Developing a Methodology for Assessing Safety Programs Targeting Human Error in Aviation

Scott Shappell, Ph.D.
Clemson University
Clemson, SC

Douglas Wiegmann, Ph.D.
Mayo Clinic
Rochester, MN

There is a need to develop an effective methodology for generating comprehensive intervention strategies that map current and proposed safety programs onto well-established forms of human error. Two separate studies were conducted using recommendations from NTSB accident investigations and several joint FAA and industry working groups. The goal was to validate a proposed framework for developing and examining safety initiatives targeting human error in aviation. The results suggest five approaches to reducing human factors associated with aviation accidents. When combined with the Human Factors Analysis and Classification System (HFACS) the resulting Human Factors Intervention Matrix (HFIX) will provide a useful tool for evaluating current and proposed aviation safety programs.

INTRODUCTION

The NTSB, FAA, and other safety organizations have committed extraordinary resources to prevent civilian aviation accidents. As a result, aviation in the U.S., particularly commercial aviation, has become one of the safest modes of transportation. Still, accidents can happen, often repeating the same sequence of events played out many times before. As a result, we are left with the regrettable truth that there are really very few “new” accidents, just different players.

So if there really are few “new” accidents, why has the aviation accident rate remained relatively stable over the last several years? Perhaps it has something to do with the current state of aviation safety. Truth be told, the industry is extremely safe and the easy fixes have been identified and remedied. What remains to be addressed is that small fraction of accidents attributable to perhaps the most complex problem facing aviation today – human error.

Unfortunately, while previous safety programs may have impacted other areas of aviation, there has been little evidence that they have had a significant impact on any specific type of human error. That is to say, the percentage of accidents associated with aircrew error (e.g., skill-based errors, decision errors, perceptual errors, and violations) has remained relatively stable since 1990 (Wiegmann & Shappell, 2003).

What this implies is that intervention strategies implemented in the 1990’s had at best ubiquitous effects on the errors and violations committed by aircrew. More likely however, there has been no sustained impact of any particular intervention program (Shappell, Detwiler, Holcomb, Hackworth, Boquet, & Wiegmann, in review). The latter should come as no surprise given that few studies of aircrew and supervisory error had been conducted using a human factors approach to accident causation. Furthermore, there has been no systematic human factors examination of the current or proposed safety programs aimed at addressing human error. For that matter a “human factors” analysis of safety programs even possible?

At least one study (Wiegmann & Rantanen, 2003) suggests that such an analysis can be performed using a set of standards derived from the same body of literature used to develop HFACS. In their book Wiegmann and Shappell (2003) describe an intervention taxonomy clustered around four broad categories:

1. Environment (control of temperature, noise, vibration, lighting, etc)
2. Human (personnel selection, incentives, training, teamwork, communication, etc.)
3. Machine (engineering design, capacity, etc.)
4. Task (ordering/timing of events, procedures, standardization, etc.)

Using this framework, Wiegmann and Rantanen (2003) examined a variety of technologies developed by NASA’s aviation safety program. From energy absorbing seats, restraints, and structures to synthetic vision, each safety program was classified within one of the four intervention categories. Their initial classification revealed that NASA’s primary intervention strategies targeted the machine rather than the human, environment, or task. Furthermore, when they examined NASA technologies using the HFACS framework, it was determined that nearly half of the technologies that NASA was developing were rated as having no impact on aircrew error. Those that did primarily targeted decision errors, by providing better information, automation, and training. An even smaller percentage of the technologies were aimed at aircrew error in general and only one of the products primarily targeted skill-based errors – the number one problem facing both commercial and general aviation.

Purpose

This report describes two studies that build upon the methodology originally described by Wiegmann and Shappell (2003) and used by Wiegmann and Rantanen (2003) with NASA safety programs. The first study describes an independent validation of the four intervention methodologies using safety recommendations from the NTSB. The second describes the examination of proposed FAA aviation safety programs using a prototype intervention matrix that maps the unsafe acts of operators (i.e., skill-based errors, decision errors, perceptual errors, and violations) onto several intervention approaches.

STUDY 1:
ANALYSIS OF NTSB RECOMMENDATIONS

Investigating accidents, identifying potential interventions, and issuing safety recommendations are central to any safety program. Ideally, safety recommendations, when adopted by organizations, will positively influence future operations in the field and thereby improve overall system safety. However, recommendations are just that … *recommendations* and as such are not always adopted. Moreover, they are often based solely on isolated events or at best a few events over a very short period of time rather than more global analyses of the system as a whole. While these interventions may solve a local or single-point problem, they often do not have far-reaching impact.

Further complicating matters, many domains like aviation and their corresponding safety boards have traditionally strong relations with quantitative disciplines like engineering and physics. While these organizations may be especially adept at dealing with mechanical issues they tend to be less robust when dealing with organizational or human-centered aspects of accidents such as human error, organizational failure, communication, and risk assessment (Stoop, 2003).

Recognizing this, the NTSB, like many safety entities has integrated human factors experts into their organization presumably leading to recommendations that address the entire system rather than a single engineering or mechanical aspect, per se. However, employing human factors experts alone does not necessarily translate into a breadth of interventions. A reasonable question to ask then is what specific intervention approaches does the NTSB employ? In other words, does the NTSB tend to be uni-dimensional (like NASA) or multi-dimensional with regard to specific intervention approaches?

Method

NTSB Safety Recommendations

To examine this question, aviation safety recommendations associated with commercial (14 CFR Part 121 – air carrier and Part 135 – commuter) aviation accidents occurring between 1998 and 2004 were obtained from the NTSB database. Of the 147 commercial aviation accidents reports that were completed at the time of this study, 622 unique safety recommendations were identified. However, several of the recommendations consisted of compound solutions. In those cases, the original recommendation was separated into sub-recommendations yielding a revised list of 872 unique recommendations for further analysis.

Clustering Process

The recommendations were then independently clustered into categories by two based on their similarities. The analysts were not instructed to use any predefined taxonomy or classification scheme. They were simply instructed to independently assign each recommendation to categories of their choosing based upon the nature of the recommendation.

Not surprising given the vagueness of the instructions, there were some differences in the terms used by the two analysts but there were also strong similarities. Wherever disagreements occurred, the analysts were asked to discuss their clustering heuristic and to agree on a single classification scheme. In the end, all 872 recommendations were classified based on their underlying similarities by two independent analysts, who later came to a consensus on the number and labels for each of these clusters.

Results

Ultimately, the analysts generated nine unique categories of recommendations, which included the design of parts/displays, procedures, communication, training, requests to conduct focused studies, rules, manuals, inspection, and human resources. These nine categories were then further grouped into four larger categories based on their similarities: 1) administrative/organizational; 2) mechanical/ engineering; 3) human/crew; and 4) task/mission. Each category and their accompanying subcategories are briefly described in Table 1.

Distribution of recommendations

From a global perspective, it appears that roughly two-thirds of the recommendations were administrative/ organizational or mechanical/ engineering fixes while nearly a quarter of the recommendations were aimed at either the task or mission. Surprisingly few interventions directly targeted operators (aircrew) even though previous studies repeatedly show that more major accidents have been attributed to human error than to any other single cause (Wiegmann & Shappell 2003, Boquet, et.al., in review; Detwiler, et.al., in review; Shappell, et.al., in review). It has also been observed that wider systemic issues, including the managerial and regulatory context of aviation operations, were also mentioned in a large number of reports (Holloway & Johnson, 2004; Johnson, in review), even though this does not appear to be reflected in the accident record.

A closer examination revealed that similar to Wiegmann and Ratannen's study of NASA safety programs, design fixes constituted the largest percentage of any individual type of recommendation made by the NTSB (23.17%) - nearly twice as many as any other category. Considerably fewer recommendations were aimed at procedures, training, information management/ communication, and the other subcategories.

Summary

When examining the breadth and scope of NTSB recommendations even at this level, it appears that current aviation safety recommendations tend to focus more on improving the design of systems or some manner of organizational change rather than focusing on the individuals in the field. While these recommendations are obviously well-intentioned and often specific to a particular accident, they may be misplaced or narrow in scope. This may help explain why the percentage of accidents associated with human error has not changed over the last 15 years (Wiegmann & Shappell, 2003; Wiegmann, et. al, in press; Shappell, et. al., in review).

Table 1. Proposed categories and sub-categories of NTSB recommendations.

Administrative/Organizational

Rules/Regulations/Policies: Issuing, modifying, establishing, amending, and/or reviewing policies, rules, or regulations.
Information Management/ Communication: Improvements in disseminating, storing, archiving and publishing information. Also included are recommendations regarding collection of data, issuing information bulletins, advisory circular and reporting activity.
Research/ Special Study: Conducting research to determine the impact of recent technological advances or call for special studies to review processes, develop/validate methodologies, evaluate the feasibility of safety equipment, and/or conduct surveys.
Human Resource Management: Adequacy of staff in specific situations, the need for additional personnel, and the evaluation of individual skills of employees.

Mechanical/Engineering

Design/Repair: Specific manufacturing changes including the design of parts. Also included is the modification, replacement, removal and/or installation or repair of parts and equipment.
Inspection: Maintenance inspections, overhauling, detecting damage including day-to-day operations such as inspecting fuel, oil level, and recommended safety checks.

Human / Crew

Training: Reviewing, developing, and implementing training programs. Also included is the training of personnel in handling emergencies.

Task/Mission

Procedures: Amending, reviewing, modifying, revising, establishing, developing, and validating procedures.
Manuals: Reviewing, revising, issuing, amending, and modifying manuals, bulletins, checklists, and other instructions or guidance.

The findings of Study 1 suggest that there are at least four broad categories of interventions that appear tenable within the aviation industry: Administrative/ Organizational, Human/Crew, Mechanical/ Engineering, and Task/Procedure. These four approaches differed slightly from those previously proposed by Wiegmann and Shappell (2003) and utilized by Wiegmann and Rantanen (2003) to analyze NASA safety programs. One category that naturally surfaced from the present analysis, but was missing from the Wiegmann and Rantanen study, was Administrative/Organizational interventions. In contrast, "environmental" interventions did not appear in the current study but were present in the NASA study (Wiegmann & Rantanen, 2003).

In the end, the question is not whether or not there are three, four, five, or more approaches to identify potential accident interventions as much as there is definitively more than one. Exactly what those approaches are remains to be fully explored. However, the five approaches identified between the present study and the investigation conducted by Wiegmann and Rantanen (2003) is a reasonable first start.

STUDY 2
HFIX ANALYSIS OF JSAT/JSIT RECOMMENDATIONS

Identifying viable approaches for intervening is only the first step. The ability to map interventions onto specific types of human error is also important. In other words, simply generating a variety of interventions across several domains, whether they are human, mechanical, environmental, etc., is likely to be ineffective unless such interventions directly target the problem area.

Given that human error continues to be the largest contributor to commercial and general aviation accidents, it makes sense to map different interventions against specific error forms. What is needed is a theoretical framework that captures the underlying causal mechanisms of human error that align with the intervention approaches identified in Study 1.

Such an error framework already exists and is widely used within the aviation industry. This framework, the Human Factors Analysis and Classification System (HFACS) describes two general categories of unsafe acts that operators commit: *errors* – the honest mistakes individuals make every day, and *violations* – the willful disregard for the rules and regulations of safety[1]. Within those two overarching categories, HFACS describes three types of errors (decision, skill-based, and perceptual) and two types of violations (routine and exceptional). Each has been described extensively in previous reports (e.g., Wiegmann & Shappell, 2003).

Human Factors Intervention Matrix (HFIX)

A prototype matrix, called the Human Factors Intervention Matrix (HFIX), pits the unsafe acts individuals commit against the five different intervention approaches presented above (Figure 3). The utility of such a framework seems intuitive. For example, if one were interested in developing interventions to address decision errors, the goal would be to identify prospective interventions within each approach (i.e., organizational/administrative, human/crew, etc.), thereby ensuring that the widest array of interventions were considered. By mapping prospective interventions onto the matrix it would be readily apparent if the scope of a proposed program was uni- or multi-dimensional.

[1] A complete description of the entire HFACS framework including all 4 tiers and 19 causal categories can be found in Wiegmann and Shappell, 2003.

	Organizational/ Administrative	Human/ Crew	Technology/ Engineering	Task/ Mission	Operational/ Physical Environment
Decision Errors					
Skill-based Errors					
Perceptual Errors					
Violations					

Figure 3. The "Human Factors Intervention matriX" (HFIX).

Alternatively, a framework like HFIX could be used proactively to determine which areas an organization has "covered" and where gaps exist in the current safety program given current trends in the error data. For instance, if you knew that the largest threat to safety within your organization was skill-based errors, followed by decision errors, violations, and perceptual errors (as is the case with general and commercial aviation in the U.S.), HFIX could be used to determine if your proposed and future interventions have the potential to address those needs and which areas are currently being targeted.

Hence, the purpose of Study 2 was to determine if such an approach could be used within the FAA and which types of human error might be affected by current and future interventions. In a sense, this analysis would provide a "benchmark" of current FAA intervention efforts.

FAA *Safer Skies* Initiative

As part of the FAA's *Safer Skies* initiative, several Joint Safety Analysis Teams (JSATs) and Joint Safety Intervention Teams (JSITs) were formed from experts in the government, private sector, industry, and academia to address civilian aviation accidents. Particularly germane to this study were outcomes derived from the JSAT and JSIT teams examining accidents associated with:

- Controlled flight into terrain
- Approach and landing
- Loss of control
- Runway incursions
- Weather
- Pilot decision making

Method

JSAT and JSIT recommendations

Final reports from the selected JSAT and JSITs were collected by researchers at the Civil Aerospace Medical Institute (CAMI). After eliminating duplicate recommendations, a comprehensive list of recommendations was compiled electronically for classification. The final list of 614 unique recommendations was then randomized to reduce bias.

Categorization of the Data

Eighteen Master of Aeronautic Science candidates were recruited from Embry-Riddle Aeronautical University to classify the recommendations. Each had experience in the aviation community as either a pilot, maintainer, or at an administrative level and all had successfully completed a minimum of one graduate level human factors course.

After a roughly 4-hr training session on the HFACS and HFIX frameworks, participants were randomly assigned to one of six groups. Each 3-person team was then randomly assigned roughly 1/6th of the recommendations to classify. Each team member was instructed to independently classify each recommendation into only one of the five intervention approaches (i.e., organizational/administrative, human/crew, mechanical/ engineering, task/mission, or physical environment). In addition, they were instructed to identify any HFACS Unsafe Acts categories they felt the intervention would impact.

After the initial rating, team members were permitted to discuss their classification within their group to resolve any differences. A final, consensus, classification for each recommendation was then provided for further analysis.

Results

The results of both classification tasks are presented in Figure 4. Several observations can be made from the data. First, as with the NTSB recommendations large percentages (36.6%) of JSAT/JSIT recommendations were directed at organizational/administrative levels. Likewise, several (22.2%) of the recommendations involved technological/engineering approaches. However, unlike the NTSB where relatively few recommendations targeted the human, nearly 1/3 of those obtained from the JSAT/JSITs did so.

INTERVENTIONS

HFACS UNSAFE ACTS	Organizational/ Administrative	Human/ Crew	Technology/ Engineering	Task/ Mission	Operational/ Physical Environmen	
Decision Errors	25.9%	26.9%	13.4%	5.7%	0.8%	72.6%
Skill-based Errors	13.8%	20.7%	12.5%	2.4%	0.2%	49.7%
Perceptual Errors	9.0%	12.7%	12.4%	2.4%	1.1%	37.6%
Violations	14.3%	7.8%	2.8%	1.3%	0.7%	26.9%
	36.6%	32.6%	22.2%	7.3%	1.3%	

Figure 4. Percentage of JSAT/JSIT recommendations classified by intervention approach and specific HFACS unsafe act addressed.

When examining the HFACS classifications, remember that unlike the specific approaches to accident interventions where subjects were instructed to select only one approach, they were permitted to select all of the HFACS Unsafe Act categories that they felt would be impacted by a given recommendation. Therefore, unlike the intervention approaches whose percentages added up to 100%, the total percentages associated with each Unsafe Act category did not.

Perhaps not unexpected, interventions aimed at decision errors were associated with nearly three out of every four JSAT/JSIT recommendations examined. In contrast, skill-based errors were associated with roughly 50% of the recommendations followed by perceptual errors (37.6%) and violations (26.9%). Of note, these numbers are slightly different than the percentage of accidents associated with each type of error where skill-based errors account for between 45-80% of the accidents depending on whether one is talking about commercial or general aviation respectively. Likewise, roughly 1/3 of the accidents were associated with decision errors yet 72.6% of the interventions have some component that will potentially affect pilot decision making.

This is not to say that there should be a one-to-one relationship between the percentage of accidents associated with a given error category and the percentage of recommendations aimed at addressing these errors. After all, it may take more effort to address one error form than another, or more interventions may naturally address pilot decision-making. In either case, the global analysis presented here suggests that additional review of this apparent incongruity is necessary.

Perhaps more important however, was the mapping of each intervention within both the intervention approach and the HFACS Unsafe Acts category (Figure 4). As can be seen (white boxes), three of the 20 possible boxes (organizational/ administrative by decision error, human/crew by decision error, and human/crew by skill-based error) contained 20% or more of the JSAT/JSIT interventions. On the surface this appears to reflect a narrow rather that a broad approach to accident intervention/mitigation by these committees. Not that the interventions contained within these categories will not be effective, just that other, potentially equally viable, interventions may have been overlooked.

It is interesting to note however, that if one examines those boxes that contained between 10-20% of the possible interventions, nearly all of the remaining boxes among the organizational/ administrative, human/crew, and technology/ engineering approaches were included. What were not accounted for were human/crew and technology/engineering approaches dealing with violations of the rules and regulations. Obviously, these approaches might prove beneficial if an organization wanted to modify or curtail a particular unsafe pattern of behavior (e.g., flight into instrument conditions while on a visual flight rules flight plan) through training or technological means.

More notable was the general lack of interventions targeting the specific task/mission of the aircrews or the environment they are faced with. Perhaps a closer examination of the operations these aircrews are engaged in or the environments they are expected to operate in is warranted. In any event, there may have been options that were not considered by these select committees along these lines.

Summary

The results from Study 2 using JSAT/JSIT interventions, although clearly more multi-dimensional than NASA's safety programs, still did not appear to fully address the current accident trends in commercial and general aviation. At least on the surface, it appears that there are weaknesses in the safety program that should be addressed.

For example, there was an apparent bias toward interventions aimed at pilot decision making, particularly those utilizing organizational and human approaches. While this is not inherently bad, previous HFACS analyses suggest that additional effort should be placed on skill-based errors and violations, two areas that appear underrepresented given current trends in the accident data.

Also noteworthy, few interventions attempted to modify/change the task itself or the environment. A closer examination of the actual types of errors may suggest changes in routes people fly or the actual type of flights being flown.

CONCLUSIONS

While HFIX may prove useful when generating comprehensive intervention strategies, organizations simply cannot implement every recommendation. Other factors may need to be considered before employing a given intervention. Factors such as *effectiveness* (i.e., what is the likelihood that it will work?), *cost* (i.e., Can the organization afford the intervention?), *feasibility* (i.e., how easy will it be to implement the intervention or does it actually exist?), and *acceptability* (i.e., will the workforce accept the proposed intervention?) all must be considered.

As such, HFIX may actually be $HFIX^3$ mapping human error against the intervention approaches and evaluation criteria (Figure 5). Although it may appear complex, in reality organizational decision makers utilize this third dimension all the time. However, even without this third dimension, the mapping of specific interventions onto a matrix that combines the five intervention approaches with general categories of human error can provide a broader perspective of the FAA's safety programs.

In sum, safety recommendations are not simply based on empirical findings surrounding an accident. Rather, they are based on one's philosophical view of what actually constitutes a "cause" of an event, coupled with one's own biased view of how changes in human or system behavior can even be accomplished. Therefore, thinking “outside the box" when it comes to generating intervention strategies is extremely difficult to do; yet failure to do so can leave other potentially viable and effective alternatives unexplored. Ideally, the HFIX framework will help safety professionals do just that.

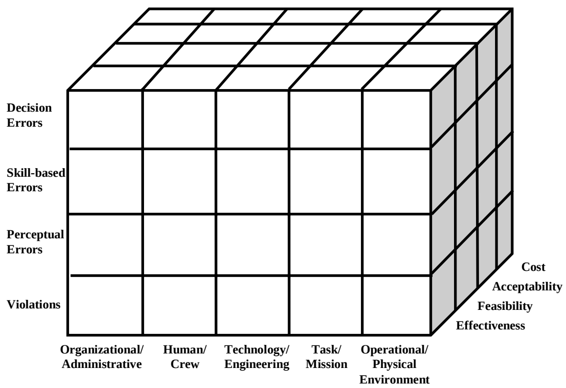

Figure 5. The HFIX^3 framework.

REFERENCES

Boquet, A., Detwiler, C., Holcomb, K., Hackworth, C., Wiegmann, D., and Shappell, S. (in review). Human error associated with air medical transport accidents in the United States. Office of Aerospace Medicine Technical Report. Civil Aerospace Medical Institute, Oklahoma City, OK 73125.

Detwiler, C., Hackworth, C., Holcomb, K., Boquet, A., Pfleiderer, E., Wiegmann, D., and Shappell, S. (in review). Beneath the tip of the iceberg: A human factors analysis of general aviation accidents in Alaska verses the rest of the United States. Office of Aerospace Medicine Technical Report. Civil Aerospace Medical Institute, Oklahoma City, OK 73125.

Holloway, C., and Johnson, C. (2004). Distribution of Causes in Selected U.S. Aviation Accident Reports Between 1996 and 2003. *22nd International System Safety Conference, Providence, Rhode Island, Aug2-6. 2004.*

Johnson, C. (in review). On the over emphasis of human error as a cause of aviation accidents: systemic failures and human error in US NTSB and Canadian TSB aviation reports 1996-2003. Submitted to *Ergonomics.*

Reason, J. (1990). *Human Error*. New York: Cambridge University Press.

Shappell, S. and Wiegmann, D. (2003). Reshaping the way we look at general aviation accidents using the human factors analysis and classification system. *Proceedings of the Twelfth International Symposium on Aviation Psychology*, Dayton, Ohio.

Shappell, S. and Wiegmann, D. (2004). HFACS Analysis of Military and Civilian Aviation Accidents: A North American Comparison. *Proceedings of the Annual Meeting of the International Society of Air Safety Investigators*, Gold Coast, Australia.

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., and Wiegmann, D. (in review). Human error and commercial aviation accidents: A comprehensive, fine-grained analysis using HFACS. Office of Aerospace Medicine Technical Report. Civil Aerospace Medical Institute, Oklahoma City, OK 73125.

Stoop, J. (2002). Accident investigations: Trends, Paradoxes and opportunities. *International Journal of Emergency management.* 1(2), *170-182.*

Wiegmann, D., Boquet, A., Detwiler, C., Holcomb, K., Faaborg, T., and Shappell, S. (in press). Human error and general aviation accidents: A comprehensive, fine-grained analysis using HFACS. Office of Aerospace Medicine Technical Report. Civil Aerospace Medical Institute, Oklahoma City, OK 73125.

Wiegmann, D., and Rantanen, E. (2003). Defining the Relationship Between Human Error Classes and Technology Intervention Strategies. *Final Technical Report AHFD-03-15/NASA-03-5.*

Wiegmann, D. and Shappell, S. (2001). Human error analysis of commercial aviation accidents: Application of the Human Factors Analysis and Classification System (HFACS). *Aviation, Space and Environmental Medicine*, *72*, 1006-1016.

Wiegmann, D. and Shappell, S. (2003). *A human error approach to aviation accident analysis: The human factors analysis and classification system.* Burlington, VT: Ashgate.