed unintentionally, or omitted unintentionally. An important difference between such errors and unsafe acts involving controlled processing is that skill-based errors arise in an almost random, non-systematic way (Rasmussen, 1982). Although the precise timing of an individual skill-based error is impossible to predict, the overall background rate at which such errors occur can be estimated. For example, the probability of a wrong switch being activated, has been estimated at .001, and the probability of a wrong number being keyed into a push-button phone is around .03 (Kirwan, 1994). Each performance of a task provides an opportunity for such errors to occur. Although performance-shaping factors can change the overall probability of such errors, the probability of two independent skill-based errors occurring on the same task by

chance is usually low. Nevertheless, accident and incident reports do contain occasional cases of skill-based error molecules, as the following example illustrates:

> An aircraft mechanic forgot to nip up a fuel line, leaving the connection finger tight. At the end of the task, the mechanic also forgot to perform a leak check that could have revealed his lapse. Although both of these errors were memory lapses, neither error was the cause of the other. The aircraft subsequently experienced a fuel leak. (Incident reported to Hobbs and Williamson).

Systems that are vulnerable to single 'automatic' errors. The lower left region of Figure 1 represents cases in which a single skill-based error is sufficient to breach the integrity of a system. Although skill-based behaviour is usually more reliable than behaviour under conscious control, human operators generate skill-based errors at a low but persistent rate. Rather like a game of roulette, given enough opportunities, skill-based errors are virtually certain to occur. While it may be acceptable for recreational activities such as rock climbing, or even personal transportation to be in the lower left sector of the safety space, public transportation systems should never operate in this region. Nevertheless, complex transport systems can be found in this sector of the safety space as the following examples illustrate:

> In August 1999, a freight train was standing on a siding to enable the Indian Pacific passenger train to pass on the main line. As the Indian Pacific approached, one of the drivers from the freight train inadvertently pressed a button that activated the electrically operated points. He immediately recognized his error, but had no opportunity to undo his action. As a result, the Indian Pacific was diverted from the main line and collided with the stationary freight train. (Western Australian Department of Transport, 1999).

> Reaching for fuel cross-feed valves on the overhead panel of an A340, the captain inadvertently shut off all four engine-driven hydraulic pumps. As a result of the loss of hydraulic pressure, the autopilot disengaged, and the aircraft pitched up, then down, before level flight was regained. Several passengers were injured during the event. The buttons for the hydraulic pumps were located adjacent to the cross-feed buttons, the two sets of controls looked similar and were both activated by the same pushbutton switching action. No guards were fitted to the hydraulic switches. (Which Switch? 1998).

Safety-critical systems that cannot tolerate the incapacitation of an operator can also be located in the lower left sector of the safety space. In October 2003, a New York ferry collided with a dock, killing ten passengers. Early reports suggested that the ferry captain may have become unconscious in the moments leading up to the accident. Despite a policy requiring two ferry operators to be present during docking, it appears that only one person was at the helm as the ferry approached the dock (Kennedy, McIntire, Rashbaum & O'Donnell, 2003).

Although the safety space concept illustrated in Figure 1 has been developed with transport systems in mind, the principles are relevant to other critical systems. A radiation therapy machine that delivered fatal doses of radiation due to a skill-based slip by a technician (Casey, 1998) or computer systems that can 'crash' due to a single keyboard error (Mellor, 1994) are each examples of fragile systems. Skill-based errors such as the following can cause disruptions costing millions of dollars in the space of minutes:

> London – An inexperienced computer operator pressed the wrong key on a terminal in early December, causing chaos at the London Stock Exchange. The error at stockbrokers Greenwell Montagu led to systems staff working through the night in an attempt to cure the problem (Norman, 1988, p.105).

Hardening systems against errors

The essential factor that determines where a system lies on the safety space is the presence of effective defences that capture errors, quarantine their effects, or lessen the chances of the action occurring. For example, since the 19th century, many rail systems have been equipped with systems to prevent points from being changed in front of an oncoming train, yet such a defence was not present in the case of the train collision referred to earlier. In airline operations, the practice of servicing redundant systems independently reduces the chances of multiple maintenance errors, and thus increases the robustness of the system.

Two points should be noted at this stage. First, accidents can occur to systems that lie at all places on the safety space, including seemingly robust systems. Second, different elements of a system may occupy different sectors of the safety space. For example, flight operations and maintenance may each have different levels of resilience. Although the central activities involved in system operation may be well-defended against most single errors, threats may come from unexpected quarters, such as the actions of aircraft cleaners, or spare part shipping practices (Walters and Sumwalt, 2000).

Evaluating system resilience. In many cases, the fragility of a system only becomes apparent after an accident has occurred. Traditional risk assessments often involve multiplying the chances of independent events to arrive at an overall probability of failure. Yet the *number* of errors that separate a system from failure is a potentially more important indicator of risk. This principle has been recognized in the safety policy of the US National Aeronautics and Space Administration, which requires that systems where failure could result in loss or damage, or personal injury must be able to withstand single errors. In circumstances where loss of life or a mission-critical event could occur, systems must be able to withstand two human errors (National Aeronautics and Space Administration, 2000).

The greatest value of the safety space concept may be to proactively identify systems that are separated from disaster by a single error. While the difficulties of such an analysis should not be underestimated, techniques such as Human HAZOPS (Kirwan, 1994) may reveal at least some of the human error scenarios that can threaten a system. Information on potential single-point human failures can be gathered via incident reports, critical incident interviews or from subject matter experts. After a rail accident in which an unauthorized person intentionally moved a set of manually operated points on the main Melbourne to Adelaide rail line, an evaluation identified several other locations in the vicinity where a single unsafe act could have had similar effects (Australian Transport Safety Bureau, 1999).

Wood's 'drift toward failure'. Although a system may appear to be robust in the face of error, defences and safety nets that exist on paper may not exist in practice. It is important therefore to evaluate the operational reality of the system rather than the theoretical ideal. For example, in a survey of maintenance mechanics, most reported that they had omitted functional checks in the previous 12 months (Hobbs & Williamson, 2000). As a result, some repair procedures that on paper contain an error-capturing check may be less error-tolerant in practice than the procedure designers intended. As Woods and Cook (2002) have noted, as time passes and planned defences lose their effectiveness, complex systems have a tendency to 'drift toward failure'. This drift can be seen as a gradual movement of a system from the robust toward the fragile sector of the safety space.

Reason's dangerous defences. In some cases the resilience of a system can be increased by adding defences, or strengthening existing defences against error. However, as Reason (1997) notes, adding new defences to an already complex system may *reduce* safety rather than increase it. For example, in the early years of the AIDS epidemic, it was common practice for health care workers to re-cap used hypodermic syringes before disposing of them. Although this procedure was intended to reduce the risk of needle-stick injuries, the result was an increased rate of these potentially fatal slips. Nevertheless, despite such cases, there will be situations where the careful design and placement of defences can help to manage the risks stemming from human error.

Risk compensation. Defences against human error must be designed to suit the type of unsafe act involved. According to the risk homeostasis theory of Wilde (2001), safety interventions may fail to increase the overall level of safety if people compensate by changing their behaviour to maintain what they perceive as an acceptable level of risk. For example, installing a safety rail on a winding road may result in people driving *faster* than before, with little overall effect on the rate of road accidents. For unsafe acts involving controlled processing, the best defences may be discrete ones that do not intrude into conscious awareness, and guard against the error hazard without the person being constantly reminded that a level of protection exists. For example, vibration strips on the side of highways warn drivers that they have deviated onto the verge, yet few drivers would be aware that this safety defence is present. Even fewer would take risks such as driving when fatigued *because* they were aware that this defence was present. When it comes to errors of automatic processing however, because the action is unintended, the problem of risk homeostasis is less likely to apply. Defences against slips and lapses may take the form of physical barriers (such as gated switches) and need not be as subtle as defences against errors of controlled processing.

Conclusions

Human error presents an unremitting threat to the safety of complex systems. In most circumstances, errors must combine to form molecules, or clusters, before a serious system disturbance will occur. The safety space diagram in Figure 1 shows that fragile systems are susceptible to single errors, particularly the inevitable slips and lapses that characterise human performance. Many of our daily activities fall into this sector of the safety space. We play sport, drive motor vehicles, cross busy roads and engage in other activities where a single slip could result in an accident. Yet while these risks may be tolerable on an individual level, society demands a lower level of risk for complex technological systems. The proactive identification of fragile systems should be a key objective for managers of systems that can do harm to customers or to the public at large. Evaluating the resilience of a complex system to human error is easier said than done, and there are no guarantees that all potential single-point human failures can be uncovered. The inner workings of technological systems can be opaque and resistant to scrutiny, and errors may have consequences that were never anticipated by the system designers (Perrow, 1984). No system is free from risk, and even systems that appear to be robust will have areas where improvements can be made to increase the resilience against error molecules. Many of these safety improvements will take the form of incremental, quantitative changes such as improved training or procedures. Fragile systems on the other hand, require reforms rather than repairs. A hazard analysis, no matter how imprecise and imperfect, will have been worthwhile if it uncovers even a few scenarios where a system is vulnerable to a single human error.

References

Australian Transport Safety Bureau. (1999). Collision between freight train 9784 and ballast train 9795. (Rail investigation report R1/2000). Canberra, Australian Capital Territory: Author.

Battmann, W., & Klumb, P. (1993). Behavioural economics and compliance with safety regulations. *Safety Science, 16,* 35-46.

Casey, S. (1998). *Set phasers on stun*. (2nd ed). Santa Barbara, CA: Aegean.

Heinrich, H. W. (1941). *Industrial accident prevention. A scientific approach.* New York: McGraw-Hill.

Hobbs, A., & Williamson, A. (2000). *Aircraft maintenance safety survey: Results* (Air Safety Occasional Paper 101). Canberra: Australian Transport Safety Bureau.

Hobbs, A., & Williamson, A. (2003). Associations between errors and contributing factors in aircraft maintenance. *Human Factors, 45*, 186-201.

Kennedy, R., McIntire, M., Rashbaum, W., & O'Donnell, M. (2003, October 25). Ferry crash raises issues: were rules enforced? *The New York Times*, pp. A1, A15.

Kern, T. (1999). *Darker shades of blue*. New York: McGraw Hill.

Kirwan, B. (1994). *A practical guide to human reliability assessment.* London: Taylor & Francis.

Lawton, R. (1998). Not working to rule: Understanding procedural violations at work. *Safety Science, 28*, 77-95.

Lawton, R., & Parker D. (1998). Individual differences in accident liability: A review and integrative approach. *Human Factors, 40,* 655-671.

McDonald, N., Corrigan, S., Daly, C., & Cromie, S. (2000). Safety management systems and safety culture in aircraft maintenance organisations. *Safety Science, 34*, 151-176.

Mellor, P. (1994). CAD: Computer Aided Disaster. *High Integrity Systems, 1*(2), 101-156.

Miller, D. P., & Swain, A. D. (1987). Human error and human reliability. In G. Salvendy (Ed.), *Handbook of Human Factors* (pp. 219-250). New York: John Wiley & Sons.

National Aeronautics and Space Administration. (2000). *NASA Safety Manual* (NASA Publication No. NPG 8715.3). Washington, DC: Author.

Orlady, H. W., & Orlady, L. M., (1999). *Human factors in multi-crew flight operations*. Aldershot, UK: Ashgate.

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

Parker, D., Reason, J., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics, 38*, 1036-1048.

Perrow, C. (1984). *Normal accidents: Living with high risk technologies.* New York: Basic Books.

Rasmussen, J. (1982). Human errors: A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents, 4*, 311-333.

Reason, J. (1979). Actions not as planned: The price of automatization. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness. Vol 1: Psychological issues* (pp. 67-89). London: Academic.

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

Reason, J. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

Reason, J., & Hobbs, A. (2003). *Managing maintenance error: A practical guide*. Aldershot, UK: Ashgate.

Reason, J. Parker, D., & Lawton, R. (1998). Organizational controls and safety: The varieties of rule-related behaviour. *Journal of Occupational and Organizational Psychology, 71*, 289-304.

Redmill, F., & Rajan, J. (Eds.). (1997). *Human factors in safety critical systems*. Oxford: Butterworth Heinemann.

Rouse, W. B., & Rouse, S. H. (1983). Analysis and classification of human error. *IEEE Transactions on Systems, Man and Cybernetics SMC-13, 4,* 359-549.

Rolt, L. T. C. (1998). *Red for danger*. Stroud, UK: Sutton.

Spearman, C. (1928). The origin of error. *The Journal of General Psychology, 1*, 29-53.

Walters, J. M., & Sumwalt, R. L. (2000). *Aircraft accident analysis: Final reports.* New York: McGraw Hill.

Western Australian Department of Transport. (1999). *Collision Indian Pacific passenger train 3AP88 and freight train 3PW4N, Zanthus, WA, 18 August 1999*. Perth, WA: Author.

Which Switch? (1998, March). *Asia Pacific Air Safety, 17*, 2-3.

Wilde, J. S. (2001). *Target risk 2: A new psychology of safety and health* (2nd ed.). Kingston Ontario: PDE Publications.

Williamson, A., & Feyer, A. (1990). Behavioural epidemiology as a tool for accident research. *Journal of Occupational Accidents, 12*, 207–222.

Wilmer, W. H. (1979). The early development of aviation medicine in the United States. *Aviation, Space and Environmental Medicine, May*, 459-467.

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