Pilot error and its relationship with higher organizational levels:
HFACS analysis of 523 accidents

Wen-Chin Li & Don Harris

Department of Human Factors, School of Engineering
Cranfield University, Bedfordshire MK43 0AL, United Kingdom

Tel: +44 (0) 1234 750111 ext. 5189 (W. Li)
+44 (0) 1234 758227 (D. Harris)
Fax: +44 (0) 1234 758209
E-mail: w.li.2002@cranfield.ac.uk (W. Li)
E-mail: d.harris@cranfield.ac.uk (D. Harris)

Running Head: Empirical Analysis of HFACS

Manuscript metrics
Word count for Abstract: 195
Word count for Narrative: 3,307
Number of References: 28
Number of Tables: 2
Number of Figures: 2

-----------------------------

INSERT FOOTNOTE 1 HERE
ABSTRACT

**Introduction:** Based upon Reason’s model of human error, the Human Factors Analysis and Classification System (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 that formally describes numerically the relationship between the levels and components in the model (the organizational structures, psychological pre-cursors of errors and actual errors). **Method:** This research analyzed 523 accidents in the Republic of China (ROC) Air Force between 1978 and 2002 through the application of the HFACS framework. **Results:** The results revealed several key 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). **Conclusions:** This research lends support to Reason’s model that suggests that active failures are promoted by latent conditions in the organization. Fallible decisions in upper command levels were found to directly affect supervisory practices, thereby creating preconditions for unsafe acts and hence indirectly impaired performance of pilots, leading to accidents. The HFACS framework was proven to be a useful tool for guiding accident investigations and developing accident prevention strategies.

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

In recent years, the focus on human error in aviation accidents has shifted away from skill deficiencies and toward decision-making, attitudes, supervisory factors and organizational culture as being the primary contributory factors (3, 9 & 10). This change in emphasis has resulted in several human error frameworks and accident investigation schemes being developed (e.g. 3, 7 & 22) aimed at gaining a more complete understanding of the circumstances surrounding the accident and hence aid in the development of effective prevention strategies.

Dekker (2) proposed that human errors are systematically connected to features of operators’ tools and tasks, and error has its roots in the organizational system. In a similar vein, Orasanu and Connolly (16) suggested that decision-making often occurs in an organizational context, and that the organization influences pilots’ decisions directly by stipulating standard operating procedures, and indirectly through its norms and culture. Maurino et al. (15) pointed out that it is important to understand how decisions made by pilots are influenced by the actions of management at the higher levels in their organizations. However, there is little empirical work describing the statistical associations between organizational structures and the actual errors committed by pilots.

The Human Factors Analysis and Classification System – HFACS (20, 21, 22, 24, 25, 26, 27 & 28) is derived from Reason’s (17, 18)
organizationally based model of human error. In this model active failures (errors) of front-line operators combine with latent failures lying dormant in the system to breach its defenses. These latent failures are spawned in the upper levels of the organization and are related to management and regulatory structures. Wiegmann & Shappell (25) 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 analyzing human error accidents in US military aviation operations (21) however the framework’s developers have also demonstrated its applicability to the analysis of accidents in US commercial aviation (21, 24, 25 & 26) and US general aviation (22).

Beaubien and Baker (1) criticized the validation evidence supporting the utility of the HFACS system as it was all 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 Taiwan and India (13 & 4), and also other applications of the HFACS methodology are now being reported. The system has been used to analyze the underlying human factors causes in accidents involving remotely-piloted aircraft (23); an adaptation of the method (HFACS-ME) has been developed to investigate maintenance error (11) and a further
adaption of the system has been developed for the investigation of railroad accidents (HFAC-RR; ref. 19).

HFACS addresses human errors 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 are focused. Failures at this level can be classified into two categories; errors and violations. Level 2 (preconditions for unsafe acts - latent/active failures) addresses the latent failures within the causal sequence of events as well as the more obvious active failures. It also describes the substandard conditions of operators and the substandard practices that they perform. Level 3 (unsafe supervision - latent failures) traces the causal chain of events producing 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. However, only one of the papers cited previously (23) reports any statistical relationships describing empirically the associations between the tiers in the model in an attempt to validate the presumed causal links between latent failures and the unsafe acts of the operator. What is more, this research was concerned with the operation of uninhabited air vehicles.
This research applied the HFACS framework to 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. Once the significant paths in the framework have been identified, the development of accident intervention strategies should proceed more rapidly and effectively.

METHOD

Data

The data were derived from the narrative descriptions of accidents occurring in the ROC Air Force between 1978 and 2002. The data set comprised of all 523 accidents occurring during this 25-year period and for the purposes of analysis can therefore be considered to be the entire population of ROC Air Force accidents during this period (14). The data comprised 206 (39.4%) class-1 accidents (cost to repair over 65% of original price or resulted in death of the crew), 78 (14.9%) class-2 accidents (cost to repair between 35 and 65% of original price or crew sustained serious injury) and 239 (45.7%) class-3 accidents (cost to repair between 3-35% of original price or crewmember sustained
minor injury). Fighter aircraft were involved in 67.5% of accidents; training aircraft in 21.6% and cargo aircraft in 10.9%. Accidents occurred during missions including air interception; air combat tactics; instrument flight; cross-country; surface attack; close pattern; formation; and test flight.

The Aviation Safety Unit (ASU) is responsible for all ROC Air Force accident investigation. For each accident involving a military aircraft, the 24-hour on call Investigator-In-Charge follows a standard procedure for conducting the investigation. The initial stage collects relevant information for further analysis including the accident classification; identification details; pilots’ information; personnel involved; aircraft information; mission and flight details; history of flight; impact and post-impact information; meteorological information; radar information and transmissions to and from Tactical Air Traffic Control. The wreckage of the aircraft is then recovered for investigation by the engineering teams. The final report details the causal factors of the accident and contains recommendations for accident prevention.

Classification framework

This study used the version of the HFACS framework described in Wiegmann & Shappell (27). The first (operational) level of HFACS categorizes events under the general heading 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'. The second level of HFACS concerns 'preconditions for unsafe acts' which has seven further sub-categories: 'adverse mental states'; 'adverse physiological states'; 'physical/mental limitations'; 'crew resource management'; 'personal readiness'; 'physical environment', and 'technological environment'. The third level of HFACS is 'unsafe supervision' which includes 'inadequate supervision'; 'planned inappropriate operation'; 'failure to correct known problem', and 'supervisory violation'. The fourth and highest level of HFACS is 'organizational influences' and comprises of the sub-categories of 'resource management'; 'organizational climate' and 'organizational process'. HFACS is described diagrammatically in figure 1.

Coding process

Each accident report was coded independently by two investigators, an instructor pilot and an aviation psychologist. These investigators were trained on the HFACS framework together for 10 hours to ensure that they achieved a detailed and accurate understanding to the categories of the HFACS. The training process consisted of three half-day modules delivered by an aviation psychologist. The training contents included an introduction to the HFACS framework; an explanation of the definitions of the four different levels of HFACS; and a further detailed description of the content of the eighteen HFACS categories in the context of military operations. The raters also jointly analyzed two years of the ROC accident data to develop a common
understanding of the process and achieve a common understanding of the categories.

The presence (coded 1) or the absence (coded 0) of each HFACS category was assessed in each accident report narrative. To avoid over-representation from any single accident, each HFACS category was counted a maximum of only once per accident. The count acted simply as an indicator of presence or absence of each of the 18 categories in a given accident. In total instances of 1762 category assignments were made to described the causal factors underlying the 523 accidents coded.

The inter-rater reliabilities calculated on a category-by-category basis were assessed using Cohen’s Kappa. The values obtained ranged between 0.44 and 0.83 (see table 1). Fourteen HFACS categories exceeded a Kappa of 0.60 indicating substantial agreement (12). As Cohen’s Kappa can produce misleadingly low figures for inter-rater reliability where the sample size is small or where there is very high agreement between raters associated with a large proportion of cases falling into one category (8), inter-rater reliabilities were also calculated as a simple percentage rate of agreement. These showed reliability figures between 72.3% and 96.4%, further indicating acceptable reliability between the raters. See Li & Harris (13) for further details.
The data were cross tabulated to describe the association between the categories at adjacent levels in the HFACS analytical framework. Goodman and Kruskall’s (5) lambda (λ) was used to calculate the proportional reduction in error (PRE). The Lambda statistic is analogous to the R squared statistic for continuous data. For categorical data (such as that found in contingency tables), its value reflects the PRE when predicting the outcome category from simply the baseline prevalence as compared to using information from the predictive category (i.e. predicting lower from higher levels of HFACS). For the purposes of this study 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 theoretical assumptions underlying HFACS: from this standpoint, lower levels in the HFACS cannot affect higher levels. Finally, odds ratios were also calculated which provide an estimate of the likelihood of the presence of a contributory factor in one HFACS category being associated concomitantly with the presence of a factor in another category. However, it must be noted that odds ratios are an asymmetric measure and so are only theoretically meaningful when associated with a non-zero value for lambda.
RESULTS

The frequency of occurrence of the individual causal factors coded in the analysis of the 523 accidents is given in table 1. Relatively few categories had exceptionally low counts. Only the categories of 'organizational climate' (level-4); 'supervisory violation' (level-3) and 'adverse physiological state' (level-2) failed to achieve double figures.

The results reported in this paper pertain only to the instances where the PRE (when predicting the lower level HFACS category from the higher level category) was in excess of 5%.

Examination of the lambda statistic showed only one association between categories at level-4 and level-3 ('organizational process' with 'inadequate supervision') in which the PRE exceeded 5%. This association was between 'organizational process' with 'inadequate supervision' ($\lambda = 0.282$). It also had a high odds ratio, suggesting that poor supervisory practices were 13.561 times more likely to occur when associated with poor managerial processes in the air force.
Calculation of the lambda statistics showed that there was only one association between the level-3 and level-2 categories of 'inadequate supervision' and 'crew resource management' in which the PRE exceeded 5% ($\lambda = 0.281$). This relationship again had a high odds ratio, suggesting that poor supervisory practices were 12.78 times more likely to subsequently result in poor CRM.

Examination of the lambda statistics showed there to be eight associations between level-2 and level-1 HFACS categories where the PRE exceeded 5%. These were 'adverse mental states' with 'decision errors' and 'skill-based errors'; 'physical/mental limitations' with 'decision errors' and 'skill-based errors'; 'crew resource management' with 'decision errors' and 'skill-based errors'; and 'personal readiness' with the categories of 'decision errors' and 'skill-based errors'. These relationships are summarized in table 2 and described diagrammatically in figure 2. All these relationships were also associated with high odds ratios, suggesting that inadequate performance in the higher level HFACS categories was associated with much increased levels of poor performance at the lower levels.
DISCUSSION

It can be seen from the data presented in table 1 that the vast majority of HFACS categories had large numbers of instances of occurrence in the data set, which allows reasonable confidence in the pattern of results obtained. Only three categories had low frequencies of occurrence. Interestingly, these low frequency counts were not associated with any one HFACS level in particular, but it can be suggested that the low numbers may reflect either the sensitivity of the issues they address (e.g. 'adverse physiological states' and 'supervisory violations') or because they deal with a less tangible issues ('organizational climate'). In a similar vein Harris (6) noted that in the post-hoc coding of incident data the categories showing the lowest level of reliability were those that required either a great deal of inference on the part of the assessors when coding the data or the ones that dealt with more abstract concepts, such as inferring a lack of 'situational awareness' or in this case identifying 'organizational climate' as a contributory factor. It is suggested that perhaps the raters may have been disinclined to utilize this category in this study in preference to something less abstract where there was more tangible evidence available from the accident report narratives, for example 'organizational process'.

Reason (17 & 18) has suggested that there is a 'many to one' mapping of the psychological precursors of unsafe acts to the actual errors themselves, making it difficult to predict which actual errors will occur as a result of which preconditions. This research, using the
HFACS framework developed by Wiegmann & Shappell (20, 21, 22, 24, 25, 26, 27, 28) goes some way to supporting this assertion. There are associations between factors at higher organizational levels, the psychological contributory factors and the errors committed by pilots (see figure 2). However, some care needs to be taken when interpreting these relationships as the frequency counts within categories were all derived from accidents. It is unknown (and 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.

Reason (17) proposed that latent conditions are present in all systems and they are an inevitable part of organizational life. For example, resources are normally distributed unequally in organizations. The original decision on how to allocate resources may have been based on sound commercial arguments, but such inequities may create reliability or safety problems for someone somewhere in the system at some later point. This investigation showed that at HFACS level-4 ‘organizational influences’, the category of ‘organizational process’ had an association with the category of ‘inadequate supervision’ at level-3 (‘unsafe supervision’). Poor ‘organizational processes’ were associated with inadequacies in supervision and hence were ultimately, albeit indirectly, at the root of many operational errors resulting in accidents. Well-developed ‘organizational processes’ that are consistently adhered to are key to all safety management systems. The commitment to safety must come
right from the very highest levels of the organization if it is to be successful in this respect (18). Both Reason (17 & 18) and Wiegmann & Shappell (27) 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 support for this relationship. Furthermore, the odds ratios associated with ‘supervisory failures’ were over 13 times more likely to occur in the presence of a concomitant failure in the category of ‘organizational process’.

The category of ‘inadequate supervision’ was the key factor at HFACS level-3. It had a particularly strong association with the level-2 category of ‘Crew Resource Management’. The failure of senior officers in a supervisory position to provide guidance and operational doctrine to pilots through promoting good resource management practices was subsequently indirectly associated with active, operational failures. Again, the odds ratios suggested failures as a result of poor CRM practices were almost 13 times more likely to occur in the presence of a concomitant failure in the category of ‘supervisory failure’. This would seem to be a key area for risk reduction.

Reason (17) suggested that human behavior is governed by the interplay between psychological and situational factors. Several pre-conditions for unsafe acts (at HFACS level-2) show strong associations with the active failures of the operators at level-1.
These level-2 factors show Reason’s classic ‘many to one’ mapping of psychological precursors to active failures in all of the level-1 categories with the exception of the category of ‘violations’ which suggests that a completely different mechanism is at play here to cause such failures (see table 2 and figure 2).

It will be noted from the results presented diagrammatically in figure 2 that there are relatively few associations between HFACS categories at adjacent organizational levels, where the lower level categories were associated with a moderately large proportional reduction in error (lambda statistic) with a higher-level category. However, there are several key associations which strongly suggest where safety interventions are likely to have the greatest potential impact.

The results suggest that interventions at HFACS levels 1 and 2 would only have limited effect in improving overall safety. As an example, 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 facilities; oversee CRM training; 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 from the highest levels of management. Furthermore, on a ‘dollar-for-dollar’ basis, interventions at higher levels are also likely to be more cost effective in the net safety benefits they realize. Specific targets for remedial safety action should be aimed
in the areas that share the strongest and greatest number of demonstrable associations with lower levels in the organization (for example, ‘organizational process’, ‘inadequate supervision’ and ‘Crew Resource Management’. These are the categories at the root of the paths of association with other HFACS categories that have very high values for the odds ratios associated with them. This would further suggest that the greatest gains in safety benefit could be achieved by targeting these areas.

CONCLUSIONS

The large sample of accident and incidents in the present study has allowed extensive analysis of the inter-relationships between the categories and levels in the HFACS thereby providing some empirical evidence to support its theoretical structure. The accidents and incidents analyzed all occurred in the ROC Air Force thus the patterns of inter-relationships reported may be culturally specific. However there is no reason why this analytical methodology cannot be employed on other data sets to establish if the patterns observed hold good in other cultures, thereby providing further evidence to support the HFACS methodology.

This study provides an understanding, based upon empirical evidence, of how actions and decisions at higher levels in the organization promulgate throughout the ROC Air Force to result in operational errors and accidents. This has not previously been done
with data analyzed using the HFACS. There are a few, clearly defined paths that relate errors at level-1 (the operational level) with inadequacies at both the immediately adjacent and higher levels in the organization. This research draws a clear picture that go some way to support Reason’s (17) model of active failures resulting from latent conditions in the organization. Furthermore, the HFACS framework has been proven to be a useful tool for guiding accident investigations and for targeting potentially cost-effective remedial safety actions.

REFERENCES


24. Wiegmann DA, Shappell SA. Applying the Human Factors Analysis and Classification System to the Analysis of Commercial Aviation


TABLE I

<table>
<thead>
<tr>
<th>HFACS Category</th>
<th>Frequency</th>
<th>Percentage</th>
<th>Cohen's Kappa</th>
<th>Kappa Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level-4, Organizational process</td>
<td>76</td>
<td>14.5</td>
<td>0.59</td>
<td>87.4</td>
</tr>
<tr>
<td>Level-4, Organizational climate</td>
<td>4</td>
<td>0.8</td>
<td>0.44</td>
<td>96.4</td>
</tr>
<tr>
<td>Influences</td>
<td>184</td>
<td>35.2</td>
<td>0.77</td>
<td>86.4</td>
</tr>
<tr>
<td>Level-3, Supervisory violation</td>
<td>8</td>
<td>1.5</td>
<td>0.69</td>
<td>96.2</td>
</tr>
<tr>
<td>Unsafe</td>
<td>12</td>
<td>2.3</td>
<td>0.54</td>
<td>95.8</td>
</tr>
<tr>
<td>Supervision</td>
<td>24</td>
<td>4.6</td>
<td>0.71</td>
<td>94.6</td>
</tr>
<tr>
<td>Level-2, Technology environment</td>
<td>44</td>
<td>8.4</td>
<td>0.61</td>
<td>89.9</td>
</tr>
<tr>
<td>Preconditions for Unsafe Acts</td>
<td>74</td>
<td>14.1</td>
<td>0.80</td>
<td>92.2</td>
</tr>
<tr>
<td>Physical environment</td>
<td>29</td>
<td>5.5</td>
<td>0.70</td>
<td>72.3</td>
</tr>
<tr>
<td>Crew resource management</td>
<td>146</td>
<td>27.9</td>
<td>0.80</td>
<td>89.7</td>
</tr>
<tr>
<td>Physical/mental limitation</td>
<td>73</td>
<td>14.0</td>
<td>0.69</td>
<td>90.4</td>
</tr>
<tr>
<td>Adverse physiological states</td>
<td>2</td>
<td>0.4</td>
<td>0.44</td>
<td>96.4</td>
</tr>
<tr>
<td>Adverse mental states</td>
<td>184</td>
<td>35.2</td>
<td>0.75</td>
<td>86.0</td>
</tr>
<tr>
<td>Level-1, Violations</td>
<td>160</td>
<td>30.6</td>
<td>0.70</td>
<td>84.9</td>
</tr>
<tr>
<td>Unsafe Acts of Operators</td>
<td>116</td>
<td>22.2</td>
<td>0.67</td>
<td>85.1</td>
</tr>
<tr>
<td>Skilled-based errors</td>
<td>226</td>
<td>43.2</td>
<td>0.71</td>
<td>83.4</td>
</tr>
<tr>
<td>Decision errors</td>
<td>223</td>
<td>42.6</td>
<td>0.68</td>
<td>81.5</td>
</tr>
</tbody>
</table>

Table I FREQUENCY COUNTS AND INTER-RATER RELIABILITY STATISTICS FOR EACH HFACS CATEGORY FOR ALL 523 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>
<thead>
<tr>
<th>Precondition for Unsafe Acts</th>
<th>Lambda (PRE)</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>With Unsafe Acts of Operators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse mental state *</td>
<td>.269</td>
<td>4.364</td>
</tr>
<tr>
<td>Decision errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adverse mental states *</td>
<td>.283</td>
<td>4.518</td>
</tr>
<tr>
<td>Skill-based errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical/mental limitation *</td>
<td>.211</td>
<td>7.730</td>
</tr>
<tr>
<td>Decision errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical/mental limitation *</td>
<td>.164</td>
<td>4.735</td>
</tr>
<tr>
<td>Skill-based errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crew resource management *</td>
<td>.215</td>
<td>3.724</td>
</tr>
<tr>
<td>Decision errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crew resource management *</td>
<td>.195</td>
<td>3.299</td>
</tr>
<tr>
<td>Skill-based errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal readiness *</td>
<td>.058</td>
<td>3.613</td>
</tr>
<tr>
<td>Decision errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal readiness *</td>
<td>.075</td>
<td>5.231</td>
</tr>
<tr>
<td>Skill-based errors</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II**

GOODMAN AND KRUSKALL’S LAMBDA AND ODDS RATIOS SUMMARISING SIGNIFICANT ASSOCIATIONS BETWEEN CATEGORIES AT THE LEVEL OF ‘PRECONDITIONS FOR UNSAFE ACTS’ AND AT THE SUBSEQUENT LEVEL OF ‘UNSAFE ACTS OF OPERATORS’
Footnote 1

Wen-Chin Li is a Lieutenant Colonel in the Aviation Training Division, Air Force Academy, Republic of China, and is currently on sabbatical undertaking research in the Human Factors Group, Cranfield University.

Dr. Don Harris is Reader in Human Factors Engineering, Cranfield University. Requests for reprints should be sent to Don Harris, Human Factors Group, School of Engineering, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, United Kingdom.

Figure Captions

FIGURE II. PATHS BETWEEN CATEGORIES AT THE FOUR LEVELS IN THE HFACS FRAMEWORK IN WHICH THE PRE ($\lambda$) VALUE IS IN EXCESS OF 5%.

Indicates Lambda (PRE) Value in excess of 5%