CLASSIFYING CREW PERFORMANCE FAILURES IN COMMERCIAL AVIATION ACCIDENTS: CAN WE GET THE NUMBERS RIGHT?

Terry L. von Thaden & Kelly S. Steelman
Human Factors Division
University of Illinois at Urbana-Champaign
Champaign, IL

The purpose of accident investigation is to identify and organize a description of the process through which interacting elements and conditions result in an unfortunate event. Unfortunately, if accident investigators lack a cohesive definition of human factors issues, the investigation into human performance may be unsystematic, incomplete, or leave significant points unresolved. This may, in turn, hinder scientific analysis of the data from which future principles or procedural rules can be revealed. Case in point is the classification of Crew Resource Management (CRM) accidents in aviation. We analyzed the National Transportation Safety Board's (NTSB) aviation accident data from 1990-2000 for CRM failure, based in part on James Reason's theory of system disasters. We identified 95 human performance accidents involving flightcrews in large air carrier operations (FAR Part 121). Examination of the NTSB-assigned sequence of events revealed that 27% of accidents involved a CRM-related latent error; however, an additional 16% involved active errors committed by the flightcrew, not additionally coded as a latent crew failure. Without consistent and systematic classification of error from the initial investigation, the final analysis of CRM-related failures may ultimately be underrepresented, and thus the concept of CRM may remain ill-defined.

Background

The purpose of accident investigation is to identify and organize a description of the process through which interacting elements and conditions result in an unfortunate event. Unfortunately, if accident investigators lack a cohesive definition of human factors issues, the investigation into human performance may be unsystematic, incomplete, or leave significant points unresolved. This may, in turn, hinder scientific analysis of the data from which future principles or procedural rules can be revealed, which can lead to better training, system recovery, and possible technological safety advances. These unresolved human performance factors leave the doorway open for similar occurrences to continue breaching system safety with similar disastrous results. This paper will look at the issue of classification by analyzing crew performance failures in commercial aviation accidents.

Failure Classification

James Reason distinguishes two types of system error, active and latent. Active errors are those whose effects are directly experienced, while latent errors involve adverse consequences that may lie dormant in the system becoming apparent only when they combine with other facets to penetrate a system's defenses.(Reason, 1990).

In order to quantify the number of human error-related accidents, as well as to identify the specific modes of failure, a comprehensive taxonomy of human error is required. The Human Factors Analysis and Classification System (HFACS) (Shappell and Wiegmann, 2000; Wiegmann & Shappell, 2003), based on Reason's Swiss Cheese Model of human error provides a systematic methodology of analyzing and classifying accident data. HFACS was developed as a framework of human error through which the causal factors of aviation accidents can be systematically investigated and understood. In addition to classifying the active failures, or unsafe acts, that may ultimately lead to an accident, HFACS also provides a taxonomy for identifying the latent failures, or preconditions, that contribute to accidents.

HFACS classifies the accident sequence-of-event data into active and latent failures. The first level of HFACS, labeled Unsafe Acts, is composed of active failures - errors and violations that ultimately lead to an accident. These include decision errors, skill based errors, perceptual errors, routine violations and exceptional violations. Latent failures, which may lie dormant in the system, are broken into three categories: Preconditions for Unsafe Acts, Unsafe Supervision, and Organizational Influences. Preconditions for Unsafe Acts covers both the substandard conditions and substandard practices of the operator; Unsafe Supervision refers to those causal factors which can be attributed to failures at the supervisory level; and Organizational Influences are those failures which are committed at the organizational or upper level. For this paper our focus is only on the categorization of the unsafe acts of the operators and the preconditions for these unsafe acts,

specifically those involving Crew Resource Management.

Crew Resource Management

Crew Resource Management (CRM), has been identified as a precursor to the unsafe acts of pilots. In HFACS, failures in CRM are classified under Preconditions of Unsafe Acts – Substandard Practice of Operators. This category was established "to account for the occurrences of poor coordination among personnel." (Shappell & Wiegmann, 2000, p.8).

During the 1970s, there was a number of high profile commercial aviation accidents involving failures in communication and coordination on behalf of the flightcrews. This resulted in increased attention to the human error aspects of failures of interpersonal communications, decision-making, and leadership in aviation operations, rather than stick-and-rudder skills (Murphy, 1980). As a consequence, the joint industry and National Aeronautics and Space Administration's (NASA) workshop, *Resource Management on the Flightdeck* was sponsored by NASA in 1979 (Cooper, White, & Lauber, 1980). This conference was an outgrowth of NASA research into the causes of air transport accidents, and the workshop focused on the reduction of "pilot error." The label *Cockpit Resource Management* was applied to the process of training crews to reduce "pilot error" by making better use of the human component, or resources, on the flightdeck (Cooper, White, & Lauber, 1980). Many of the air carriers represented at this meeting committed to developing new training programs to enhance the interpersonal aspects of flight operations, with United Airlines taking the lead developing a formal training program in 1980. Since then, both civilian and military CRM training programs have spread in the United States and around the world (Prince & Salas, 1993). CRM has gone through numerous permutations in the subsequent years, changing from *cockpit* resource management to *crew* resource management to reflect the incorporation of the entire gamut of personnel responsible for the safety of a flight.

The original focus of CRM was based on human factors concepts such as Edward's (1972) SHEL model, a conceptual systems model to explain the interactions between the software, hardware, liveware and environment that comprise aviation operations. Other concepts include the Transcockpit Authority Gradient (Edwards, 1975), which expanded upon the SHEL model by defining the optimal working crewmember relationship with the captain's role neither over nor under emphasized.

A necessary and vital component of CRM is Cockpit Task Management (CTM), which entails the four primary functions of piloting an aircraft: aviation, navigate, communicate, and systems management. CTM has been prioritized as first aviate, then navigate, then communicate, then administrate systems management. As characterized by Abbott (1993) the task of managing a flight involves primary control of the aircraft (aviation); navigation; communications concerned with all flight deck interchanges, including interactions with air traffic control and other crewmembers; and systems management, which concerns the system within which the flightcrew interacts that must be monitored or managed as the tasks are executed. "Task management is the first-level function of managing tasks and associated resources involved in conducting the mission. This is both a supervisory and a supporting function to the other three major flight deck functions. This function involves monitoring, scheduling, and allocating the tasks and task resources between and for each major function" (Abbott & Rogers 1993). While CTM is a major principle in technical skills proficiency for each individual member of a flightcrew enabling them to perform the duty of flying the aircraft (active), coordinated crew task management has historically been granted the least amount of attention, though it is the prevailing embedded activity on the flightdeck.

Jensen defines CRM in terms of pilot judgment as, "The application of aviation judgment to a multi-person crew environment" (Jensen, 1995, p. 6). He further states, "It is helpful to apply the term 'judgment' when referring to single pilot operations and CRM when referring to multi-person crews in aviation. CRM adds the social dimension to judgment and requires a strong emphasis on interpersonal communication. In every other way, the two terms mean the same" (p. 6).

Helmreich and Foushee (1993) developed a conceptual model of CRM taking into account group processes and performance. They define this as the application of human factors in the aviation system which includes "optimizing not only the person-machine interface and acquisition of timely and appropriate information, but also interpersonal activities, including leadership, effective team formation and maintenance, problem-solving, decision-making, and maintaining situation awareness" (p.4).

The most recent Federal Aviation Administration

(FAA) definition refers to CRM as "the effective use of all available resources: human resources, hardware, and information. Other groups routinely working with the cockpit crew, who are involved in decisions required to operate a flight safely, are also essential participants in an effective CRM process". (FAA Advisory Circular 120-51E, 2004, p.2).

FAA Advisory Circular (AC) 120-51 presents guidelines for developing a CRM training curriculum. Within this AC is a set of behavioral markers set forth by the FAA which include:

- Communications Processes and Decision Behavior
  - Briefings
  - Inquiry/Advocacy/Assertion
  - Crew Self-critique
  - Communications/Decision
- Team Building and Maintenance
  - Leadership/Followership/Concern for tasks
  - Interpersonal Relationships/Group Climate
- Workload Management and Situation Awareness
  - Preparation/Planning/Vigilance
  - Workload Distributed/Distractions Avoided.

Crew Resource (Mis)Management, as defined by HFACS, encompasses good communication skills and coordination among members of the flight crew (Wiegmann & Shappell, 2003). This includes coordination within the cockpit as well as between the aircraft and air traffic control, maintenance control, and other support personnel that may occur before, during and after the flight. Shappell and Wiegmann outlined a non-inclusive list of CRM errors for accident analysis:

- Failed to back up,
- Failed to communicate/coordinate,
- Failed to conduct adequate brief,
- Failed to use all available resources,
- Failure of leadership, and
- Misinterpretation of traffic calls.

While these failures do not make up a verbatim list of the behavioral markers included in the FAA's definition of CRM, there is certainly considerable overlap.

Aviation Accident Investigation

In the Unites States the National Transportation Safety Board (NTSB) is responsible for conducting independent investigations of all civil aviation accidents to determine their probable cause. Each investigator's expertise varies according to their background, which can be represented as technical, operations, or human performance. The NTSB determines the probable cause through gathering factual evidence and holding meetings to gain consensus on determining a sequence of events leading to and resulting in an incident or accident. The sequence of events data contains information about what happened (findings), when it occurred in the chain of events, whether it is determined to be a cause or a factor, and in the case of human factors events, personnel and their relationship to the event.

At times, investigators may judge factors differently given their individual background and training. In the case of accidents involving flightcrews there appears to be considerable variation in the way these accidents are analyzed.

Perhaps some variation in assigning cause factors to multi-pilot operation accidents lies in the confusion around what constitutes a CRM error. This relates to the ambiguity of the CRM definition itself. Lauber's definition describes very generally what CRM means, but not the specific skills through which it can be achieved nor the behaviors which cause it to breakdown. These skills and behaviors can be determined circuitously by looking at the specific behavioral markers that have been set forth by the FAA. Where this confusion is most apparent however, lies in the divide between CRM and CTM. The basic concepts as defined under AC 120-51E maintain that CRM is "based on an awareness that a high degree of technical proficiency is essential for safe and efficient operations. Demonstrated mastery of CRM concepts cannot overcome a lack of proficiency. Similarly, high technical proficiency cannot guarantee safe operations in the absence of effective crew coordination" (p. 5). This ambiguity between what is a latent crew error and an active unsafe act on behalf of the individual crewmember becomes more apparent upon the analysis of NTSB investigated accidents.

Method

We analyzed the NTSB's data of over 46,000 incidents and accidents for ten-year period from 1990–2000. Of these 46,000 incidents and accidents, 12,222 represented commercial aviation accidents. Seven hundred and seventy of these accidents were determined to have human factors causes. From this, 95 accidents corresponded to FAR Part 121 scheduled air carrier operations. Each of these 95 accidents corresponds to flight duties performed by multi-pilot crews. A comprehensive analysis of these accidents was conducted by analyzing the HFACS-coded failures for both CRM coded and non-CRM coded accidents. Each HFACS coding represents a distinct categorization of the NTSB assigned cause

factors. Each accident can have multiple HFACS codes but each coding must correspond to the discrete NTSB assigned cause factor in the sequence of events.

Of the 95 Part 121 accidents, 26 (27%) were attributed to a breakdown in CRM, while 69 (73%) were not attributed to CRM factors. We examined the NTSB assigned findings in the sequence of events data to determine how CRM and other codes were established for each accident as they pertain to the flightcrew by analyzing the event and the involved personnel identified by the NTSB.

Results

Among the 26 CRM coded accidents, 33 cause/factors were coded. Of these, 44% involved crew group coordination (CGC), 22% involved improper or inadequate supervision by the pilot-in-command, 10% involved inadequate monitoring, 6% involved the failure to use or comply with a checklist, and 6% were due to inadequate preflight (see Figure 1). The remaining errors were the result of improper planning decisions, inaccurate weight and balance, inadequate weather evaluation, and failure to comply with unsafe/hazardous condition warnings.

CAUSE/FACTOR	PERCENT
Crew/Group Coordination Inadequate Not Attained Not Maintained Not Performed	44%
Supervision Improper/Inadequate	22%
Monitoring Inadequate	10%
Aircraft Preflight Inadequate	6%
Checklist Not Complied With/Not Used	6%
Aircraft Weight and Balance Inaccurate	3%
Planning/Decision Improper	3%
Unsafe/Hazardous Condition Warning Not Complied With	3%
Weather Evaluation Inadequate	3%

Figure 1. Overview of CRM cause factors associated with 26 Part 121 accidents.

A comprehensive analysis of all 95 accidents (both CRM and non-CRM related) reveals that 85% of CRM failures occur in combination with at least one unsafe act, while 15% do not have an associated unsafe act. Additionally, many non-CRM coded accidents are attributed to unsafe acts by the *flightcrew*, yet they are not coded as CRM accidents. Twenty-six percent of the non-CRM coded accidents were attributed to at least one failure by the *flightcrew* but since each cause/factor in the assigned sequence of events can be assigned only one corresponding HFACS code, an event involving a decision, skill-based, or perceptual error or violation cannot be dually coded as a CRM failure even if it is committed by the flightcrew since the NTSB did not assign another coding in the sequence of events. In the data, sometimes these are coded as CRM and other times, not.

Analysis of the sequence of event data for all of the 95 accidents in these multi-pilot operations revealed that an additional 50% of the accidents met the criteria for CRM accidents, yet could not be coded as CRM in HFACS since the NTSB sequence of event modifiers preclude it (Figure 2). When this is assessed in terms of the accidents coded as non-CRM (N=69), as high as 70% of the accidents classified as non-CRM may be miscoded. Examples of this include: *(In Flight) Planning/ Decision Improper, PIC; Wrong Runway Selected, PIC. E*vents such as these occurred with enough time elapsed between the captain's decision and the accident for another crewmember to correct the situation. Other events include: *Weather Briefing Inadequate, PIC; Ice/Frost Removal From Aircraft Not Performed, PIC.* As part of a crew's usual and customary preflight preparations these items are CRM issues to be checked and verified by crewmembers. Even the multiple coding of a coordination failure as*: Checklist Not followed, Pilot in Command, Checklist Not followed, Second Pilot, Checklist Not followed, Flight Engineer* leads to each event being coded for unsafe acts on behalf of each separate crewmember, but not an overall coding of CRM.

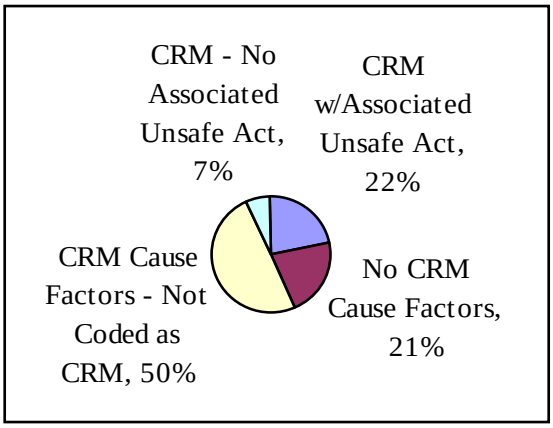

Figure 2. Itemization of accidents in Part 121 multi - pilot operations.

When the data are considered in this light, clearly these multi-pilot accidents could very well be classified as CRM accidents, yet inconsistencies in the way they are initially classified leads to the incongruity in their classification with a human factors taxonomy.

Assessing these accidents one step further we found that 7% of the CRM-coded accidents were correctly coded as CRM, yet had no associated Unsafe Acts. Forty-three percent of these accidents appeared to be consistently identified in the NTSB assigned coding, lending them to appropriate HFACS categorization as CRM accidents with their associated Unsafe Acts or as non-CRM accidents.

Conclusions

CRM covers a wide range of crew behavior. With a broad but not explicit definition of CRM and without the ability to re-code events, it is difficult to accurately assess the role that CRM plays in aviation accidents. Although analyses of HFACS-coded data reveal that CRM is a factor in aviation accidents, the actual frequency of CRM-related events may be, in fact, under-represented. With the possibility of as many as 57% of Part 121 scheduled operations accidents that may be improperly coded, this issue clearly needs more research to understand the training and judgment of the initial assessors.

Knowledge of underlying accident factors is an assumed specialization of accident investigators. Each investigator has expert skill and each has his own specific technique. Across a series of accidents and across multiple investigators, the association of accident cause/factor data may not be well documented. This in turn may lead individual investigators to piece together entirely differing cause/factors in the sequence of event data. Due to this inconsistency, even a rigorous taxonomy cannot produce consistently valid results. In turn, the opportunity to contribute to the body of data from which future training guidelines, procedures, or regulations can be derived is lost in translation. The result is that the investigative process, for all its good intent and laboring, may come up short and occasionally contradictory. It appears, then, that the nature of assigned sequence of events may not be compatible with a taxonomy used in error classification. Therefore a thorough representation of underlying accident cause/factors may not be available.

References

Abbott, T. (1993). Functional categories for future flight deck designs (NASA Technical Memorandum TM- 109005). Hampton, VA: NASA Langley Research Center.

Abbott, T. S., & Rogers, W. H. (1993). Functional categories for human-centered flight deck design. *Proceedings of the 12th Digital Avionics Systems Conference*. New York: AIAA/IEEE.

Cooper, G. E., White, M. D., and Lauber, J. K. (Eds). (1980). *Resource management on the flightdeck: Proceedings of a NASA/industry workshop (NASA CP-2120).* Moffett Field, CA: NASA-Ames Research Center.

Edwards, E. (1972). Man and machine: Systems for safety. In *Proceedings of the British Airline Pilots Technical Symposium*. Pp. 21-36. London: BAPA.

Edwards, E. (1975). *Stress and the airline pilot.* (paper presentation) British Airline Pilots Medical Symposium, London

Federal Aviation Administration (2004). Crew resource Management Training. Advisory Circular 120-51E. Washington DC: Author.

Helmreich, R. and Foushee, C. (1993). Why crew resource management? In E. L. Weiner, B. G. Kanki, & R. L. Helmreich (Eds.). *Cockpit resource management*. (pp. 3-45). San Diego: Academic Press.

Jensen, R. (1995). *Pilot judgment and crew resource management*. Burlington: Ashgate.

Murphy, M. (1980). Review of aircraft incidents. *In Cooper, G. E., White, M. D., and Lauber, J. K. (Eds). Resource management on the flightdeck: Proceedings of a NASA/industry workshop (NASA CP-2120).* Moffett Field, CA: NASA-Ames Research Center.

Prince, C., and Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. L. Weiner, B. G. Kanki, & R. L. Helmreich (Eds.). *Cockpit resource management.* (pp. 337–366). San Diego: Academic Press.

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

Shappell, S. and Wiegmann, D. (2000). The Human Factors Analysis and Classification System – HFACS. Federal Aviation Administration, Office of Aviation Medicine Report (DOT/FAA/AM-00/7).

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