A New Tool for Investigating and Tracking Human Factors Issues in Incidents

Capt. Sharon G. Jones, Ph.D. & Captain Bruce Tesmer[1]
University of Texas Team Research Project & Continental Airlines

ABSTRACT

The role in a comprehensive safety program of self-report data regarding threat, error, countermeasures, and consequences related to operational incidents is discussed. Discussion highlights the capability of incident related data, supplied by operational personnel using a new tool, to inform integrated analyses and support organizational interventions. Crew member narratives and responses to structured questions regarding demographics, circumstances, behaviors, outcomes, and corrective action proposals capture the complexity of individual events while facilitating tracking of behavioral and systemic trends. Results of questionnaire field testing by safety investigators during flight crew debriefings are presented within the framework of a threat and error model.

INTRODUCTION

The common goal of preventing 'pilot error' accidents defines the point of intersection between safety, flight operations, and flight training. Interdependence among constituencies at this interface suggests that courses of action should be coordinated in their bases and implementation for maximal effectiveness. No single initiative in any one department could protect an organization against the varied threats, errors, and non-terminal, but suboptimal, outcomes being documented through Line Operations Safety Audits (LOSA) (Klinect & Wilhelm, in press). Rather, each factor's contribution to creating hazards and its capability of forestalling or mitigating them must be determined so that an informed, efficient allocation of organizational resources and delegation of duties may be accomplished.

INTEGRATED, PROACTIVE PHILOSOPHY

As an industry, aviation is maturing in its preference for proactive intervention in potentially hazardous circumstances over post-accident remediation. Circumstances no longer need to deteriorate to the point of creating undesirable consequences before being addressed. Deterioration, like rust, is a gradual process and can be arrested and reversed at any point, assuming it is detected and interpreted as evidence of impending failure. Through focused assessment and early prioritization and intervention, organizations can enhance performance and productivity while avoiding the costs associated with accidents and incidents.

The complexities of an aviation operation cannot be captured with anything less than a comprehensive, multivariate system of interdependent, complementary, data collection programs. Performance is influenced by the operational environment as well as by individual and team attitudes and capabilities and their applications. Thus, a variety of data are essential to determining the nature, scope, and implications of threats to organizational performance and to minimizing risk at the earliest possible opportunity.

Organizational weaknesses may become evident in trainee performance, normal line operations, incidents, or accidents. Each of these realms can be uniquely informative with regard to pilot aptitude and achievement, skills application, environmental characteristics and demands, and defense adequacy. However, when data collection programs are developed and implemented in isolation, their capability to converge on a common issue is diminished greatly.

Ideally, the technical, behavioral, and contextual data collected in the interest of bolstering safety and effectiveness should fit within a common theoretical or conceptual model. While data may differ in type (electronic, narrative, or structured), source (observational, self-report, survey, interview, or equipment monitoring), perspective (objective to introspective), level of detail, level of analysis (aggregate to case study), or subject (training, operations, incidents, or accidents), they can be united by a common purpose and guiding philosophy. The advantages of an integrated approach include common terminology, compatible categorization schemes, corroboration or validation of results, and compatible goals and assessments of intervention effectiveness.

[1] Research reported here was conducted with the support of FAA Grants 92-G-017 and 99-G-004, Robert Helmreich, Principal Investigator.

INCIDENT DATA

Several programs exist for the collection of confidential self-reports on operational incidents, most notably National Aeronautics and Space Administration's (NASA's) Aviation Safety Reporting System (ASRS), the British Airways Safety Investigation System (BASIS), and American Airlines' Aviation Safety Action Partnership (ASAP). All of these have achieved high credibility and admirable participation due to their excellent accessibility, demonstrated responsiveness, and uncompromised capability to maintain reporter confidentiality. Narrative and demographic data obtained through them have instigated constructive, proactive improvements to the system. These characteristics and successes set the standard for any subsequent efforts.

Framework and Focus

Despite their advantages, the previously mentioned programs lack two important elements. The first is the benefit of a guiding conceptual framework for assembling and arranging information so that it makes sense in a total, integrated way. The second is a module for collecting human factors data in a structured format that provides a quantitative framework within which other data resources can be viewed and interpreted.

The remainder of this paper discusses the integration of a human factors oriented incident reporting program into the conceptual model of threat and error already in use as a framework for auditing operational behavior. The data collection and trend monitoring program for incidents discussed herein is intended to be a voluntary, confidential, non-jeopardy, self-report system. It has been designed to support proactive, safety-related, organizational initiatives by providing data that facilitate human factors trend monitoring.

Unmitigated, undesired states

Within the threat and error model, incidents can be viewed not only as operational outcomes, but also as the result of failure to defend against or mitigate an undesired state. Some incidents may be captured electronically through Flight Operations Quality Assurance (FOQA) data, if the associated undesired states have been defined as exceedances. However, as is the case with undesired states observed during audits of normal operations, many recorded exceedances are mitigated before evolving into incidents. Thus, organizations need an independent source of data specifically designed to capture characteristics of the rare subset of undesired states, known as incidents, neither trapped by system defenses nor mitigated through crew behavior. Study of the precursors, correlates, and consequences of such events assists identification of the limits of acceptable risk and deviation from normative behavior.

Self-Report data

Although both electronic records from FOQA and LOSA data on observed behaviors document the types of undesired states that precede incidents, neither provides information on crew member aptitudes, attitudes, intentions, expectations, perceptions, or situational interpretations. FOQA data indicate what happened, where it happened, and when it happened; LOSA data describe events as well as their behavioral and contextual correlates. Only self-report data provide the crew member perspective crucial to interpreting aircraft state and crew behavior during incidents.

Quantitative, human factors data

ASAP and ASRS rely on labor intensive analysis of voluminous, idiosyncratic, narrative data by subject matter experts. Data collection instruments used by these programs were designed for this methodology rather than for monitoring and quantitative analysis of specific human factors trends within any conceptual framework. Therefore, a new instrument for systematically recording threats, errors, countermeasures, and consequences has been developed and tested by the University of Texas Aerospace Crew Research Project in cooperation with members of the Air Transport Association (ATA) and representatives of ASRS.[2, 3]

Based on current theory and research results, the new questionnaire solicits detailed, structured data regarding the demographic, behavioral, and environmental precursors of incidents. Its strength lies in its ability to reveal the nature of operational threats and errors as they are perceived by the crew and to track the practice of CRM core behaviors and skill sets during incidents.

The data collection form contains five sections: demographics, consequences, narrative event

---

[2]The participation of Continental Airlines and American Airlines deserves special note.

[3]Copies of the questionnaire may be obtained from the authors or the University of Texas Aerospace Crew Research Website.

description, structured questionnaire, and corrective action / prevention suggestions. Variables in the demographic section of the form are compatible with ASRS data fields, and those required for ASRS submission are highlighted for respondents. Items on the structured questionnaire address the presence and consequence (positive or negative) of 34 behaviors and 22 circumstances during an incident. Crew member responses indicate the utility of CRM behaviors and other system defenses as countermeasures against threat and error and as tools for undesirable state mitigation.

Although the incident questionnaire enables crew members to document those behaviors and other defenses that "assisted recovery" from an undesirable state, the occurrence of an incident necessarily implies the absence of at least one potential mitigator of hazardous circumstances. Therefore, crew members must convey their belief that a missing factor could have served as an effective countermeasure by reporting that it "helped cause" the incident. Thus, care must be taken to determine for each causal factor whether the error is one of commission or omission. Consider, for example, a report that briefings helped cause an incident. In this case, it is more likely that a briefing which could have mitigated circumstances was absent than that a briefing was conducted which caused or exacerbated them.

PILOT TESTING THE QUESTIONNAIRE

Despite the advantages of collecting, storing, and analyzing incident data, difficulties can impede implementation of self-report programs. However, an organization can realize the benefits of the new questionnaire even in the absence of an incident reporting program by using it as a tool for post-incident debriefings of crews by safety department representatives. In fact, the form has been used for this purpose during a field test described below. This test demonstrated the manner in which the conceptual framework employed during questionnaire design enhances event analysis. During the test, two incidents at one carrier were debriefed using the new tool. Crew responses to the incident questionnaire can be used to frame each incident within the threat and error model.

The first event could not have been discovered through any source other than self-report. In it, a First Officer (FO) with 300 hours in seat and type flew a precision approach at night under instrument conditions with wind and snow before landing beyond the touchdown zone on a short, snow-covered runway. The aircraft nearly departed the pavement before stopping (only 20' remained).

In the second event, the Captain descended prior to reaching the Final Approach Fix (FAF) on a non-precision approach conducted in mixed weather conditions (ceiling 2000' broken, visibility 5 miles, and rain) at a busy airport. The descent was initiated because the Captain, who was minimally familiar with the area, mistakenly believed a visual, ground reference to be collocated with the instrument fix. The FO encouraged the descent after misinterpreting instrument indications (distance to a Flight Management System (FMS) waypoint was confused with Distance Measuring Equipment (DME) distance to the localizer facility). After being warned of an Air Traffic Control (ATC) Minimum Safe Altitude Warning (MSAW), the crew climbed back to FAF altitude before completing the approach.

The responses of each of these crews to the new incident questionnaire were used to construct Table 1, in which the two incidents are summarized according to elements of the threat and error model. The near runway over-run involved two environmental threats (weather and runway condition) and a skill-based proficiency error. The crew response was to exacerbate the situation with the consequence of this being an undesired state (an unstable approach) that was mitigated just prior to runway departure. The premature descent involved two environmental threats (a high volume airport and mixed weather conditions) and one aircraft threat (non-standard instrument displays). There was a knowledge-based proficiency error (unfamiliarity with local procedures and landmarks) as well as a procedural one (use of visual references during an instrument approach). The crew exacerbated these threats and errors by descending inappropriately, and their response resulted in an undesired state (vertical deviation) resulting in an incident that was detected by ATC.

Event	**Near Over-Run**	**Early Descent**
Threat	Environment	Environment Aircraft
Error	Proficiency	Proficiency Procedural
Error Response	Exacerbate	Exacerbate
Error Outcome	Undesired State	Undesired State
State Response	Mitigate	No Response
State Consequence	Recovery	Incident

Table 1: Two incidents seen through the lens of the threat and error model.

Both of these incidents involved proficiency errors that, when exacerbated by the crew, resulted in undesired states. However, the sources, manifestations, and consequences of inadequate proficiency differed in the two events, as did the countermeasures that might have mitigated its effects. When prompted by the questionnaire, crew members provided detailed information on contributing factors as well as accurately identifying knowledge and skills that could have served as effective countermeasures. The following is a list of potential countermeasures that can be extrapolated from the four reports received during pilot testing of the questionnaire: 1) anticipating illusions and deferring to instrument indications over external, visual references, 2) discussing during briefings the multiple reasons for abandoning a course of action and parameters for rejecting the current plan of action (including, but not limited to, inability to acquire the runway environment visually), 3) coordinating the transition from instruments to visual references, 4) crew self assessment, 5) considering crew member experience during action planning, 6) transferring aircraft control when one crew member is inadequately oriented, and 7) utilizing more conservative procedures than those required by law or company policy.

QUESTIONNAIRE BENEFITS

Benefits of the new incident questionnaire became apparent during pilot testing. First, comprehensive coverage of factors effecting performance and explicit wording of structured questions on crew behavior prompt recall as well as clarifying and expanding crewmember comprehension of events. Thus, the organization achieves a more thorough and explicit understanding of the factors that create or mitigate incidents, crew members learn the value of their input in affecting organizational change. Second, differences in perspective become apparent when both crew members report, and each individual achieves a better understanding of the other's perceptions, interpretations, and goals.

RESPONDING TO INCIDENT DATA

Whether the questionnaire is used for trend monitoring or incident debriefing, resultant data can differentiate between threats beyond human control and weaknesses in malleable resources. Inexorable threats include not only natural phenomena, such as terrain and weather, but also indelible human cognitive, affective, physiological, perceptual, and anatomical characteristics and limitations. Each of these can be prevented from becoming consequential only through strengthened defenses. Manipulable resources include human knowledge and skill and the application thereof, information systems, and equipment. Improving any of these can help minimize error commission and maximize threat recognition while optimizing responses to both threat and error.

When incident data are collected as part of a trend monitoring program, the impracticality and inadvisability of fully investigating each incident necessitate development of criteria and plans for both aggregate and case-study analysis. Data may be grouped according to commonalities in consequence, phase of flight, error type, or level of automation in use. Questionnaire item groupings provide an additional categorization scheme: policies, procedures, publications; aircraft; environment; support; cockpit crew member characteristics; use of autoflight; operational conduct; and management of flight. However, no definitive basis yet exists for valid and meaningful categorization of questionnaire data. Initial analyses may be guided by theory and literature, but only the emergence of trends in actual data will support the construction of multivariate profiles that explicate relationships between threat, error, response, and consequence.

In early program stages, the best anyone can do is count the total number of responses databased in each category, establish an average of and variance in number of variables or categories addressed in each report, and search for correlations between demographics, consequences, and contributors. Individual reports with high numbers of contributors may indicate complex situations; those with low numbers may suggest low margins of safety or inadequate defenses separating threats from potential consequences. Heavy weighting of responses in any category may be indicative of extreme, but isolated, weakness.

Any carrier will want to focus on catastrophic consequences such as injury, damage, unusual attitudes, and improper location. Issues emerging during operational audits or quality assurance trend monitoring may lead to interest in additional factors. For example, a carrier that discovers an abundance of intentional non-compliance may explore events involving deviation from normal procedures and issues of checklist use. Another that documents numerous unstable approaches might single out incidents occurring during the approach phase for analysis.

Organizations with no other guiding data might focus on issues raised by research. For example, a comprehensive review of accidents conducted by the NTSB (1994) found relationships between accidents and first time crew pairings and hours since crew member awakening. Other factors, such as automation usage and lack of proficiency, have been implicated in crew error during cross-airline operational observation (Klinect & Wilhelm, in press). The authors' incident questionnaire supports examination of the relationship of any of these factors to undesirable outcomes.

Developing Training From the Data

Historically, training footprints have been driven by FAA requirements. Although compliance with FAA standards is an industry imperative, their generic nature makes them inadequate as templates for efficient deployment of training resources to develop the knowledge and skills required to deal with particular operational environments in specific organizational cultures. Normative data from line audits document the variety and prevalence of threats with which crews must deal and of errors to which they are susceptible. Data on the factors associated with incidents indicate which threats and errors combine to create unacceptable consequences. Both kinds of data provide clues to the effects of crew behavior in preventing, mitigating, or exacerbating undesired sates. Together, they provide priorities for training and a context for application of knowledge and skills developed during training.

Armed with such data, training programs can exceed FAA standards by embedding required material or maneuvers in the types of scenarios that will require their retrieval and practice on the line. For example, the technical conduct of missed approaches can be practiced within the context of external threats that necessitate decisions to abort in actual operations. These include confusion and instability associated with visual illusions, ATC commands that lead to excessive speed and/or altitude, or ambiguity in instrument indications leading to position confusion at a high threat or international destination.

One example of a behavioral goal for training is crew recognition and rejection of suboptimal ATC commands. Although rarely addressed in flight crew training, during which ATC is simulated by the simulator operator or instructor, flight crew response to ATC commands accounted for 24% of the instances of linked threat and error (but only 6% of total threat) documented by LOSA (Klinect and Wilhelm, in press). Self-report data on crew acquiescence to suboptimal ATC commands could help clarify the nature of this linkage and establish what particular knowledge or skill crew members need to develop in order to trap this threat rather than allowing it to instigate error.

CHALLENGES IN QUESTIONNAIRE DEVELOPMENT

Development of this, and any other, research tool requires attention to content, format, and plan of analysis. Thus, the development team must include experts on subject matter, information systems design, and data analysis who work together closely. This is a difficult proposition, considering the rapid advances being made in the state of Human Factors knowledge and the complexity of database software and analytic techniques. Realistically, limitations on human and economic resources prevent many organizations from obtaining sufficient expertise in all these areas to create their own set of data collection tools. These are the reasons that resources from various organizations have been pooled in the cooperative design of this tool, which can be used as a starting point for any carrier interested in obtaining incident related data. It is expected that ongoing analyses, in addition to results from other data collection programs such as FOQA and LOSA, will prompt tailoring of the instrument to individual organizations.

SUMMARY

The threat and error model used to develop and implement LOSA has also proven to be an effective framework for the design and testing of a structured, human factors centered, incident reporting form. The availability of an overarching framework and data from a compatible, but independent, initiative to guide its design and use supported construction of a tool with unique capabilities and potential value to data collection programs. The process of designing and testing this incident reporting questionnaire demonstrated that the process of integrating safety related initiatives from the outset is both feasible and worthwhile, regardless of the methodological challenges during development and economic or political barriers to instituting non-jeopardy, self-report programs. When data collection tools and methodologies accommodate description of normal operations and incidents in common terms of threats, errors, responses, and outcomes, then effective, coordinated action can be taken to minimize threats and errors and to optimize responses and outcomes.

REFERENCES

Helmreich, R.L., Klinect, J.R., and Wilhelm, J.A. (this volume). Models of Threat, Error, and Response in Flight Operations. In Jensen, R.S. (Ed.), *Proceedings of the Tenth International Symposium on Aviation Psychology*, The Ohio State University.

Klinect, J.R., Wilhelm, J.A., Helmreich, R. L. (this volume). Threat and Error Management: Data from Line Operations Safety Audits. In Jensen, R.S. (Ed.), *Proceedings of the Tenth International Symposium on Aviation Psychology*, The Ohio State University.

National Transportation Safety Board (1994). *A review of flightcrew-involved, major accidents of US air carriers, 1978 through 1990* (Safety study NTSB/SS-94/10). Washington, DC: author.

University of Texas Crew Research Project Website: www.psy.utexas.edu/psy/helmreich/nasaut.htm