Routes to failure: Analysis of 41 civil aviation accidents from the Republic of China using the human factors analysis and classification system

Li, W-C. , Harris, D. and Yu, C-S.

Routes to failure: analysis of 41 civil aviation accidents from the Republic of China using the Human Factors Analysis and Classification System

Abstract

The Human Factors Analysis and Classification System (HFACS) is based upon Reason's organizational model of human error. HFACS was developed as an analytical framework for the investigation of the role of human error in aviation accidents, however, there is little empirical work formally describing the relationship between the components in the model. This research analyses 41 civil aviation accidents occurring to aircraft registered in the Republic of China (ROC) between 1999 and 2006 using the HFACS framework. The results show statistically significant relationships between errors at the operational level and organizational inadequacies at both the immediately adjacent level (preconditions for unsafe acts) and higher levels in the organization (unsafe supervision and organizational influences). The pattern of the 'routes to failure' observed in the data from this analysis of civil aircraft accidents show great similarities to that observed in the analysis of military accidents. This research lends further support to Reason's model that suggests that active failures are promoted by latent conditions in the organization. Statistical relationships linking fallible decisions in upper management levels were found to directly affect supervisory practices, thereby creating the psychological preconditions for unsafe acts and hence indirectly impairing the performance of pilots, ultimately leading to accidents.

Keywords: Accident Analysis; Human Error; Human Factors Analysis and Classification System (HFACS)

1 INTRODUCTION

When investigating the causes of aviation accidents, the human factors focus has now shifted away from investigating skill deficiencies and has moved toward other factors such as decision-making, attitudes, supervisory factors and organizational culture (Diehl, 1989; Jensen, 1997). This change in emphasis has resulted in human error frameworks and accident investigation schemes being developed that investigate and categorise the organizational factors and psychological precursors surrounding the accident in an attempt to develop a more complete understanding of the circumstances and hence aid in the development of effective prevention strategies.

The Human Factors Analysis and Classification System - HFACS (Shappell and Weigmann, 2001, 2003, 2004; Weigmann and Shappell, 1997, 2001a, 2001b, 2001c, 2003) is perhaps the most widely used human factors accident analysis framework. HFACS was developed from Reason's organizationally based model of human error (Reason, 1990, 1997). In this model active failures (errors) of front-line operators (in this case pilots) combine with latent failures lying dormant in the system to breach its defences. These latent failures are spawned in the upper levels of the organization and are related to management and regulatory structures.

HFACS addresses human errors and the factors underpinning them at four levels. The framework is described diagrammatically in figure 1. Level 1 (unsafe acts of operators - active failures) is the level at which the majority of accident investigations have been focused in the past. These are the behaviours of the flight crew on the flight deck that contribute directly to the accident. Failures at this level can

be further classified into two sub-categories; errors and violations. Errors fall into three basic error types (skill-based, decision, and perceptual). Violations, however, are instances of the willful disregard of rules, which subsequently result in an accident. Level 2 (preconditions for unsafe acts - latent/active failures) addresses the psychological pre-cursors to the active failures at level 1, such as the substandard conditions of the operators and the operating environment which predispose them to making an error. Level 3 (unsafe supervision - latent failures) traces the causal chain of events producing the unsafe acts up to the level of the front-line supervisors. Level 4 (organizational influences - latent failures) describes the contributions of fallible decisions in upper levels of management that directly affect supervisory practices, as well as the conditions and actions of front-line operators. Each higher level affects the next downward level in HFACS framework. Reinach and Viale (2006) argue that the use of a classification system to investigate incidents and accidents has several benefits compared to a less structured, more ad hoc process. It provides a consistent and format for accident/incident data collection and analysis; it ensure the accident and incident investigations are undertaken systematically and comprehensively; it helps counteract investigator's heuristics and biases and it allows the comparisons of contributing factors across industries.

Wiegmann & Shappell (2001b) claim that the HFACS framework bridges the gap between theory and practice by providing safety professionals with a theoretically based tool for identifying and classifying the human errors in aviation mishaps. Given that the system focuses on both latent and active failures and their inter-relationships, it facilitates the identification of the underlying causes of human

error. HFACS was originally designed and developed as a human error framework for investigating and analysing human error accidents in US military aviation operations (Shappell and Weigmann, 2001) however the framework's developers have also demonstrated its applicability to the analysis of accidents in US commercial aviation (Shappell and Weigmann, 2001; Weigmann and Shappell, 2001a, 2001c) and US general aviation (Shappell and Weigmann, 2003). The system has also been used to analyse the underlying human factors causes in accidents involving remotely-piloted aircraft (Tvarnyas, Thompson and Constable, 2006); air traffic control incidents (Scarborough and Pounds, 2001) and has been further developed to investigate both maintenance error in aircraft (Krulak, 2004) and for the investigation of railroad accidents (Reinach and Viale, 2006).

Although the HFACS framework has received wide acceptance as a tool for investigating and analysing aircraft accidents there have been some criticisms of it. Beaubien and Baker (2002) noted that it was often difficult to collect information about the latent conditions from incident or accident reports. Furthermore, they added that all the evidence collected relating to its reliability and validity had been collected and analyzed by the developers of the framework. However, other authors have now successfully used and proven the system outside the US, for example in India and the Republic of China (Gaur, 2005; Li and Harris, 2005a). These studies have also shown the framework to have a high degree of inter-rater reliability when coding accident narratives. Dekker (2001) suggested that the HFACS framework has only a slight link between human error and working environment and there is some confusion between categorization and analysis. The simple assignment of errors to categories has no explanatory power and suggests neither what contributes to these errors being made nor

remedial actions to avoid them in the future. He added that the framework merely repositioned human errors by shifting them from the forefront to higher up in the organization instead of finding solutions for them. Although HFACS was based directly on the organizational theory of failure espoused by Reason (1990; 1997) at the time it was derived there was little or no quantitative data to support the theoretical model upon which it was based. However recently several studies have uncovered statistical relationships describing the cause/effect relationships between various components at the different levels in the analysis and classification system, giving support to the underpinning theory behind the framework (Li & Harris 2006a; Li & Harris, 2006b; Tvaryanas, Thompson & Constable, 2006). These analyses begin to describe statistically how actions and decisions at higher managerial levels promulgate down through the organization to result in operational errors and accidents.

However, it needs to be noted that the studies by Li & Harris (2006a, 2006b) and Tvaryanas, Thompson & Constable (2006) were undertaken in the context of military flight operations and the operation of uninhabited air vehicles and only in the case of the Li and Harris papers, did the analyses progress to the highest organizational levels. Li and Harris (2006a) clearly showed that deficiencies in Organizational Processes and Organizational Climate (HFACS level 4) increased the likelihood of deficiencies in Inadequate Supervision (HFACS level 3) thirteen-fold. The category of 'inadequate supervision' had a particularly strong association with the level 2 category of 'crew resource management' (CRM). It was suggested that failures of senior officers in supervisory positions to provide guidance and operational doctrine to pilots by promoting good CRM practices was indirectly associated with active, operational failures.

Again, the odds ratios suggested failures as a result of poor CRM were almost 13 times more likely to occur in the presence of a concomitant failure in the category of 'supervisory failure.' Finally, Li and Harris (2006a) observed that several latent pre-conditions for unsafe acts (at HFACS level 2) showed strong associations with active failures at level 1. These level 2 factors exhibited Reason's classic 'many to one' mapping of psychological precursors to active failures in several of the level 1 categories.

INSERT FIGURE 1 ABOUT HERE

As stated in the previous paragraph, this research applied the HFACS framework to analyse Republic of China (ROC) Air Force accidents. The objective was to provide probabilities for the co-occurrence of categories across adjacent levels of the HFACS to establish how factors in the upper (organizational) levels in the framework affect categories in lower (operational) levels. It was suggested that once the significant paths were identified the development of intervention strategies should proceed more rapidly and effectively. Civil airlines have very quite different organizational imperatives and organizational structure. Harris and Morley (2006) argue that military aviation is a much less open system than is commercial civil aviation (see Katz and Kahn, 1978). The armed forces exert far more control over flight operations and, at least in peacetime, the operating environment. Air Forces are responsible for much of the maintenance of the aircraft they fly; they operate their own airfields and provide their own Air Traffic Control services; they train their own pilots and engineers and all personnel are indoctrinated into the

military culture and way of working. In contrast, civil airlines operate into a wide range of airports (none of which they own); aircraft maintenance is often provided by third parties; aircraft ramp servicing is almost invariably provided by a range of external suppliers and ATC is provided by the air traffic services providers from the countries into which they either operate or overfly. From an organizational perspective, airlines can be considered to be considerably more 'open' than military aviation. As a result, it needs to be established if the same routes to failure exist in the operation of commercial aircraft.

This study examines 41 accidents occurring to civil aircraft in the ROC using the HFACS framework. Using data from ROC civil aviation accidents allows a direct comparison (without any cultural confounding effects) of the organizational routes to failure with the earlier work undertaken with the air force of the same country.

2 METHOD

2.1 Data

The aviation accident reports were obtained from ROC Aviation Safety Council between 1999 and 2006. A total of 41 accidents and reportable incidents occurred within this time period. All accidents and serious incidents conformed to the definition within the 9th edition of the Convention on International Civil Aviation, Annex 13 (International Civil Aviation Organisation, 2006). There were 24 different types of aircraft involved in the accidents analysed, including commercial jets airliners (Airbus A300, A320 and A330; Boeing B737 and B747;

McDonnell-Douglas MD11, MD82, MD83 and MD90): private jets (Bombardier BD700): turbo-prop powered aircraft (ATR 72-200; De Havilland Canada DASH-8-300, Fokker 50) and commercial helicopters (Bell UH-1H, 206 and 430; Boeing 234; Eurocopter BK117). Full copies of all these accident reports may be found on the ROC Aviation Council web site (http://www.asc.gov.tw/asc_en/accident_list_1.asp).

2.2 Classification framework

The version of the HFACS framework described in Wiegmann and Shappell (2003) was utilised in this study (see figure 1). Level 1 of the HFACS categorises events under the headings of 'unsafe acts of operators' that can lead to an accident. This comprises four sub-categories of 'decision errors'; 'skill-based errors'; 'perceptual errors' and 'violations'. Level 2 of HFACS is concerned with 'preconditions for unsafe acts'. This has seven sub-categories within it: 'adverse mental states'; 'adverse physiological states'; 'physical/mental limitations'; 'crew resource management'; 'personal readiness'; 'physical environment', and 'technological environment'. Level 3 of HFACS is concerned with 'unsafe supervision' which includes the four categories: 'inadequate supervision'; 'planned inappropriate operation'; 'failure to correct known problem', and 'supervisory violation'. Level 4, the highest level in the framework is labelled 'organizational influences' and comprises of three sub-categories: 'resource management'; 'organizational climate' and 'organizational process'.

2.3 Coding process

Two aviation human factors specialists coded each accident report independently. The analysts had previously been trained together on the use of the analysis and categorization framework to ensure that they achieved a detailed and accurate understanding of it. This training consisted of three half-day modules delivered by an aviation psychologist. The training syllabus included an introduction to the HFACS framework; explanation of the definitions of the four different levels of HFACS; and a further detailed description of the content of the eighteen individual HFACS categories. Subsequent to this the raters then jointly analysed two years of the ROC air force accident data to develop a shared understanding of the categorization process and achieve a common understanding of the categories. Prior to undertaking the present study these analysts also undertook the analysis of 523 ROC air force accidents reported elsewhere (Li and Harris, 2006a, 2006b).

The presence or the absence of each HFACS category was evaluated from the narrative of each accident report. Each HFACS category was counted a maximum of only once per accident, thus this count acted simply as an indicator of the presence or absence of each of the 18 categories within a given accident.

Where there were discrepancies in the categorisation of an accident, the raters convened and resolved their observations.

3 RESULTS

3.1 Inter-rater reliability

Prior to resolution of discrepancies in coding between the raters, the inter-rater reliabilities, calculated on a category-by-category, basis were assessed using Cohen's Kappa. In half the categories the Kappa values was in excess of 0.40, which is regarded as being acceptable (Landis and Koch, 1977). For the remainder of the categories, though, the Kappa value failed to achieve this level. Values for Kappa range between 0 and 1.00, where 0 represents complete independence between raters and 1.00 is indicative of perfect agreement (Cohen, 1960). Below 0.40 is regarded as a poor level of inter-rater reliability. However, Cohen's Kappa has several weaknesses as an index of inter-rater reliability. Low observed frequencies can distort Kappa values, deflating its value where there is actually a very high level of agreement. Cohen's Kappa becomes unreliable when the vast majority of observations fall into just one of the categories and there is also a high percentage of agreement between raters in this category (as in the category 'adverse physiological states' – see table 1). In such a case Cohen's Kappa will be low as the statistic is based upon expected probabilities calculated from the marginal observed totals (Huddlestone, 2003). Gwet (2002a, 2002b) also observed that Kappa does not take in account raters' sensitivity and specificity. Gwet also observed that Kappa becomes unreliable when raters' agreement is either very small or very high. As a result, inter-rater reliabilities were also calculated as a simple percentage rate of agreement. These showed reliability figures of between 63.4% and 95.1%, indicating acceptable reliability between the raters.

3.2 Overview of analysis rationale

The data were cross-tabulated to describe the strength of association between the categories at adjacent levels in the HFACS framework. Chi-square (χ^2) analyses were performed to estimate the statistical strength of association between the categories in the higher and lower levels of the framework. However, as the χ^2 test is a simple test of association these analyses were supplemented with further analyses using Goodman and Kruskal's lambda (λ) which was used to calculate the proportional reduction in error (PRE). The lambda statistic is a measure (ranging from 0 to 1, where 1 represents certainty) of the extent that knowledge of the category of one variable improves the prediction of the other variable. Lambda (Goodman and Kruskal, 1954) has the advantage of being a directional statistic. The lower level categories in the HFACS were designated as being dependent upon the categories at the immediately higher level in the framework, which is congruent with the framework's underlying theoretical assumptions. From a theoretical standpoint, lower levels in the HFACS cannot adversely affect higher levels. Higher levels in the HFACS are deemed to influence (cause) changes at the lower organizational levels, thus going beyond what may be deemed a simple test of co-occurrence between categories. Finally, odds ratios were also calculated to provide an estimate of the likelihood of the presence of a contributory factor in one HFACS category being associated with the concomitant presence of a factor in another category. However, it must be noted that odds ratios are an asymmetric measure and so are only really theoretically meaningful when associated with a non-zero value for lambda.

3.3 Results

INSERT TABLE 1 ABOUT HERE

In total 330 instances of human error, describing the underlying causal factors of the 41 accidents, were recorded using the HFACS framework.

Initial results found that there were 107 instances of acts (32.4% of all errors recorded) at the level of 'unsafe acts of operators' (level 1). Errors at this level were implicated in 85.4% (35) of accidents. There were 84 (25.4%) instances of errors at the level of 'Preconditions for unsafe acts' level (level 2). Errors recorded at this level were implicated in 82.9 % (34) of all the accidents analysed. At the level of 'unsafe supervision' (level 3), there were 82 (24.8%) instances of error that were implicated in 75.6% (31) of accidents. Finally, at the 'organizational influences' level in the HFACS model (level 4) there were 57 (17.3 %) instances of error recorded, implicated in 68.3% (28) of accidents in the sample (see table 1).

To keep the overall type I error rate across the whole analysis to $p<0.05$, using a Bonferroni adjustment (Sankoh, Huque and Dubey, 1997) the individual alpha level for each analysis was set to $p<0.001$. Using this value, 16 relationships between categories at adjacent HFACS levels showed significant associations (see table 2).

INSERT TABLE 2 ABOUT HERE

Analysis of the strength of association between categories at HFACS level-4 'organizational influences' and HFACS level-3 'unsafe supervision' indicated that of a possible 12 relationships, six pairs of associations were significant ($p<0.001$) between categories at adjacent levels. 'Organizational process' was significantly associated with all four supervisory factors at level-3: 'inadequate supervision'; 'planned inappropriate operations'; 'failed to correct a known problem'; and 'supervisory violations'. Resource management was significantly associated with two categories at level-3, 'inadequate supervision' and 'planned inappropriate operations'. These statistically significant relationships are summarized in table 2 and are described diagrammatically in figure 2. Organizational climate had no significant associations with any level-3 categories.

INSERT FIGURE 2 ABOUT HERE

It will be noted from table 2 that in several of the tests of association performed between categories at HFACS level-4 and 3 very high odds ratios are observed, all of which are associated with non-zero values for lambda. Inadequate supervision is over ten times more likely to occur when there are organizational level issues associated with poor resource management. Similarly, inadequate supervision is nine-times more likely to occur in the presence of poor organizational

processes, an issue also associated with inflating the likelihood of failing to correct known problems by a similar amount.

Analysis of the strength of association between categories at HFACS level-3 'unsafe supervision' and level-2 'pre-conditions for unsafe acts' shows just three pairs of associations to be significant ($p<0.001$) between categories at these adjacent levels. These were 'inadequate supervision' at level-3 versus 'CRM' at level-2; 'planned inappropriate operations' at level three with the 'physical environment' at level-2 and 'supervisory violation' versus 'personal readiness'. Of these comparisons it can be seen that poor CRM was over eight times more likely to occur in the presence of 'inadequate supervision' at the higher organizational level.

The analysis of the strength of association between categories at HFACS level-2 'pre-conditions for unsafe acts' and level-1 'unsafe acts of operators' shows seven pairs of associations to be significant between the categories at these two adjacent levels. These mainly fall into two distinct groupings. In the first group the level-2 category 'CRM' was significantly associated with three categories of unsafe act: 'decision errors'; 'skill-based errors' and 'violations'. In the second grouping 'adverse mental states' was also significantly associated with three categories of unsafe act: 'decision errors'; 'skill-based errors' and 'perceptual errors'. The final relationship was between the HFACS level-2 category of the 'physical environment' and the level-1 category of 'perceptual error'.

Of particular note are the downward relationships the category of poor CRM has with the three level-1 categories. Inspection of the associated odds ratios show between (approximately) a 30-40 fold

increase in the likelihood of error or violation in the presence of poor CRM practices.

4 DISCUSSION

It can be seen from the data presented in table 1 that the majority of HFACS categories had large enough numbers of instances of occurrence in the data set to allow reasonable confidence in the pattern of results obtained. All categories also exhibited good levels of inter-rater reliability ranging from 63.4% for the HFACS level-4 category of 'resource management' to 95.1% for the level-2 category of 'adverse physiological states'. These were as good as, or in excess of the levels reported in previous studies (e.g. Gaur, 2005; Li and Harris, 2005a; Weigmann and Shappell, 1997).

Reason (1990, 1997) proposed that latent conditions promoting unsafe acts are inevitably present in all systems. The original decision on how to allocate resources made at the highest levels in the organization may originally have been based on sound commercial arguments but such inequities can create reliability or safety problems in other, operational parts of the system. The analyses in this paper clearly show that inadequacies at HFACS level-4 ('organizational influences') had associations with further inadequacies at HFACS level-3 ('unsafe supervision). See table 2 and figure 2. The category of 'organizational process' is a particularly important factor at this highest organizational level. Poor 'organizational processes' were associated with inadequacies in all categories at the level of 'unsafe supervision' and hence indirectly were ultimately at the root of many operational errors resulting in

accidents. In the HFACS framework inadequacies in 'organizational processes' includes such issues as imposing excessive time pressures on staff; poor mission scheduling; poor incentivization; management failing to set clearly defined objectives; poor risk management programmes; inadequate management checks for safety; and failing to establish safety programmes. 'Resource management' (which involved the selection, staffing and training of human resources at an organizational level; excessive cost cutting; providing unsuitable equipment, and a failure to remedy design flaws) also showed strong relationships with the two level-3 categories of 'inadequate supervision' and 'planned inappropriate operations). Both Reason (1990) and Wiegmann and Shappell (2003) hypothesized that inappropriate decision-making by upper-level management can adversely influence the personnel and practices at the supervisory level, which in turn affects the psychological pre-conditions and hence the subsequent actions of the front-line operators. This study provides statistical support for this hypothesized relationship. A similar pattern of results was also found in the analysis of 523 ROC air force accidents previously reported by Li and Harris (2006a, 2006b).

The level-2 ('preconditions for unsafe acts') category of 'CRM' was perhaps the key factor in HFACS framework. CRM encompasses issues such as a lack of teamwork on the flight deck; poor communication between flight crew; failures of leadership and inadequate crew briefings. Inadequacies in CRM practices were particularly influenced by the level-3 category of 'inadequate supervision'. This category in the HFACS framework encompasses issues such as a failure to provide proper training or adequate rest periods; a lack of accountability; failure to track qualifications and performance of personnel; the use of untrained supervisors and a general loss of situation awareness at

the supervisory level). The risk of 'inadequate supervision' was itself greatly inflated by poor organizational processes and resource management issues. The earlier study of Li and Harris (2006a; 2006b) also observed that the issue of inadequate supervision was the key link between inadequacies at the higher organizational levels and poor CRM.

Reason (1990) suggested that human behaviour is governed by the interplay between psychological and situational factors. In particular two pre-conditions for unsafe acts (HFACS level-2 – 'CRM' and 'adverse mental states') showed a strong statistical relationship with the active failures of the operators at level-1 (see table 2 and figure 2). This was also observed in the earlier study of ROC air force accidents. Poor CRM is associated with 'decision errors' (instances in this category included, selecting inappropriate strategies to perform a mission; improper in-flight planning; making an inappropriate decision to abort a take-off or landing, or using improper remedial actions in an emergency) and 'skill based errors' (for example, inappropriate stick and rudder coordination; excessive use of flight controls; glideslope not being maintained, and adopting an improper airspeed or altitude). However it is also noticeable that poor CRM practices are associated with the level-1 category of 'violations', which encompasses issues such as intentionally ignoring standard operating procedures (SOPs); neglecting SOPs; applying improper SOPs; and diverting from SOPs. Adverse mental states are those conditions that affect pilot performance, such as loss of situational awareness; task fixation; distraction and mental fatigue due to stress. It is perhaps not too surprising that such mental states pre-dispose accident involved pilots to all the main categories of human error.

Reason (1990, 1997) has suggested that there is a 'many to one' mapping of the psychological precursors of unsafe acts and the actual errors themselves, making it difficult to predict which actual errors will occur as a result of which preconditions. The results of this study using the HFACS framework support this assertion. It is also evident that accidents rarely involved just a single error. Overall every accident involved an average of over eight individual errors distributed across the various organizational levels and even across different organizations. These erroneous actions need not be confined to either the flight deck or the airline. For example, the report on the investigation of the runway incursion occurrence involving a TransAsia Airways aircraft (flight GE543) concluded that there was inadequate planning and implementation in airport construction safety procedures by both the Taiwanese Civil Aviation Authority (CAA) and the Military Authorities (both problems in 'organizational process', HFACS level-4); there was an error in issuing an approval to land after the curfew time ('decision error', HFACS level-1) and insufficient cooperation and coordination between the CAA and the military base authorities prior to commencing construction work ('inadequate supervision', HFACS level-3). There was also a lack of communication between ATC and workers on the ground (a coordination and personnel resource management ['CRM'] issue, HFACS level-2) and finally there was the action of workers to enter the airfield without approval by ATC ('violation', HFACS level-1) (Aviation Safety Council, 2004).

The errors finally resulting in an accident may also be separated by considerable periods of time. To illustrate, in the China Airlines CI-611 accident, the principal causal factors could be tracked back to

a failure to follow the Structural Repair Manual for the repair of a tail strike occurring to the aircraft 22 years prior to the accident. It was concluded that there was no work audit system and a failure to provide both flight safety inspection training and a handbook for inspectors ('resource management' and 'organizational process' issues, HFACS level-4). There was also inadequate supervision when conducting the Corrosion Prevention and Control Program ('inadequate supervision', HFACS level-3). There was a failure of communication between the Maintenance Operation Center and the Maintenance Planning Section in the airline (a 'CRM' issue, HFACS level-2) and a failure to provide adequate equipment for structural inspections ('technology environment', HFACS level-2). This contributed directly to the line operator's failure to detect structural defects in the rear pressure bulkhead ('perceptual' and 'skill-based errors', level-1). The pressure bulkhead finally ruptured causing the aircraft to break up in flight (Aviation Safety Council, 2002; Li & Harris, 2005b).

There are statistically significant associations between causal factors at the higher organizational levels, psychological contributory factors and errors in decision-making, skilled performance and the violations committed by pilots (see table 2 and figure 2). It can even be suggested that poor organizational processes at the highest levels in the organization result in poor supervisory oversight, which itself results in poor CRM resulting in wilful violations of the SOP. However, some care needs to be taken when interpreting the statistical relationships presented within figure 2. In a few categories the frequency counts are moderately small. Furthermore, the frequency counts within categories were all derived from accidents. It is unknown (and at the higher organizational levels perhaps it is unknowable) how often instances

within the various HFACS categories have occurred in day-to-day operations that have not resulted in an accident. Thus, the relationships between HFACS levels and categories should not be interpreted outside the accident causal sequence. However, some insight into the relative frequency of instances of error that have not resulted in an accident at HFACS levels 1 and 2 may perhaps be gained by the retrospective analysis of LOSA (Line Operations Safety Audit) data. These are data gained through the observation (by trained observers) of crews on the flight deck engaged in everyday line operations, undertaken for the purpose of identifying the day-to-day threats to commercial aviation safety (see Klinect, Wilhelm & Helmreich, 1999; Helmreich, Klinect & Wilhelm, 2001). For example, in LOSA data collected by Thomas (2003) it was reported that in line operations, crews often did not demonstrate effective error detection. More than half of all errors observed remaining undetected by one or both of the flight crew. Further, fine-grain analysis of this type of data may give some insight into the frequency with which undetected errors committed at the lower HFACS levels do not subsequently result in an accident. Nevertheless, the results of this study of civil aviation accidents occurring in the ROC show a remarkable similarity to the study of military accidents conducted in the air force of the same country (Li & Harris 2006a; LI & Harris 2006b).

5 CONCLUSIONS

This study provides an understanding, based upon empirical evidence, of how actions and decisions at higher managerial levels in the operation of commercial aircraft in the ROC result in errors on the flight deck and subsequent accidents. The results show clearly

defined, statistically-described paths that relate errors at level-1 (the operational level) with inadequacies at both the immediately adjacent and also higher levels in the organization. To reduce significantly the accident rate these 'paths to failure' relating to these organizational and human factors must be addressed. This research draws a clear picture that supports Reason's (1990) model of active failures resulting from latent conditions in the organization. Furthermore, the mechanisms relating to operational errors seem to be common to both civil and military operation in the ROC despite the different organizational imperatives and organizational structures underlying these types of operation (see Harris and Morley, 2006). Further work needs to be undertaken to establish if a similar pattern of results are found in other countries and cultures.

The results suggest that interventions at HFACS levels 1 and 2 would only have limited effect in improving overall safety. For example, CRM inadequacies are associated with subsequent error in three out of four categories at HFACS level-1. However, improving CRM practices alone is unlikely to have a major impact on safety unless the supervisory processes (level-3) and organizational processes (level-4) are in place to provide things such as training facilities; the mechanisms to oversee CRM training and monitor its effectiveness, and respond to any further changes required in the training program. All of these activities require organizational commitment and capacity, which can only be provided by the highest levels of management. Furthermore, on a 'bang for the buck' basis, interventions at higher levels are likely to be the most cost effective in the net safety benefits they realize. Remedial safety actions should be aimed at the higher organizational areas that share the greatest numbers of associations with factors at lower organizational levels, in this case

'organizational processes' and 'resource management' at HFACS level-4. These are the categories at the root of the paths of association with many other HFACS categories. This strongly suggests that the greatest gains in safety benefits could be achieved by targeting actions in these areas.

References

Aviation Safety Council, 2002. In-flight Breakup Over Taiwan Strait Northeast of Makung, Penghu Island China Airlines Flight CI-611, Report no. ASC-AOR-05-02-001. Author; Taipei, ROC.

Aviation Safety Council, 2004. TransAsia Airways Flight 543 Runway Incursion Investigation Report, Report no. ASC-AOR-04-10-001. Author; Taipei, ROC.

Beaubien, J.M., Baker, D.P., 2002. A Review of Selected Aviation Human Factors Taxonomies, Accident/Incident Reporting Systems and Data Collection Tools. Int. J. Appl. Aviat. Studies 2, 11-36.

Cohen J., 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Dekker, S.W.A., 2001. The re-invention of human error. Human Factors and Aerospace Safety 1, 247-266.

Diehl, A., 1989. Human Performance/System Safety Issues in Aircraft Accident Investigation and Prevention. In: Proceedings of 11th International Symposium on Aviation Psychology, Columbus, OH

Gaur, D., 2005. Human Factors Analysis and Classification System Applied to Civil Aircraft Accidents in India. Aviat. Space Environ. Med. 76, 501-5.

Goodman, L., Kruskal, W.H., 1954. Measures of association for cross-classifications. Journal of the American Statistical Association 49, 732-764.

Gwet, K., 2002a. Kappa statistics is not satisfactory for assessing the extent of agreement between raters. Statistical methods for inter-rater reliability assessment, 1, 1-6. Retrieved 24 November 2006 from www.stataxis.com/files/articles/inter_rater_ reliability_ dependency.pdf

Gwet, K., 2002b. Inter rater reliability: Dependency on trait prevalence and marginal homogeneity. Statistical methods for inter-rater reliability assessment, 1, 1-9. Retrieved 24 November, from www.stataxis.com/files/articles/inter_rater_reliability_ dependency.pdf.

Harris, D., Morley, F.J.J., 2006. An open systems approach to safety culture: actions, influences and concerns. Keynote address to Evolving System Safety 2006 (7th International Symposium of the Australian Aviation Psychology Association). Manly, 9-12 November, 2006.

Helmreich, R.L., Klinect, J.R., & Wilhelm, J.A., 2001. System safety and threat and error management: The line operations safety audit (LOSA). In: Proceedings of 11th International Symposium on Aviation Psychology. Ohio State University, Columbus, OH.

Huddlestone, J.A., 2003. An Evaluation of the Training Effectiveness of a Low-Fidelity, Multi-Player Simulator for Air Combat Training. Unpublished PhD Thesis. College of Aeronautics, Cranfield University, England.

International Civil Aviation Organisation, 2006. Convention on International Civil Aviation ICAO (9th edition) Document 7300/9. ICAO, Monteal, Canada.

Jensen, R.S., 1997. The Boundaries of Aviation Psychology, Human Factors, Aeronautical Decision Making, Situation Awareness, and Crew Resource Management. Int J Aviat Psychol 7, 259-68.

Katz, D., Kahn, R.L., 1978. The social psychology of organizations (2nd Edition). John Wiley, New York.

Klinect, J.R., Wilhelm, J.A., & Helmreich, R.L., 1999. Threat and error management: Data from line operations safety audits. In: Proceedings of 10th International Symposium on Aviation Psychology. Ohio State University, Columbus, OH.

Krulak., D.C., 2004. Human Factors in maintenance: Impact on aircraft mishap frequency and severity. Aviat. Space Environ. Med. 75, 429-432.

Landis, J.R., Koch, G.G., 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33, 159-174.

Li, W.C., Harris, D. 2005a. HFACS Analysis of ROC Air Force Aviation Accidents: reliability analysis and cross-cultural comparison. Int. J. Appl. Aviat. Studies, 5, 65-81.

Li, W.C., Harris, D. 2005b. Where safety culture meets national culture: the how and why of the China Airlines CI-611 accident. Human Factors and Aerospace Safety 5, 345-353.

Li, W.C., Harris, D. 2006a. Pilot error and its relationship with higher organizational levels: HFACS analysis of 523 accidents. Aviat. Space Environ. Med. 77, 1056-1061.

Li, W.C., Harris, D., 2006b. Breaking the Chain: An Empirical Analysis of Accident Causal Factors by Human Factors Analysis and Classification System (HFACS). In: Proceedings of International Society of Air Safety Investigators Seminar 2006. September 11-14, Cancun, Mexico.

Reason, J.T., 1990. Human Error. Cambridge University Press, Cambridge.

Reason, J.T., 1997. Managing the risks of organizational accidents. Ashgate, Aldershot, UK.

Reinach, S. and Viale, A. (2006). Application of a human error framework to conduct train accident/incident investigations. Aviat. Space Environ. Med. 30, 396-406.

Sankoh, A.J., Huque, M.F., Dubey, S.D., 1997. Some comments on frequently used multiple endpoint adjustments methods in clinical trials. Statistics in Medicine 16, 2529-2542.

Scarborough, A., Pounds, J., 2001. Retrospective human factors analysis of ATC operational errors. In: Proceedings of 11th International Symposium on Aviation Psychology, Columbus, OH.

Shappell, S.A., Wiegmann, D.A., 2001. Applying Reason: the Human Factors Analysis and Classification System (HFACS). Human Factors and Aerospace Safety 1, 59-86.

Shappell, S.A., Wiegmann, D.A., 2003. A Human Error Analysis of General Aviation Controlled Flight Into Terrain Accidents Occurring Between 1990-1998, Report no. DOT/FAA/AM-03/4, Federal Aviation Administration, Washington, DC.

Shappell, S.A., Wiegmann, D.A., 2004. HFACS Analysis of Military and Civilian Aviation Accidents: A North American Comparison. In: Proceedings of International Society of Air Safety Investigators, Australia, Queensland, 2-8 November, 2004.

Thomas, M.J.W., 2003. Improving organisational safety through the integrated evaluation of operational and training performance: an adaptation of the Line Operations Safety Audit (LOSA) methodology. Human Factors and Aerospace Safety 3, 25-45.

Tvarnyas, A.P., Thompson, W.T., Constable, S.H., 2006. Human factors in remotely piloted aircraft operations: HFACS analysis of 221 mishaps over 10 years. Aviat. Space Environ. Med. 77, 724-732.

Wiegmann, D.A., Shappell, S.A., 1997. Human Factors Analysis of Postaccident Data: Applying Theoretical Taxonomies of Human Error. Int. J. Aviat. Psychol. 7, 67-81.

Wiegmann, D.A., Shappell, S.A., 2001a. Human Error Analysis of Commercial Aviation Accidents: Application of the Human Factors Analysis and Classification System. Aviat. Space Environ. Med. 72, 1006-1016.

Wiegmann, D.A., Shappell, S.A., 2001b. Applying the Human Factors Analysis and Classification System to the Analysis of Commercial Aviation Accident Data. In: Proceedings of 11th International Symposium on Aviation Psychology. Ohio State University, Columbus, OH.

Wiegmann, D.A., Shappell, S.A., 2001c. Human Error Perspectives in Aviation. , Int. J. Aviat. Psychol. 11, 341-357.

Wiegmann, D.A., Shappell, S.A., 2003. A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System. Ashgate, Aldershot, UK.

TABLE CAPTIONS

Table 1 Frequency counts (post-rater resolution) and inter-rater reliability statistics (prior to resolution) for each HFACS category for all 41 accidents. Note that the percentages in the table will not equal 100%, because in many cases more than one causal factor was associated with the accident.

Table 2 Significant chi-square test of association ($p<0.001$) and associated values for Goodman and Kruskal's Lambda for the analysis of upper level and adjacent downward level categories in the HFACS framework for the data derived from the analysis of the 41 accidents occurring to ROC commercial aircraft between 1999 and 2006. All tests have 1 degree of freedom. All other comparisons were non-significant.

Table 1

HFACS Category		Frequency	Percentage	Cohen's Kappa	Inter-rater Reliability (percentage Agreement)
Level-4, Organizational Influences	Organizational process	25	60.9	.510	75.6%
	Organizational climate	9	21.9	.492	85.4%
	Resource management	23	56.1	.262	63.4%
Level-3, Unsafe Supervision	Supervisory violation	23	56.1	.360	68.3%
	Failed correct a known problem	15	36.6	.358	73.2%
	Planned inappropriate operations	18	43.9	.310	68.3%
	Inadequate supervision	26	63.4	.369	70.7%
Level-2, Preconditions for Unsafe Acts	Technology environment	7	17.1	.000	70.7%
	Physical environment	19	46.3	.610	80.5%
	Personal readiness	8	19.5	.547	87.8%
	Crew resource management	28	68.3	.459	73.2%
	Physical/mental limitation	7	17.1	.109	80.5%
	Adverse physiological states	1	2.4	.000	95.1%
	Adverse mental states	14	34.1	.229	68.3%
Level-1, Unsafe Acts of Operators	Violations	28	68.3	.584	80.5%
	Perceptual errors	21	51.2	.517	75.6%
	Skilled-based errors	29	70.7	.547	78.0%
	Decision errors	29	70.7	.597	82.9%

Table 2

Significant association between upper level and adjacent downward level categories in the HFACS framework	Chi-square		Lambda	Odds ratio
	Value	p		
HFACS Level 4 association with Level 3 categories				
Resource management * Planned inappropriate operations	14.009	.000	.500	2.500
Resource management * Inadequate supervision	10.777	.001	.333	10.857
Organizational process * Planned inappropriate operations	15.105	.000	.500	2.813
Organizational process * Inadequate supervision	22.563	.000	.667	9.000
Organizational process * Failed to correct a known problem	10.408	.001	.200	9.500
Organizational process * Supervisory violation	20.251	.000	.667	3.000
HFACS Level 3 association with Level 2 categories				
Inadequate supervision * CRM	9.744	.001	.231	8.250
Planned inappropriate operations * Physical environment	8.946	.001	.421	7.367
Supervisory violation * Personal readiness	10.752	.001	.000	0.000
HFACS Level 2 association with Level 1 categories				
Adverse mental states * Decision error	12.476	.001	.000	0.000
Adverse mental states * Skill based error	12.476	.001	.000	0.000
Adverse mental states * Perceptual error	10.124	.001	.045	12.000
CRM * Decision errors	20.882	.000	.583	43.333
CRM * Skill-based errors	20.882	.000	.583	43.333
CRM * Violations	17.973	.000	.583	27.778
Physical environment * Perceptual errors	10.896	.001	.500	10.000

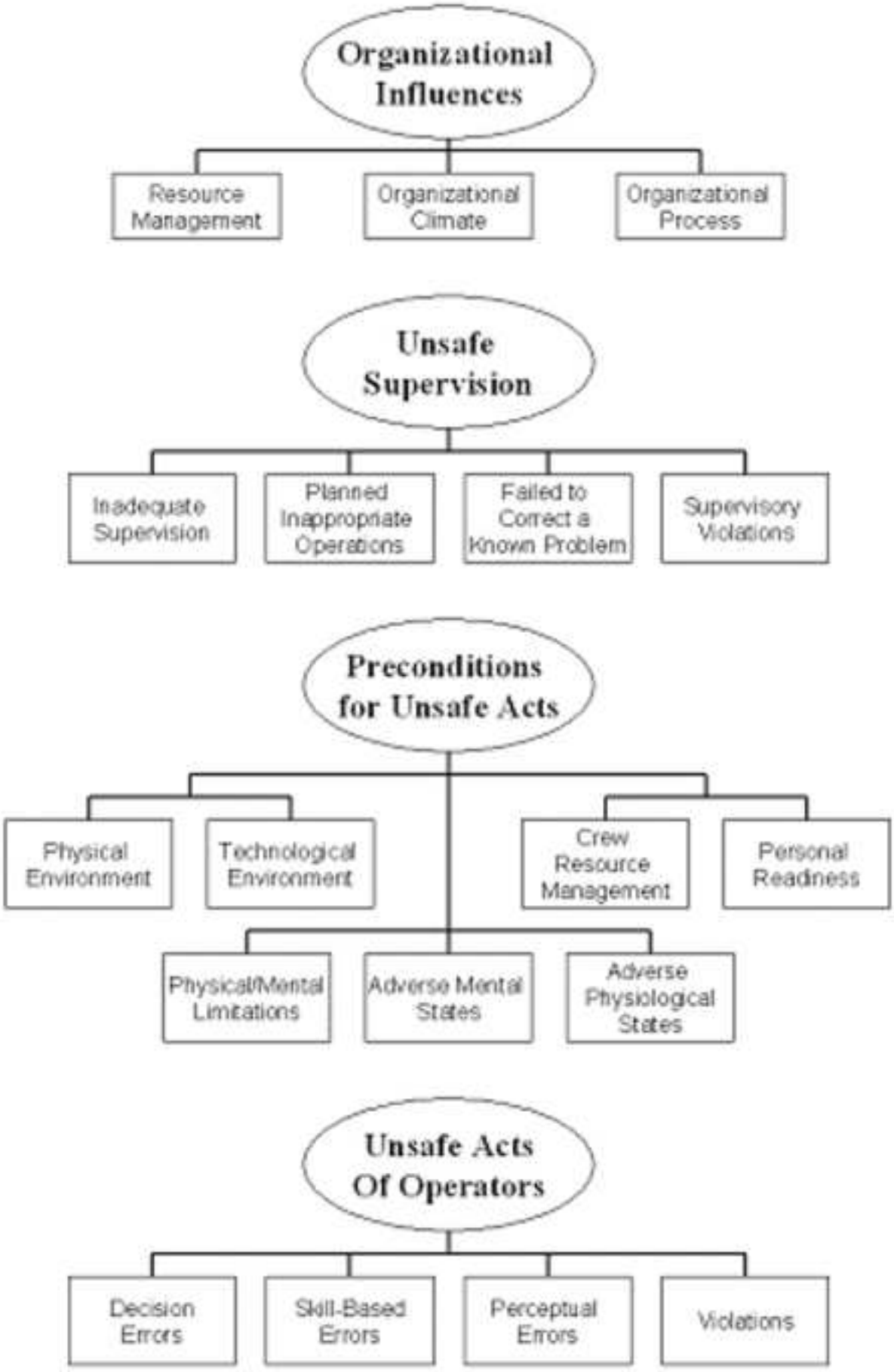
FIGURE LEGENDS

Figure 1 The HFACS framework. Each upper level is proposed to affect items at the lower levels (Wiegmann and Shappell, 2003).

Figure 2 Paths between categories at the four levels in the HFACS framework showing the significant associations ($p<0.001$) using chi-square ($\chi 2$) and lambda (λ)

Figure(s)

Click here to download high resolution image



Figure(s)

Click here to download high resolution image