The Dissertation Committee for James Ray Klinect certifies that this is the approved version of the following dissertation:

Line Operations Safety Audit: A Cockpit Observation Methodology for Monitoring Commercial Airline Safety Performance

Committee:

Robert Helmreich, Supervisor

Clarke Burnham

Randy Diehl

Daniel Maurino

Ashleigh Merritt

Line Operations Safety Audit: A Cockpit Observation Methodology for Monitoring Commercial Airline Safety Performance

by

James Ray Klinect, B.A.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

The University of Texas at Austin
August, 2005

Dedication

To my patient and supporting wife, Stephanie. We finally did it!

Acknowledgements

There are so many people to acknowledge who have contributed to this dissertation and my post-graduate career. First and foremost, I am deeply grateful to Bob Helmreich. You gave me an opportunity and let me run with it. For that, I will be forever indebted. Your pathological optimism, kindness, and guidance have influenced what I am today. I sincerely hope that I have made you proud with my research.

I would also like to thank my dissertation committee members, Clarke Burnham, Randy Diehl, Ashleigh Merritt, and Dan Maurino. Ashleigh, your honest critique of my ideas and writing advice were priceless. Most of all, thanks for being such a great friend. I would be remiss if I didn't say a few words about Dan Maurino. Dan, you are without question, my role model in the aviation industry. I appreciate everything that you have done for LOSA but especially, everything that you have done for me. Also, a special thank you goes to my dissertation support crew, George Schorn and Kay Henry, for their proof-reading and to Shane Allua for her statistical advice.

I also appreciate the support that I have received from The University of Texas Human Factors Research Project over the years, particularly Randy Law, Bill Hines, Dave Musson, Michelle Harper, and Chris Henry. Most of all, I would like to acknowledge my fellow Texan and surrogate mother, Lou Montgomery. I cannot put into words what you have done for Stephanie and me. All I can say is thank you, thank you, and thank you.

This dissertation does not happen without my partners at The LOSA Collaborative. John Wilhelm, I consider you to be my graduate coach and "go to" guy when I need solid insight. You did so much for me personally while you were at UT and even more, outside the "40 Acres". Thank you for always being there. Pat Murray, you

are a godsend. I consider you to be one of my closest confidants and inspirations. I must also thank you for always being there. Stephen Ingham, I appreciate your hard work with the data and am thankful for our friendship. I would also like to tip my hat to all of the external LOSA Collaborative observers for their time spent in the cockpits.

I must express my sincere gratitude to all of the airlines, pilots, and observers that have participated in LOSA. I hope the work presented in this dissertation is one small contribution that helps you in your safety management pursuits. Along the way, I have met some amazing people that have contributed to LOSA development. While it would be impossible to list everyone in the space provided, I would like to recognize two people in particular. First, thanks go to Bruce Tesmer of Continental Airlines who took the first chance on LOSA. Your "can do" attitude gave me confidence and for this, I will always be grateful. Finally, Robert Sumwalt, you gave me passion for this industry. I admire your integrity and professionalism and it is something that I try to copy every day. Thanks for providing such a wonderful example.

Finally, I would like to thank all of my family and friends that have supported me through the years. Thank you, Dad, for your sound advice, and for teaching me how to think about things and deal with people and to my Mom, who taught me the value of creativity, striving for perfection, and never settling for anything less. Last but not least, I want to thank my wife, Stephanie. You have made some awesome sacrifices, many of which I could never make up for in several lifetimes. Thanks for your free spirit, keeping me honest, providing me with balance, and for being the best life partner. Adapting the words of Tom Petty and the Heartbreakers, there is no doubt that I'm the one that got lucky. I love you more than ever before and look forward to our time together as we turn to the next chapter of our lives.

Line Operations Safety Audit: A Cockpit Observation Methodology for Monitoring Commercial Airline Safety Performance

Publication No._____________

James Ray Klinect, Ph.D.
The University of Texas at Austin, 2005

Supervisor: Robert Helmreich

This dissertation presents a field observation method called the Line Operations Safety Audit (LOSA) designed to provide a proactive snapshot of system safety and flight crew performance *before* an incident or accident. The data indicators underlying this effort are based on a conceptual framework known as Threat and Error Management (TEM). This framework proposes that threats (such as adverse weather or aircraft malfunctions), pilot errors (such as selecting a wrong automation mode or missing a checklist item), and undesired aircraft states (such as altitude deviations or speed exceedances) are everyday events that flight crews must successfully manage to maintain safety. By having cockpit observers collect TEM data, LOSA provides an opportunity, never before realized, to understand the complex interactions among operational context, flight crew processes, and outcomes during routine flights. This type of insight benefits both airlines and researchers. For airlines, LOSA provides a diagnosis of operational performance strengths and weaknesses without relying on adverse safety events for such information. For researchers, LOSA addresses the shortage of field findings in aviation by providing TEM performance data gathered in its natural context.

LOSA has been developed and refined since 1996 with projects conducted at over 20 major international and regional airlines from 10 different countries. Using this experience as a foundation, this dissertation describes the rationale underlying LOSA as well as its methodology, data analysis strategies, and safety implications for the aviation industry. Some highlights include: a discussion of the 10 operating characteristics designed to gain pilot trust and lessen the tendency to "fake good" during an observation; the instrumentation, observer selection procedures, training objectives, and quality control checks used to enhance data reliability and validity; a multistage approach to data analysis and interpretation that demonstrates the transformation of LOSA data into knowledge that can drive airline safety management practices; and initial findings from an archive of over 2600 observations collected during the last five years. The dissertation concludes with a discussion of current regulatory, pilot association, and airplane manufacturer support for LOSA, and the efforts under way to expand its methodology to other domains within aviation.

Table of Contents

List of Tables xi

List of Figures xii

List of Illustrations xiii

Chapter One: Introduction 1

Chapter Two: Literature Review and Dissertation Research Objectives 5

- Human Factors in Aviation Safety 5
- Organizational Factors in Aviation Safety 19
- Airline Safety Data Sources 25
- LOSA: A Proactive Safety Measure 29
- Dissertation Research Objectives 30

Chapter Three: LOSA Methodology 31

- History of LOSA Development 31
- LOSA Data Quality Concerns 33
- LOSA and its Operating Characteristics 34
- LOSA Data Collection Overview 41
- LOSA Instrumentation 41
- LOSA Data Measures 43
- LOSA Observer Selection and Training 60
- LOSA Observation Procedures 67
- LOSA Data Reliability Checks 68
- Chapter Summary 74

Chapter Four: LOSA Data Analysis and Interpretation 76

- A Three-Stage Approach to LOSA Data Analysis 76
- Stage One – LOSA Indices and Organizational Profiles 77
- Stage Two – Drill-Down Analyses 83
- Stage Three – Targets for Enhancement 86
- A Demonstration of LOSA Data Analysis 86

LOSA Data Analysis in Research 95
Chapter Summary 100

Chapter Five: Discussion 101
Review: LOSA Rationale and Dissertation Research Tasks Completed 101
LOSA Implications for Aviation Safety 104
LOSA and its Future Research Directions 108
Conclusion 110

Appendix A: LOSA Observation Form 111

Appendix B: Threat Sub-Categories 118

Appendix C: Error Sub-Categories 121

Appendix D: Undesired Aircraft State Sub-Categories 130

Appendix E: LOSA Observer Feedback Form 131

Appendix F: Sample Error Organizational Profile (Zed Airlines) 133

Appendix F: Sample UAS Organizational Profile (Zed Airlines) 134

Bibliography 135

Vita 143

List of Tables

Table 2.1 – Human Error Perspectives in Aviation 13
Table 2.2 – Threat and Error Management Definitions 18
Table 3.1 – LOSA Threat Types and Examples 47
Table 3.2 – Primary Threat Management Variables 51
Table 3.3 – LOSA Flight Crew Error Types and Examples 53
Table 3.4 – Primary Error Management Variables 57
Table 3.5 – LOSA Undesired Aircraft State Types and Examples 58
Table 3.6 – Primary Undesired Aircraft State Management Variables 60
Table 3.7 – Sample TEM Training Exercise 65
Table 3.8 – Sample LOFF Results (Six Airlines) 71
Table 4.1 – LOSA Prevalence and Mismanagement Indices 79
Table 4.2 – Threat, Error, and Undesired Aircraft State Types 80
Table 4.3 – Sample Error Organizational Profile 82
Table 4.4 – Zed Airline's Threat Organizational Profile 88
Table 4.5 – Event Description Drill-Down Analysis: Aircraft Malfunction Threats 90
Table 4.6 – Sample Narrative Data 91
Table 4.7 – Sample Threat Description Data 92
Table 4.8 – Fleet Drill-Down Analysis: Aircraft Malfunction Threats 93
Table 4.9 – Geographic Base and Observation Count: Ten Airline Archive 96
Table 4.10 – LOSA Threat Indices: Ten Airline Archive 97
Table 4.11 – LOSA Error Indices: Ten Airline Archive 98
Table 4.12 – LOSA Undesired Aircraft State Indices: Ten Airline Archive 99
Table 5.1 – Dissertation Research Tasks Completed 104

List of Figures

Figure 2.1 – The IPO Model of Flight Crew Performance 10
Figure 2.2 – Threat and Error Management Framework 16
Figure 2.3 – The “Swiss Cheese” Model of Defenses 23

List of Illustrations

Illustration 3.1 – LOSA Data Collection Tool Opening Screen 42
Illustration 3.2 – LOSA Data Collection Tool: Predeparture Tab 45
Illustration 3.3 – LOSA Data Collection Tool: Threats Tab 50
Illustration 3.4 – LOSA Data Collection Tool: Errors Tab 56

Chapter One: Introduction

The next time you travel on a business trip or that much needed vacation getaway, the chances are you will use a commercial airline. Since the first scheduled passenger flight in 1914, the primary business objective for any airline is to fly as many passengers as possible with the least amount of operating cost.[1] However, there is an important condition. Airlines must offer their service in a safe and reliable manner. Any major safety event not only affects an airline's reputation but can change the industry as a whole.

Consider the unfortunate circumstances surrounding the World Trade Center terrorist attacks on September 11, 2001. Before this event, airlines were experiencing record growth and carrying more passengers than ever before (ICAO, 2004). However, after the attacks, the entire aviation industry went into a sharp decline (Boeing, 2004). The flying public had lost confidence in aviation safety and as a result, many people stayed away from airline services. While the events on September 11th were unique, they serve as an important reminder that any high-profile airline accident has the potential to shake societal views of aviation safety. It is this concern that drives airlines to manage with sound business practices while knowing they truly succeed or fail on the strength of their safety performance.

Fortunately for the flying public, airlines have historically had a stellar safety record. Between 1994 and 2003, the accident rate in commercial aviation was roughly

---

[1] On January 1, 1914, the St. Petersburg-Tampa Air Line operated the first scheduled passenger flight from St. Petersburg to Tampa, Florida. The airline up to that time was primarily a mail carrier. Tom Benoist, a flying boat builder and part owner of the airline, customized one of his bi-wing sea planes to carry pilot Tony Jannus and passenger Abram Pheil on the 23 minute flight across the bay to Tampa. After the inaugural flight, the popularity of the single-seat passenger flight increased primarily among the many Northern tourists visiting Florida during a winter break. However, the service only lasted for a few months as the Northern tourists started to return home at the end of the winter season.

one in one million flights worldwide (Boeing, 2004).[2] This degree of safety is remarkable, especially when compared with other forms of transport. For instance, archival findings show that people are two times more likely to suffer a fatal injury in a passenger train than when flying on a commercial airplane and 30 times more likely to die in an automobile (NSC, 2003). While statistics such as these testify to the success airlines have in managing risk, there are several environmental changes on the horizon that have the potential to erode current safety levels. Possibly, the most serious of these changes are projected increases in air traffic.

The commercial aviation industry is projected to grow at a rate of 5% more passengers every year for the next 10 years (Boeing, 2003). To satisfy this demand, the number of airline departures is expected to increase by 30% to about 26 million flights a year (ICAO, 2001). Assuming the airline accident rate remains stable as it has for the past 10 years, Wiegmann and Shappell (2003) suggest such air traffic increases could lead to an increase in the frequency of airline accidents. If this is the case, commercial aviation could have one fatal accident for every two weeks of the year by 2010.[3] In an industry that survives by providing safe transport, an accident frequency such as this could cause an irretrievable loss of public confidence.

While increases in air traffic might yield more accidents a year, there are other factors that could also compromise aviation safety. Perhaps the most serious are capacity concerns at major airports and air traffic services. The Federal Aviation Administration (FAA) recently conducted a congestion study at 35 major airports in the United States (2004). Their findings showed many potential faults. About 43% of the airports studied

---

[2] This accident rate is based upon 143 million total departures from 1994 to 2003. Events used to calculate this rate only include accidents involving a hull loss and/or fatalities. A hull loss is defined as structural damage to an airplane beyond economic repair (Boeing, 2004).

[3] Wiegmann and Shappell originally projected one accident per week (2003). However, their calculations were based on 1997 forecasts before the September 11, 2001 terrorists' attacks. The "one accident in two weeks" projection uses more recent long-term forecasts showing projected growth (ICAO, 2001).

will not be able to handle projected increases in air traffic over the next 10 years. Some of the nation's busiest airports, such as Atlanta Hartsfield, Newark Liberty, New York LaGuardia, Chicago O'Hare, and Philadelphia International, were identified as already exceeding their model capacity limits. In response to these findings, the FAA has begun collaborating with airlines to achieve more efficient flight schedules, cutting traffic at peak times, and working with airports in building extra runways.

Apart from congestion issues at the airports, there is also an issue of overcrowding in the air and its effect on Air Traffic Control (ATC). ATC services are the traffic officers of the skies. They provide pilots with directions for takeoffs and landings, keep airplanes safely separated, and oversee traffic flows in and out of various areas of airspace. These responsibilities are not limited to commercial airliners, but also extend to airplanes from the military, cargo operators, and general aviation. Consider the FAA Air Route Traffic Control Centers (ARTCC), which are responsible for managing airplanes between their arrival and destination air space. In the year 2000, ARTCC controllers handled over 46 million flights (FAA, 2001). With the projected increases in air traffic, the workload for these controllers is expected to increase by 45% to over 67 million flights by 2015 (FAA, 2001). As with airport capacity, ATC overcapacity is again a safety concern. Due to cuts in government funding, outdated technology, and a glut of retirees leading to a shortage of trained controllers, ATC is already being overextended and is hard-pressed to handle more airplanes (Fischer, 2001; GAO, 2002).

As the flying environment becomes more congested, the question becomes, "How can airlines best understand and manage safety in such times of change?" In response to this question, many industry professionals believe that airlines must improve their search for system safety vulnerabilities in normal flight operations (Helmreich, Wilhelm, Klinect, & Merritt, 2001; Maurino, 2001; Reason, 1997). The foundation of such a search is the collection and processing of safety performance data. In commercial aviation, safety performance data can come from various sources. Some of the most common are quick access recorders, training evaluations, incident reporting systems, or in rare cases, accident investigations. The problem with such sources is they either point to organizational failures when it is too late after an incident or accident has taken place, or they provide fragmented perspectives of safety performance. For example, training evaluations are designed to collect information on pilot proficiency issues so training curriculum can be improved. On the other hand, quick access recorders are limited to the capture of flight parameters, such as an airplane's heading, speed, or altitude with no insight of pilot proficiency or flight crew behaviors inside the cockpit. This dissertation proposes a solution to these safety data limitations. It is a cockpit observation method called the Line Operations Safety Audit (LOSA) that gathers system safety and flight crew performance data during regularly scheduled flights in normal operations.

Chapter Two: Literature Review and Dissertation Research Objectives

This chapter provides a presentation of previous work inspiring LOSA development and its rationale. It begins with a review of human factors research in the areas of flight crew teamwork and human error. From the micro-perspective of human factors, the discussion shifts to the macro-perspective of organizational factors in aviation safety. The research from high-risk industries provides an explanatory model of accidents and the degree of influence airlines have in their prevention. This is followed by an evaluation of major airline safety data sources and how LOSA complements these sources in providing a unique perspective of flight crew performance. Finally, the last section lists the research objectives of this dissertation.

HUMAN FACTORS IN AVIATION SAFETY

> *Human factors is the discipline that attempts to optimize the relationship between technology and the human* (adapted from Kantowitz & Sorkin, 1983).

The Early Days of Aviation Human Factors

Since Orville and Wilbur Wright conducted the first powered flight in 1903, aviation pioneers have not only spent time improving aircraft construction but have also devoted attention to the influences of the pilots at the controls. This early focus marks the beginning of aviation human factors (Roscoe, 1980). Much of the originating work in human factors focused on pilot psychomotor skills, perceptual needs, or bodily stresses involved in keeping an airplane in stable flight (Costa, 1999; Koonce, 1984). As a result, many new airplane instruments were created giving pilots immediate data on several key flight parameters, such as altitude and attitude (pitch). However, regardless of the advances in instrumentation, many of aviation's pioneers held fast to their belief that successful flying was mostly dependent on a person's piloting skills.

The focus on pilot skills continued through World Wars I and II. Pilot selection and training received the most attention as air warfare became the military tactic of choice (Flanagan, 1947). More traditional human factors experiments, such as pilot-machine interface studies, were also carried out to support the war efforts (Champanis, Gardner, & Morgan, 1947; Fitts & Jones, 1947). One such study presented by Roscoe (1997) was conducted in 1943 to research causes behind a rash of accidents involving pilots landing without lowering their wheels. Alphonse Champanis, a human factors psychologist at the time, was called on to study the problem. After examining the cockpits of airplanes involved in the incidents, he suggested the design of the wheels and flap levers was the culprit. He noted both levers were located in the same place and nearly identical with each other. As the workload increased in preparation for landing, Champanis surmised pilots would get the levers confused and select the wrong one, extending or retracting flaps instead of lowering the wheels. His solution was to attach a circular piece of rubber on top of the wheels lever to stand for wheels, and a rubber triangular wedge on top of the flap lever to stand for flaps. The thought was pilots would not only be able to see the difference between levers but they would also be able to feel the difference. The solution worked, resulting in an immediate drop in wheels-up landings. This now classic study has become an often cited example in flight control design (Roscoe, 1997).

After World War II, much of human factors research continued to concentrate on pilot selection, training, and machine interface issues (Koonce, 1999). However, another avenue of research started to grow, prompted by airplane manufacturers reducing the number of pilots in the cockpit. Most of the early commercial airplanes needed a three-person crew to fly: a pilot, copilot, and flight engineer. The flight engineer was mostly responsible for monitoring systems and had a back-up role in supporting the two pilots at the controls. However, advancements in automation later removed the flight engineer position and reduced all flight responsibilities to two pilots, a Captain (pilot) and First Officer (copilot). This decrease in pilots made airlines more profitable by cutting their labor costs per flight. At the same time, it made flight crew teamwork a more valued commodity in preserving safety.

Enhancing Aviation Safety with Teamwork: Crew Resource Management (CRM)

In June 1979, the National Aeronautical and Space Administration (NASA) hosted a now historic workshop to address crew performance issues in aviation (Cooper, White, & Lauber, 1980). According to Lauber (1993), the driving force for the meeting was human factors research identifying crew coordination failures contributing to pilot errors. As researchers were studying the more traditional human factors issues of pilot-system interface, shortfalls in less traditional crew performance areas, such as leadership, communication, decision-making, and monitoring/cross-checking, became increasingly obvious (Ruffell Smith, 1979). Analysis of accident data uncovered a similar trend (Cooper et al., 1980). Most of the findings suggested that accidents were more likely to reflect failures in team performance than shortages in “stick and rudder” skills (Weiner, Kanki, & Helmreich, 1993). With these findings in mind, workshop attendees decided that flight crew coordination needed to be taught much like training pilots to fly. The new idea became known as Crew Resource Management (CRM) and is now a training standard for commercial airlines worldwide (Weiner et al., 1993).[4]

CRM training has evolved over five generations (Helmreich, Merritt, & Wilhelm, 1999). The first generation of CRM mostly focused on addressing pilot personality to bring about effective crew coordination. Much like the early days of human factors research in the military, psychological testing was prominent in selecting pilots who would work well with others. Early detractors of CRM labeled these efforts as "charm school" that tried to soften the persona of pilots from being true pilots. In response, the second generation of CRM training moved away from personality to training pilots by providing examples of wanted flight crew behaviors and linking them to flying tasks. Acceptance for second generation CRM grew and expanded to the third generation, which underlined the context in which flight crews performed. This generation considered the systemic influences, such as organizational culture, regulations, and pressures that were outside the control of flight crews, yet had an impact on their performance. The fourth generation of CRM was known for its focus on everyday flight operations. Up to this point, CRM training was a stand-alone course that pilots attended once a year. Fourth generation CRM aimed to proceduralize crew coordination concepts and seamlessly integrate them with technical proficiency training.

In the mid-1990's, researchers noted CRM losing acceptance among pilots (Helmreich & Taggart, 1995). Some questioned whether CRM had any influence at all on crew performance and safety (Maurino, 1999; Salas, Burke, Bowers, & Wilson, 2001). This prompted a re-examination of the original CRM principles set forth at the previously discussed 1979 NASA workshop. In its fifth generation, the driving force behind CRM was to reduce pilot error and prevent accidents, not training teamwork for the sake of

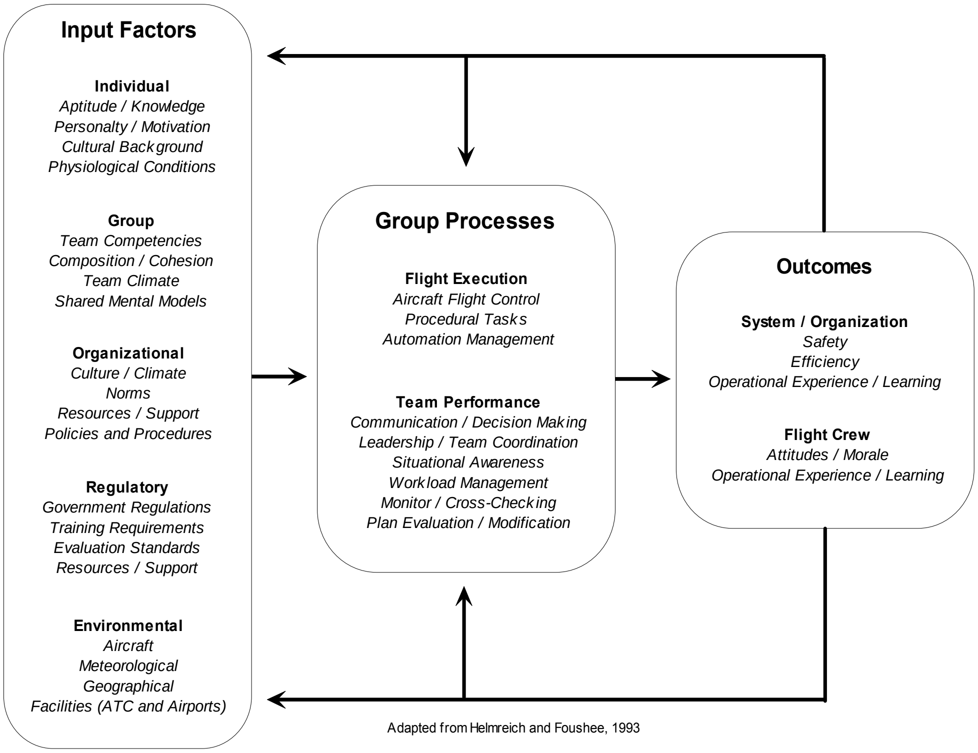
---

[4] The original name of CRM training was Cockpit Resource Management. However, it was later changed to Crew Resource Management to emphasize the crew component of flight safety performance.

teamwork. However, the realization that pilot error was unavoidable and a natural condition of human nature led some to reject the idea of an error-free flight (Helmreich & Merritt, 1998; Maurino, Reason, Johnston, & Lee, 1995). With this realization, CRM was given a renewed sense of purpose by helping flight crews to not only prevent errors when possible but also manage them through effective teamwork (Helmreich et al., 2001). This latest generation of CRM has stimulated a convergence of teamwork and human error research to better understand error management processes and their influences. However, before discussing this convergence, a review of the relevant teamwork and human error research literature is needed.

Teamwork Research: Understanding Input Factors, Group Processes, and Outcomes

To better understand teamwork in aviation, a model is helpful to show its many parts and relationships. An often cited model of flight crew teamwork is the Input-Process-Output (IPO) Model of Flight Crew Performance (Helmreich & Foushee, 1993; McGrath, 1964). The model shows flight crew performance being defined by: *input factors* influencing *group processes* that lead to flight *outcomes*. A current version of the model is shown in Figure 2.1.



Adapted from Helmreich and Foushee, 1993

Figure 2.1 – The IPO Model of Flight Crew Performance

As shown in Figure 2.1, *input factors* can originate from several different sources. Some of the factors are what pilots bring with them when they enter a cockpit, such as personality, while others are external and beyond their control, such as regulatory influences or organizational norms. A unique characteristic of all input factors is they can either act alone or interact with others to influence group processes. This multidimensional feature of the IPO model is essential since it reflects the contextual realism that flight crews often have to manage along with their regular flying duties.

The next unit of the IPO model is *group processes*, both flight execution and team performance. Flight execution includes the "stick and rudder" handling, procedural, and automation tasks needed to fly the airplane. The team performance processes are similar to the issues that sparked the beginnings of CRM. They consist of but are not limited to communication, decision making, leadership, situational awareness, leadership, workload management, monitor/cross-checking, and plan evaluation/modification.

The final unit of the IPO model is *outcomes.* The most obvious outcomes are flight safety and efficiency. However, there are other types of outcomes influenced by input factors and group processes that are less direct. These include airline operating experience, system learning, pilot attitudes, and morale. Most importantly, outcomes can have a recursive effect on future input factors and group processes.

The IPO unit receiving the most attention in teamwork research is group processes (Burke, Wilson, & Salas, 2003). This is not surprising since group processes mediate the relationship between input factors and outcomes. A few examples of research into the impact of input factors on group processes include:

- Examining the effect of aircraft malfunctions on flight crew workload management (Ruffell Smith, 1979).
- Investigating crew composition on the quality of crew communication (Foushee, Lauber, Baetge, & Acomb, 1986).
- Exploring the effects of national culture on leadership (Merritt, 2000).
- Assessing automation design on crew coordination (Weiner, 1993).
- Studying team competencies, defined as the collective knowledge, skills, attitudes and mental models of crews and their influence on various group processes, such as crew coordination and decision making (Cannon-Bowers, Tannebaum, Salas, & Volpe, 1995; Stout, Salas, & Carson, 1994).

A drawback in much of the general teamwork research, including research in aviation human factors, is limited consideration of the operational context in which team performance naturally occurs (Guzzo & Dickson, 1996). Organizational, regulatory, and environmental input factors are only a few of the contextual determinations that influence flight crew performance. This noted shortfall points to a need for more field research that takes into account the context and its effect on group processes. It is this notion that partly inspired LOSA development.

Human Error Research: Error Causation and Management

Flight crew error has been identified as a primary cause in over 60% of the commercial airline accidents worldwide (Boeing, 2004). Regardless of advances in aircraft technology, this statistic has remained relatively stable since the introduction of commercial jets in the 1950's (Boeing, 2004). Nevertheless, these findings continue to motivate research to better understand and lessen the influences of pilot error in airline safety (Helmreich, 1997; Nagel, 1988; Shappell & Wiegmann, 1997; Woods, Johannesen, Cook, & Sarter, 1994).

Much of the human error research in aviation can be separated into two major topic areas: (1) error causation, and (2) error management. First, the majority of error causation research seeks to understand the nature of human error. This objective seems simple to study, but it has historically been one of the more heated sources of debate in the literature (Senders & Moray, 1991). As in many other fields of research, definitions and assumptions often depend on the academic discipline of the researcher. Given this potential bias, Wiegmann and Shappell (2002) summarized the error causation research in aviation as stemming from five perspectives: (1) cognitive, (2) ergonomics and systems design, (3) aeromedical, (4) psychosocial, and (5) organizational. Table 2.1 provides a summary of each perspective.

Table 2.1 – Human Error Perspectives in Aviation

Perspective	Theory of Error Causation	Sample References
Cognitive	Failure in pilot "information processing" of stimuli	(Hollnagel, 1998) (Reason, 1990)
Ergonomics and Systems Design	Mismatch between pilots, airplane design, operational rules and environmental conditions	(Edwards, 1988) (Heinrich, Petersen, & Roos, 1980)
Aeromedical	Physical condition or ailments such as pilot fatigue, spatial disorientation or hypoxia	(Costa, 1999) (Previc & Ercoline, 2004)
Psychosocial	Failures in flight crew coordination and interpersonal communication	(Helmreich & Foushee, 1993) (Prince & Salas, 1999)
Organizational	Flawed organizational policies, procedures, or managerial decisions creating error-inducing conditions for pilot	(Perrow, 1984) (Reason, 1997)

While the five perspectives of human error causation research in Table 2.1 might seem independent of one another, they are not. For example, several cognitive psychologists have expanded their research to include system factors normally studied by organizational psychologists (see Rasmussen, Duncan, & Leplat, 1987; Reason, 1997; Woods et al., 1994). While this inclusion provides a deeper understanding of the interaction between cognitive ability and context, much of the effort is theoretical, often lacking empirical evidence from the environment in which the findings apply.

A second, less-examined area in human error research is called error management. Instead of identifying the antecedents of error, error management researchers study the environmental influences and psychological mechanisms involved after committing an error. There are three stages that define the error management process: error occurrence; error diagnosis; and error recovery (Bagnara & Rizzo, 1989;

Zapf & Reason, 1994).[5] The first stage, error occurrence, begins with an individual or team committing an error. Next, the error diagnosis stage involves detecting, describing, and making sense of the error. Finally, the last stage, error recovery, is defined as the act of planning and performing corrective actions to keep the error from causing further complications. While all three stages of error management are critical to safety performance, the most important is arguably error detection since an error undetected cannot be managed. This concept alone has stimulated much of the research in error management.

As in error causation research, cognitive researchers are leading the way with their study of error detection and its impact on human-machine interface design (Rizzo, Bagnara, & Visciola, 1987; Sellen, 1994; Zapf, Maier, Rappensperger, & Irmer, 1994). However, as in error causation research, much of this work has either been conducted as simulator experiments lacking operational context, or is based on data gathered from aviation mishaps, such as incident reports and accident investigations. While such research has produced valuable insight, a "restriction of range" problem exists since findings are arguably skewed to sterile research environments or negative flight outcomes, such as those deriving from incident and accident data. It leads to questioning whether these same findings would generalize to the same psychological mechanisms involved in successful error management performance.

A possible solution to the "restriction of range" problem is to conduct field studies that capture flight crew performance in its natural context. This would provide a more balanced sample of successful and unsuccessful error management cases, increasing the generalizability of findings. However, before collecting field data, a theoretical foundation is needed that couches error management performance in relation to its

---

[5] Zapf and Reason (1994) actually defined this process as *error handling*. The term *error management* was used as a replacement because it is a more familiar term within the commercial airline industry. There is no difference in definition.

operating environment. This dissertation proposes the Threat and Error Management (TEM) framework developed by The University of Texas Human Factors Research Project as one such foundation for data collection (Helmreich et al., 2001).

Threat and Error Management (TEM)

The TEM framework in Figure 2.2 is a conceptual combination of the previously discussed IPO model of flight crew performance and stages of error management. It proposes that threats and errors are part of everyday operations that flight crews have to manage to maintain flight safety. The framework itself was empirically derived from cockpit observations pre-dating LOSA that mostly focused on CRM performance. This initial derivation of TEM continues to be refined as a collaborative effort between The University of Texas researchers, airline managers, pilots and aviation subject matter experts. This collaboration ensures TEM concepts remain relevant in providing a common frame of reference and terminology for all stakeholders interested in understanding flight crew performance during normal operations. It also serves as the primary measure for LOSA.

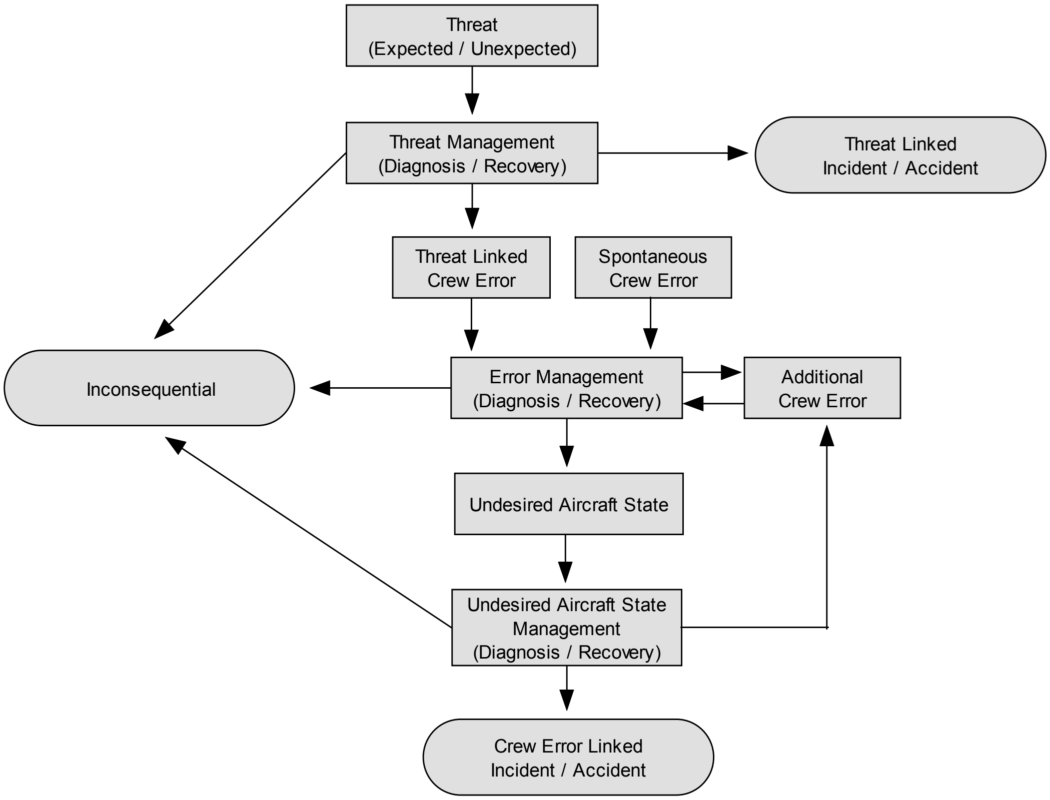


Figure 2.2 – Threat and Error Management Framework

Threats and Threat Management

On a typical day of flying, pilots have to manage various environmental complexities. Such complexities include dealing with adverse meteorological conditions, airports surrounded by high mountains, congested airspace, and errors committed by other people outside the cockpit, such as air traffic controllers, flight attendants or maintenance workers. In TEM, these events are known as threats because they all have the potential to compromise flight safety.

Much like error management, the effectiveness of a flight crew's ability to manage threats depends on their detection. However, threats, unlike errors, can be anticipated. For example, pilots can anticipate the effects of a thunderstorm by briefing their response in advance, or prepare for an often-congested airport by making sure they keep a watchful eye for other aircraft as they carry out their approach. Threats can also occur unexpectedly, such as a dispatch paperwork error that happens without warning. Regardless of whether threats are expected or unexpected, flight crews are considered the last line of defense in keeping them from affecting safety.

Errors and Error Management

From the TEM perspective, error is a crew action or inaction that leads to a deviation from crew or organizational intentions or expectations. Errors can be spontaneous, linked to threats, or part of an error chain. Examples include aircraft handling errors, such as beginning to fly to a wrong heading; automation errors, such as executing a wrong mode; procedural errors, such as failing to give a proper callout; and communication errors, such as misinterpreting an ATC runway landing clearance. Similar to threats, and regardless of the type committed, an error's effect on safety depends on flight crew detection and management.[6]

Unfortunately, not all errors are well managed. Sometimes they lead to another error or a safety-compromising event called an undesired aircraft state (UAS). Undesired aircraft states are flight crew error-induced aircraft deviations or incorrect configurations associated with a clear decline in safety margins. Often considered on the cusp of becoming an incident or accident, they also involve flight crew detection and management. Some examples of undesired aircraft states include lining up for the wrong runway during landing, exceeding an ATC speed limit during an approach, or landing

---

[6] One could argue that the original concept of a two-person cockpit was to have one pilot flying and the other pilot monitoring for threats and errors. Hence, TEM is very much a team activity in aviation.

long on a short runway. Some readers might wrongly identify events, such as equipment malfunctions or an ATC controller error clearing the flight crew to wrong altitude, as undesired aircraft states. In the TEM framework, these events are threats. Undesired aircraft states are only caused by undetected or mismanaged flight crew errors. More details on TEM definitions and examples are provided in Table 2.2.

Table 2.2 – Threat and Error Management Definitions

Element	Definition	Examples
Threat	Event, error, or aircraft state that occurs outside the control of the flight crew but still requires their management to maintain safety.	Thunderstorms Late ATC clearance Ground/ramp error Aircraft malfunction
Error	Crew action or inaction that leads to a deviation from crew or organizational intentions or expectations.	Missed checklist item Omitted approach callout Incorrect flap setting Wrong automation entry
Undesired Aircraft State	A crew–error–induced aircraft state that clearly reduces existing safety margins.	Altitude deviation Unstable approach Runway incursion Long landing

Practical Applications of Threat and Error Management

TEM was originally conceived as a data collection measure for LOSA. Now, it has become widely recognized and accepted within the airline industry. Major U.S. airlines such as Continental, Delta, and US Airways, and foreign carriers such as Cathay Pacific, Singapore Airlines, and Air New Zealand, currently use TEM as the foundation for their human factors and CRM training. Regulatory agencies have also started to realize the utility of TEM. The International Civil Aviation Organization (ICAO) has begun an effort to incorporate TEM in their operating standards and pilot licensing requirements. ICAO has also produced a human factors training manual for its 186 member states largely based on TEM performance (ICAO, 2002).

Human Factors Research Summary

This section reviewed the human factors research literature with a particular focus on teamwork and human error. As discussed, much of this research has either been theoretical with little empirical evidence, experimental simulator studies lacking true environmental context, or aviation mishap studies that are naturally predicated by failures of some sort. This review of teamwork and human error research is not intended to denigrate previous work. Instead, it underscores a need to better understand the complex relationships among operational context, flight crew performance, and safety outcomes. This dissertation proposes the TEM framework as one such measure that enables a systematic study of these relationships.

ORGANIZATIONAL FACTORS IN AVIATION SAFETY

While regulators, manufacturers and other commercial aviation industry groups have notable influence on safety, this section limits its focus to airlines. It is ultimately the airline's responsibility to provide mechanically sound airplanes, consistent ground support, effective risk management practices, and trained flight crews to fly passengers to their destinations. Deficiencies in any one of these areas can at any time, compromise airline safety. The question becomes how much control airlines have in controlling these deficiencies and preventing accidents.

Normal Accident Theory (NAT)

Are accidents preventable in complex, high-risk industries such as commercial aviation? If you subscribe to the Normal Accident Theory (NAT), you might decide the answer is no. The basic principles of NAT came from examining the 1979 Three Mile Island nuclear reactor accident that occurred outside Harrisburg, Pennsylvania (Perrow, 1981). The accident involved multiple system events that nearly induced a reactor core meltdown and the massive release of radiation. Fortunately, the radiation was contained, but the reactor was destroyed, with a clean-up cost of over one billion dollars (Rees, 1994). The accident was one of the most comprehensively studied events in U.S. nuclear history. The many investigations that followed not only focused on the malfunctions and operator error that triggered the event, but also on organizational factors, including managerial failures of communication and complacency (Hale, Baram, & Hovden, 1998; Hopkins, 2001). Perrow argued the Three Mile Island accident was inescapable due to several component failures interacting in ways previously unanticipated by system designers, safety managers, and operators. Thus, these failures rapidly intensified to a catastrophe before operators could understand what was happening. Perrow (1984) came to call these types of events "normal accidents" to represent their inevitability in many of the world's high-risk industries.

NAT contends accident prevention becomes more difficult as a system becomes more complex and tightly coupled (Perrow, 1984; Sagan, 1993). Unfortunately, this is the predicament of commercial aviation. Flying a commercial airliner is complex due to the multiplicity of input factors, defined by the IPO model discussed earlier, that can interact in unpredictable ways. Compounding this problem in aviation is tight-coupling: regimented tasks often needing quick execution under time pressure and with little room for error. The high complexity, tight-coupling scenario proposed by NAT does not paint a rosy picture for airlines. However, some industrial theorists believe that many airlines have unique organizational characteristics that strengthen their ability to prevent accidents. These characteristics have come to form the foundation of what researchers call high-reliability organizations.

High Reliability Theory

High Reliability Theory seeks to understand how organizations are able to keep such high safety standards in industries full of risk (La Porte & Consolini, 1991; Roberts, 1993). Rochlin (1993) is quick to note that an exemplary safety record is not enough to discriminate high reliability organizations (HRO) from others. It is their effective corporate design, management practices, and operational policies that reduce their exposure to accidents. HROs also promote a belief that safe performance is not good enough; it must be something that is continuously looked after and improved (Rochlin, 1993; Weick, 1987). Some researchers suggest that high reliability comes from organizations being *mindful* of managing the unexpected (Weick & Sutcliffe, 2001). Weick and Sutcliffe define organizational *mindfulness* with the following characteristics:

- Preoccupied with safety performance failures
- Reluctant to simplify and considers as many factors as possible
- Sensitive to operational performance on the front lines
- Committed to continuously build resiliency against safety hazards
- Willing to seek external perspective of performance issues

While these characteristics offer a description of effective HRO operating practices, they fail to provide information on what organizations specifically target when managing the unexpected. For airlines, these targets are provided by perhaps the most widely accepted model of accident causation in commercial aviation, the “Swiss Cheese” Model of Defenses (Reason, 1997).

The “Swiss Cheese” Model of Defenses

The “Swiss Cheese” Model of Defenses in Figure 2.3 proposes that all organizational accidents are the result of a breakdown in systemic defenses designed to prevent hazards from compromising safety margins (Reason, 1997). The hazards in commercial aviation are multi-faceted. They involve the risks associated with flying complex airplanes, on-time, with passengers, while managing a threat-laden environment. A popular adage in the aviation industry is if an airline wants to avoid safety hazards, they should keep their airplanes on the ground. Of course, this is not possible, so airlines, regulators, manufacturers and other aviation stakeholders must build system safety defenses to keep these hazards in check.

System safety defenses are designated as “hard” and “soft” safeguards (Reason, 1997). Hard systemic defenses are those mostly associated with aircraft design, such as automation, instrument displays, and aircraft warnings. A good example of a hard systemic defense is the Traffic Collision Avoidance System (TCAS). TCAS is a piece of equipment on many airplanes providing flight crews with visual and audio warnings of nearby airplanes to prevent midair collisions. Soft defenses are more common in aviation. These can include regulations, standard operating procedures, training, aircraft maintenance, air traffic control principles, and most important, the pilots, who are often the last line of defense against hazards. The slices of cheese in Figure 2.3 represent the many layers of systemic defenses (hard and soft) in commercial aviation.

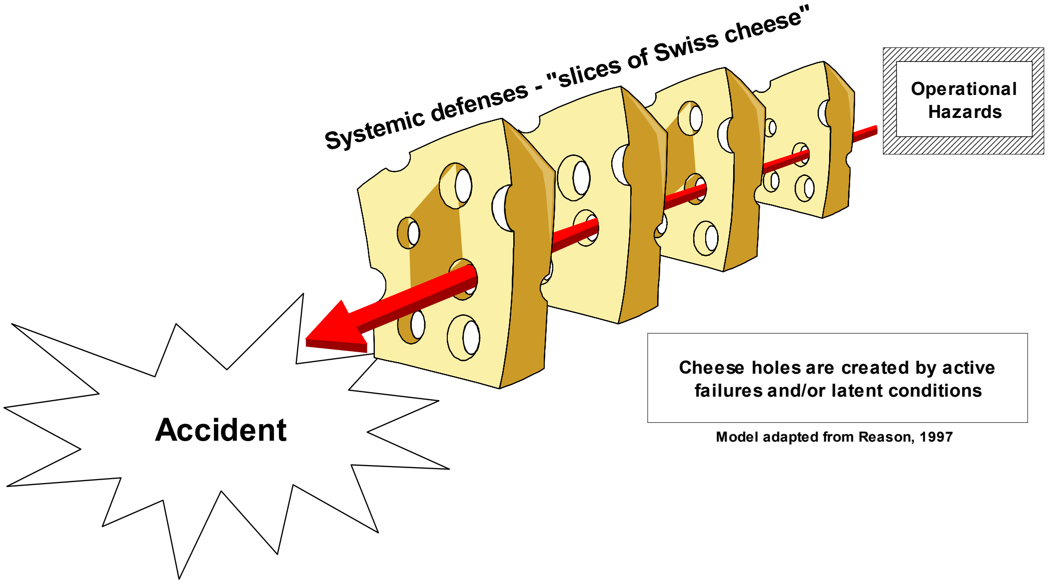


Figure 2.3 – The "Swiss Cheese" Model of Defenses

Unfortunately, not all systemic defenses are perfect in aviation; they can have weaknesses symbolized by the Swiss cheese holes of various sizes, large and small. The larger the hole, the more likely an operational hazard will penetrate through a defense. If too many holes line up, operational hazards can pass through all defenses and lead to an accident (shown by the arrow in Figure 2.3). This is perhaps best illustrated with an example.

Imagine a flight with an unexpected engine failure during the cruise phase of flight. It just so happens that the airline's pilot training curriculum for this two-engine aircraft includes this scenario in their flight simulator sessions. Therefore, the crew is well-prepared and manages the problem by following emergency procedures in a checklist called the Quick Reference Handbook (QRH) for an engine shutdown (two "soft" safeguards). While the aircraft is designed to fly with one engine ("hard" safeguard), the crew decides they need to divert to the closest airport by first gaining an

ATC clearance to do so (pilot decision-making, the final defense). Each of these example safeguards provides an added layer of protection against a known hazard.

Now, let's change the scenario. Due to budget constraints, the airline decides to reduce their annual pilot training and remove the engine shutdown simulation. Also, imagine an over-confident or forgetful flight crew that diagnoses an engine problem without referring to the QRH. This omission leads them to overlook an important step and shuts down the wrong engine leading to an accident. Each of the events in the changed scenario represents "holes" in airline training and pilot performance defenses that compromised safety. Therefore, under the HRO perspective, finding and plugging these holes is the basis of accident prevention. However, for this to happen, an understanding of what causes the holes is needed.

According to Reason (1997), holes in systemic defenses are caused by active failures and latent conditions. Active failures are unsafe acts committed by people on the front lines. In aviation, these can include errors or violations by pilots, air traffic controllers, ground personnel, or maintenance. When committed, they have an immediate impact on safety, often bringing the best designed systemic defenses into jeopardy. Unlike active failures caused by individuals at the operational levels, other defense holes, known as latent conditions, originate from higher levels of the system, such as regulators, manufacturers, and airline management. Latent conditions typically lie dormant for long periods of time, incubating until they interact with active failures, such as pilot error, to contribute to an accident (Reason, 1997; Turner, 1978). For commercial aviation, latent conditions might include poor organizational climate, flawed management decisions, inefficient aircraft design, sloppy maintenance, inadequate pilot training, ambiguous procedures or automation anomalies. Unfortunately, many latent conditions result from a *failure of foresight* and are typically identified only in accident investigations after it is too late (Pidgeon, 1997; Turner & Pidgeon, 1997).

To compound the failure of foresight, latent conditions are much more sweeping within an organization. Active failures, such as pilot error, are short lived and typically affect safety one flight at a time. On the other hand, latent conditions are more across–the–board, possibly compromising safety on all flights within an airline. From a risk management perspective, this would arguably make the search for latent conditions a top priority. So, can airlines identify and correct latent conditions *before* they contribute to an accident? The answer resides in what accident causation theory is applied; Normal Accident Theory would suggest no, High Reliability Theory and the Swiss Cheese Model of Defenses would suggest yes. Regardless of which theory is adopted, airline accident prevention from a system-centered approach requires performance data.

AIRLINE SAFETY DATA SOURCES

If commercial airlines aspire to become mindful in their search for latent conditions, they need to understand what causes their airplanes to move from the center to the edges of the "aviation safety envelope." This is accomplished by collecting and processing performance data from accidents, incidents and normal flight operations.

Accidents and Accident Data. Airline accidents are catastrophic events involving loss of life, serious injuries, or irreparable structural damage to an aircraft (NTSB, 2002a). As mentioned before, accidents are rare events in commercial aviation. However, when an accident does occur, there is often much media publicity and interest by various aviation stakeholders, such as government regulators, manufacturers, airlines, and the flying public. With such public attention, there is a lot pressure to understand causes of an accident with a formal investigation.

Accident investigations are usually conducted by governmental safety agencies. In the United States, this agency is the National Transportation Safety Board (NTSB). When an accident occurs, NTSB investigators are dispatched to the wreckage site to collect information on anything or anyone remotely involved. They can sometimes spend years and large amounts of resources conducting just one investigation.[7] The purpose of this effort is not necessarily to assign blame but to identify causes and provide recommendations for prevention (Wells, 2001). Most times, recommendations focus on removing hazards contributing to the accident by incorporating new safety design features, warning devices or procedures (Diehl, 1991). Accident investigations also provide airlines with massive amounts of data on their operations that possibly highlight other areas of safety weaknesses needing correction.

Incidents and Incident Data. Estimated to occur roughly 60 times more often than accidents, incidents represent the next most serious safety event in commercial aviation (FAA, 2001).[8] Examples of incidents could include a near-midair collision of two airplanes approaching each other on the same altitude or a runway incursion, which could occur when a taxiing aircraft enters a runway being used by landing aircraft. The primary motivation behind collecting incident data is incidents occur more often than accidents and are essentially "free lessons" for those airlines wanting to expand their

---

[7] Examples of some of the more extensive accident investigations in U.S. aviation history include the 1996 Trans World Airlines (TWA) Flight 800 in-flight breakup off the coast of Long Island, New York, caused by ignited fuel in the center wing fuel tank (NTSB, 2000); and the 2000 Alaska Airlines Flight 262 accident over the Pacific Ocean caused by an in-flight failure of the horizontal stabilizer used to control the aircraft's pitch (NTSB, 2002b).

[8] The information used in this estimate was based on the number of incidents reported to the Federal Aviation Administration (FAA) by major and commuter airlines between 1985 and 1998. The estimation is based upon the average annual number of incidents per one million departures compared to the accident rate of one in one million departures derived from Boeing (2004).

search for safety defense weaknesses (March, Sproull, & Tamuz, 1991). However, unlike accidents that are highly visible when they occur, incidents are usually more obscure, relying largely on pilot self-report to mandatory or voluntary reporting systems (Helmreich et al., 2001; Wells, 2001). Mandatory incident reporting requires pilots by law to send a descriptive account of certain types of incidents to a country's regulator. Unfortunately, the quality of reports in such systems is questionable due to pilot fear of disciplinary action. As a result, mandatory incident reporting typically provides data on aircraft malfunctions with little or no data on human factors performance issues (Wells, 2001).

Voluntary incident reporting systems were introduced to capture safety performance data outside mandatory incidents. By tapping into these otherwise "non-reportable" incidents, a wealth of new safety information is made available (Tamuz, 2001). In the United States, the Federal Aviation Administration (FAA) supports several voluntary incident reporting systems. Possibly the most well known is the Aviation Safety Reporting System (ASRS) managed by the National Aeronautical Space Agency (NASA). However, voluntary incident reporting managed internally by airlines is steadily gaining in popularity because data are kept in-house. One such system is a joint effort between airlines, pilot associations, and the FAA known as the Aviation Safety Action Program (ASAP).

In exchange for confidentiality and protection against disciplinary action, ASAP asks pilots to report events, close-calls or hazards that they see to be unsafe.[9] If pilots trust the airline's assurances to remain non-punitive, they report more often and provide more details on human factor issues that contributed to an event compared to mandatory reporting, where such issues are usually not discussed (Harper & Helmreich, 2003;

---

[9] The protection against disciplinary action does not include events caused by deliberate or criminal acts of negligence.

Wells, 2001). This effect was found at Continental Airlines, where 88% of their filed ASAP reports described events that mandatory reporting would not have captured (Gunther, 2003).

Normal Operations and Normal Operations Data. All flights not resulting in an incident or accident are considered normal operations. Data from these flights can provide airlines with valuable information on systemic defense strengths and weaknesses before a serious safety event. The primary data sources for such information are quick access recorders and line checks. Let's first examine the quick access recorder.

Quick access recorders (QAR) are digital devices that gather data on various flight parameters, such as altitude, heading, speed, flight control positions, and hundreds of other measurements from sensors located throughout an airplane. In addition, QARs can record data on these parameters at a sampling rate of every second an airplane is in operation. Not surprisingly, this can create huge amounts of data for an airline to absorb and analyze. To control potential data overload, airlines will sometimes only examine those flights with predetermined exceedances in engine settings, airspeed, altitude or heading (Chidester, 2003). Once flights with exceedances are aggregated, an airline can highlight and target safety weaknesses with corrective actions. However, QAR as a safety source is limited. There is no information about environmental context or flight crew performance issues that would allow one to examine the reasons underlying exceedances. Put simply, QAR only provides information on *what* happened, not *why*.

Another safety data source from normal operations is line checks. Often required by governmental regulators, line checks are cockpit evaluations of pilots during regularly scheduled flights to assess proficiency. Line checks are typically graded as pass or fail, and pilots that fail a line check usually have to undergo extra training or, in the worst case, lose their flying license. With so much at stake, flight crews will often display their best performance in front of an evaluator. While pilots are hard-pressed to fake competency–that is, they cannot manipulate performance of what they do not know or understand–processes such as following procedures, leadership, monitor/cross-check, and communication can be staged. As a result, much like QAR data, line checks are limited in identifying safety issues. These data are best suited for uncovering proficiency weaknesses and less so for portraying an accurate snapshot of flight crew behaviors in normal operations.

LOSA: A PROACTIVE SAFETY MEASURE

Most of aviation's understanding of safety performance is based on data concerning adverse safety events, such as those collected from incident reporting and accident investigations (Helmreich et al., 2001; Maurino, 2001). Such data sources are known as reactive measures of safety since their collection is dependent on a negative flight outcome (Reason, 1997). An over-reliance on reactive measures can sometimes induce airlines to declare their safety practices as "working" in the absence of major incidents or accidents (Weick, 1987). This is not to deny the utility of reactive measures in understanding safety but instead, their general over-dependency by airlines highlights a need for a balanced perspective provided by proactive measures.

Proactive safety measures collect data that is independent of flight outcome. For airlines, this means collecting data from flights in normal operations. Therefore, QAR programs and line checks are types of proactive measures.[10] However, as mentioned before, these data sources have limits, with QARs restricted to flight parameter data and

---

[10] QAR programs that only examine parameter exceedances could be considered another type of reactive measure, since they are event-based, requiring airlines to work backwards from a negative occurrence.

line checks best representing proficiency issues. These limitations along with an apparent over-reliance on reactive data sources point to a gap in the measurement of airline safety performance. This dissertation proposes to fill this gap with LOSA and its collection of threat and error management (TEM) data in normal flight operations.

DISSERTATION RESEARCH OBJECTIVES

This dissertation proposes to achieve the following research objectives.

1. Present and explain LOSA methods, instrumentation, procedures, and data quality checks. This includes providing information on its historical development, use of TEM measures, and operating characteristics that have come to distinguish LOSA from other observational methods. Chapter Three accomplishes this objective.
2. Outline and demonstrate LOSA data analysis and interpretation strategies. This includes presenting a multi-stage approach that provides a snapshot of threat and error management performance in its natural context of normal flight operations. Chapter Four accomplishes this objective.
3. Highlight LOSA implications for aviation research and airline safety management practices. This includes discussing future research directions for LOSA and TEM. Chapter Five accomplishes this objective.

Chapter Three: LOSA Methodology

This chapter presents the methodology for the Line Operations Safety Audit (LOSA). The first half of the chapter provides a short history of LOSA development and ten operating characteristics that have come to define its methodology. The second half presents instrumentation, procedures, and data quality checks, which allow researchers to examine the scientific rigor underlying LOSA. Because LOSA is first and foremost a proactive measure of airline safety performance, its applied features are discussed throughout the chapter.

HISTORY OF LOSA DEVELOPMENT

The forerunner to LOSA began in 1994 at the request of Delta Air Lines. After developing a new Crew Resource Management (CRM) course for their line pilots, airline management questioned whether concepts taught in training transferred to the front lines of flying. At the time, Delta's only perspective on CRM performance came from regular line checks, such as those discussed at the end of Chapter Two, and pilot training data collected during flight simulator sessions. While managers agreed these data were good at uncovering proficiency issues, they thought the data fell short in reflecting their true CRM performance standards in normal operations. This prompted a collaborative partnership between Delta and The University of Texas Human Factors Research Project (UTHFRP) to develop an observational audit method that could measure the difference between training and operational CRM performance.

Within three months, a team of observers collected over 400 jump seat observations on routine flights. Each observation required a written narrative by phase of flight (predeparture, takeoff/climb, cruise, and descent/approach/land) and behavioral performance marker ratings (Helmreich, Butler, Taggart, & Wilhelm, 1994). The markers included behaviors believed to be critical for effective CRM performance, such as leadership, communication environment, workload management, and monitoring/cross-checking actions. The phase of flight narratives rounded out the observations with what Geertz (1973) calls "thick description," which provides information about operational context, flight crew performance, and outcomes for each observation. Both sources of data taken together provided the full picture of CRM performance strengths and weaknesses on the line.

As noted in Chapter Two, the 1990's marked the proliferation of systems and human error thinking in aviation (Maurino et al., 1995; Perrow, 1999; Reason, 1997). This period also represented a paradigm shift for UTHFRP. After years of studying CRM behaviors, UTHFRP researchers realized that they could also capture system and human error performance in their observational audits. This shift in thinking stimulated the development of the Threat and Error Management (TEM) framework and coining of the term *Line Operations Safety Audit* (LOSA).

February 1996 marked the first LOSA measuring threat and error management performance. In collaboration with Continental Airlines, the first LOSA collected over 600 domestic and international observations across a four-month period. After the data were analyzed and presented to airline management, solutions were designed and implemented with many of the improvements centering on pilot training curriculum. One example is the development of an error management training course for every pilot at the airline.

Using the 1996 results as a baseline, Continental decided it was time to measure the effectiveness of their solutions with a follow-up LOSA in 2000. The following is an excerpt from an article by Continental Airline's Captain Don Gunther (2002) about their ability to compare LOSA results from 1996 to 2000.

> The 2000 LOSA, when compared to the results of 1996, showed the pilots had not only accepted the principles of error management but incorporated them into everyday operations. LOSA 2000 showed a sizeable improvement in the areas of checklist usage, a 70 percent reduction in non-conforming approaches (i.e., those not meeting stabilized approach criteria), and an increase in overall crew performance. It could be said that Continental had taken a turn in the right direction. (p. 12)

The Continental 1996 and 2000 results provided a "proof of concept" for LOSA. It showed the value of LOSA was not just the diagnostic snapshot of operational performance it provided. It also allowed Continental to benchmark the effectiveness of their organizational safety changes.

LOSA DATA QUALITY CONCERNS

There are several data quality concerns to be aware of when conducting field observations. Some of the more relevant concerns affecting LOSA are data reliability (Reid, 1982), establishing trust with those being observed (J. M. Johnson, 1975), and an accurate coding scheme to classify group processes (Bakeman, 2000). Possibly the most important of LOSA data quality issues is observation reactivity, which occurs when pilots alter their normal behaviors because of an observer's presence in the cockpit. The next section discusses this concern and how it is addressed in LOSA data collection.

Addressing Observation Reactivity in LOSA

Observation reactivity is widely recognized throughout the academic literature as a major threat to the validity of observational findings (Haynes & Horn, 1982; S. Johnson & Bolstad, 1973; Kazdin, 1982). It is an especially sensitive issue with LOSA observations because, to pilots, a cockpit observer usually means a line check. As discussed in Chapter Two, line checks can induce a great deal of pilot apprehension and pressure them to "fake good" for the evaluator (Patterson & Sechrest, 1983; Rosenberg, 1969). Therefore, if pilots sense LOSA observers as line checkers in sheep's clothing, observation reactivity is a real data collection concern. It can also single-handedly defeat the purpose of LOSA as an accurate measure of threat and error management performance in normal operations. Therefore, many of the methods and procedures in LOSA deal solely with this issue.

LOSA AND ITS OPERATING CHARACTERISTICS

The observational research design literature and years of previous field work experience in commercial aviation resulted in ten operating characteristics that have come to define LOSA. These characteristics are listed below, followed by subsections explaining each characteristic's rationale and implementation during a LOSA:

1. Jump seat observations of regularly scheduled flights
2. Voluntary flight crew participation
3. Anonymous, confidential, and non-punitive data collection
4. Joint management/union sponsorship
5. Secure data collection repository
6. Trusted and trained observers

7. Systematic observation instrument
8. Data verification roundtables
9. Data-derived targets for enhancement
10. Feedback of results to line pilots

1. Jump Seat Observations of Regularly Scheduled Flights

Since LOSA is a proactive measure of safety performance in normal operations, observations are limited to regularly scheduled flights. Formal line checks or training evaluation flights where pilots know they are being evaluated do not make for quality LOSA observations. Evaluators already increase stress on the flight crew and adding a LOSA observer to this mix not only adds to this stress but can unnecessarily overcrowd the cockpit. Therefore, it is crucial for LOSA observers to conduct observations on routine flights and explain to the pilots that they are there to collect safety data and not to evaluate performance.

2. Voluntary Flight Crew Participation

All LOSA observations are conducted with flight crew consent. This not only satisfies concerns associated with research ethics, but is another way to distinguish LOSA from line checks, which are compulsory. To gain consent, a LOSA observer should meet the pilots before the flight and ask each one for consent to be observed. If any pilot declines, observers are trained to take another flight with no questions asked. A high number of observation denials could signal a lack of organizational trust in LOSA and warrant an immediate suspension of the project. Fortunately, experience shows that few pilots refuse LOSA observations. Since 1999, the average denial rate for LOSA has been around one for every 100 observations collected.[11] In many projects, there were zero denials, which suggest a high degree of pilot confidence in LOSA and its observers.

3. Anonymous, Confidential, and Non-punitive Data Collection

As with the successful voluntary incident reporting discussed in Chapter Two, pilot trust in LOSA is a precondition for its success. The first building block for developing this trust is the airline's assurance that all data will be kept anonymous, confidential and will never be used for disciplinary reasons.[12] Anonymity is gained by requiring LOSA observers to withhold pilot names, flight numbers, dates or any other identifying bits of information from their observations. The only demographic information collected is which pilot is flying (Captain or First Officer), city pairs (departure and arrival airports), and aircraft type and series (for example, Boeing 737-700). This minimal amount of detail is enough for airlines to diagnose safety issues while at the same time providing a reasonable amount of anonymity protecting pilots against punitive action.

As well as preserving pilot anonymity, all LOSA observations are kept confidential. Observers are strictly prohibited from talking about their observations in private or public due to concerns about accidentally identifying flight crews. Observer identity is also withheld from LOSA data. This promotes more open and accurate reporting and avoids the possibility of airline managers asking observers for information about a problematic event collected during LOSA.

---

[11] These data were gathered by the author of this dissertation. LOSA observers were e-mailed during or at the end of data collection and asked to report any denials. No identifying information about the pilots or flights was collected.

[12] As with many voluntary incident reporting systems and Quick Access Recorder (QAR) programs, non-punitive protection does not extend to deliberate or criminal acts of negligence.

4. Joint Management/Pilot Association Sponsorship

Airline assurance of anonymous, confidential, and non-punitive data collection is best formalized with a letter of agreement between management and the pilot's association. This letter symbolizes a partnered sponsorship of LOSA and its intent to collect safety data that benefits all stakeholders. Once a drafted letter outlines the purpose of LOSA, it is jointly signed by both parties and sent to every pilot in the airline. Such public announcement has shown to be a powerful mechanism in strengthening trust and credibility in an observational project (Lincoln & Guba, 1985).

5. Secure Data Collection Repository

Another layer of protection for confidentiality is to make certain all LOSA observations are collected and maintained in a secure data repository. LOSA data repositories can be located within the airline, such as those used for voluntary incident reporting, or managed off-site by a neutral third party. Regardless of location, pilots must believe there is no chance of individual observations being misplaced or improperly spread throughout the airline.

6. Trusted and Trained Observers

The fundamental element of LOSA success is quality observers. It is critical that observers be respected and trusted within their airline to ensure full pilot acceptance of LOSA. Kirmeyer (1985) has suggested that observation reactivity is a "function of the observer's acceptance by their peers" (p. 370). Therefore, managers are not good candidates because of their authority over flight crews. The best candidates for LOSA observers are regular line pilots. This common ground between pilots as observers and observees eases suspicions, allows quicker adaptation to the observer, and lessens the potential for reactivity (see Kotarba, 1980; Snow, Benford, & Anderson, 1986). In addition, the operational familiarity and technical expertise of pilot observers arguably enables them to collect more complete and accurate LOSA observations than observers with less familiarity with the flying environment. However, the downside is that pilot observers are initially inexperienced, which places a greater importance on their selection and training. This issue is discussed in more detail later in this chapter.

7. Systematic Observation Instrument

Using pilots to conduct LOSA observations means that personal biases can sometimes slip in. This results in judgments of "what flight crews should have done" instead of objective descriptions of "what they did" during a flight (Singleton, Straits, & Straits, 1993). The latter are critical if LOSA observations are to be reliable and valid representations of flight crew performance in normal operations (Herbert & Attridge, 1975; Weick, 1968). This assessment of reliability and validity is made easier with structured, systematic data collection (Hartman, 1982; Heyns & Lippitt, 1954).[13] LOSA observations are systematic since LOSA measurement is based on the TEM framework. The classifications provided by TEM prompt observers as to what to look for, record, and code during an observation (Weick, 1968). Later in this chapter, LOSA instrumentation and the TEM classifications are discussed at length.

---

[13] Systematic measurement in field work is when quantifiable categories are used to simplify events or behaviors within their natural context (Weick, 1968).

8. Data Verification Roundtables

Data-driven programs like LOSA need rigorous data management procedures and consistency checks. One such technique to improve data consistency is to have members of the group from which data were collected verify their correctness (Lincoln & Guba, 1985). For LOSA, this is known as the data verification roundtable, which takes place before performing any data analysis. A roundtable consists of three or four airline managers and pilot association members that scan LOSA TEM data for mistakes. For example, an observer might code something as an aircraft malfunction threat, but the written narrative describes the threat as an ATC controller error. In this case, the narrative takes precedence and the roundtable members would correct the coding in the database. The result is a consistent and accurate database according to government flying regulations, airline policies, and standard operating procedures.

9. Data-Derived Targets for Enhancement

During LOSA data analysis, operational trends can help airlines identify active failures, latent conditions and other safety performance issues. These trends lead to forming data-derived targets for enhancement that airlines may want to address with organizational change. As with the Continental Airlines LOSA “proof of concept” story mentioned earlier, targets for enhancement in one LOSA can serve as benchmark indicators for a follow-up LOSA to measure the effectiveness of organizational change. More discussion of targets for enhancement is in Chapter Four of this dissertation, which presents LOSA data analysis and interpretation strategies.

10. Feedback of Results to Line Pilots

To ensure long-term acceptance of LOSA, airline management has an obligation to report LOSA results and targets for enhancement back to the pilots. This is typically done with special airline publications or group presentations. Whatever the method, the results must be communicated in a timely fashion. Experience has shown such feedback improves pilot acceptance of LOSA and the targets it produces. Of course, eventual organizational action on the targets and improvement is arguably the best way to ensure long-term support for LOSA.

Summary

LOSA operating characteristics originate from previous work in observational research design and much collaboration with airlines, pilots, researchers, and other aviation safety professionals. The first of these characteristics–jump seat observations during regularly scheduled flights–represents the over-arching purpose of LOSA: to capture flight crew performance in its natural context. To achieve this objective, trust in LOSA as a process and trust in LOSA data are needed.

Trust in LOSA is mostly gained by the characteristics of voluntary flight crew participation; anonymous, confidential, non-punitive data collection; joint management/union sponsorship; a secure data collection repository; data-derived targets for enhancement; and feedback of results to line pilots. These characteristics build an environment of trust by making the method of LOSA transparent and its purpose clear. If airlines fail to earn the trust of pilots, LOSA is nothing more than an elaborate line check, which wastes an opportunity to gain a unique understanding of performance in normal operations.

Trust in LOSA data is produced by reliable results and valid representations of system and flight crew performance. The remaining LOSA operating characteristics help ensure TEM data quality: a systematic observation instrument; trusted and trained observers; and data verification roundtables. The next section provides a more complete examination of these characteristics and their critical role in LOSA.

LOSA DATA COLLECTION OVERVIEW

LOSA data collection extends beyond simply asking observers to sit in the cockpit and record what they see about safety performance. To extract meaningful interpretations, field observations are: planned (Singleton et al., 1993), methodical (Weick, 1985), and quantifiable (Hawkins, 1982). These conditions are achieved through instrumentation, observers, and procedural protocol.

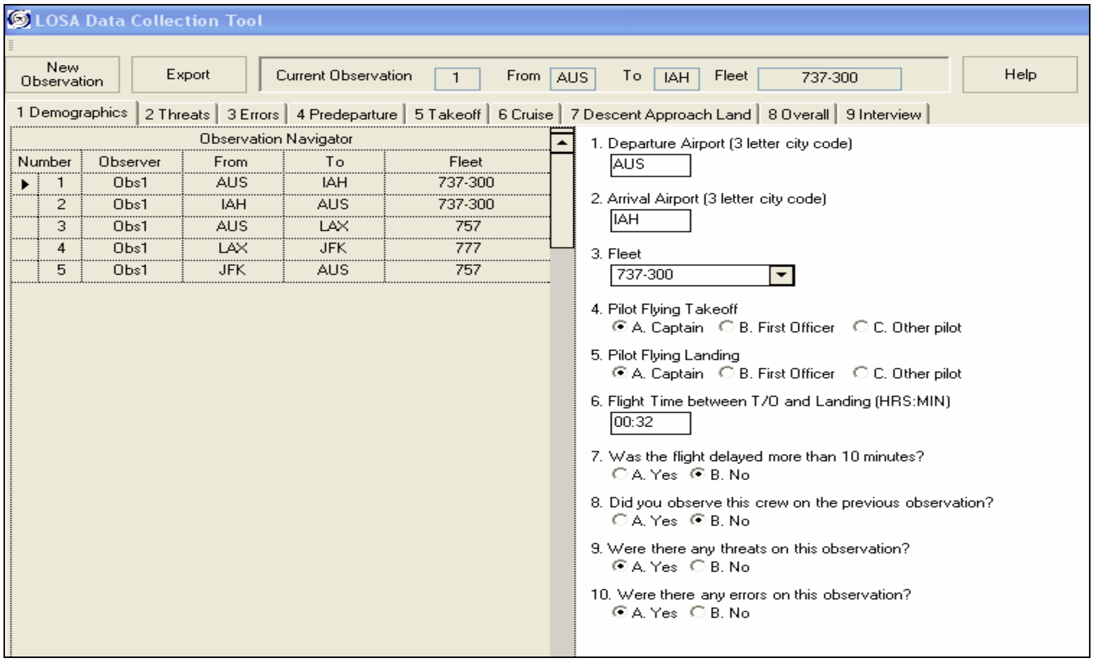
LOSA INSTRUMENTATION

As stated in the operating characteristics, LOSA uses a targeted observation instrument based on the TEM framework. The most current version of this instrument is a custom piece of software called the LOSA Data Collection Tool.[14] This software stems from an older, paper-based measure called the LOSA Observation Form. In presenting the most current LOSA data collection method, the discussion in this dissertation will be limited to the LOSA Data Collection Tool. However, for those readers interested, the LOSA Observation Form is in Appendix A. It contains the same TEM measures and is

---

[14] The LOSA Data Collection Tool is the property of the LOSA Collaborative, which is a private organization that implements LOSA worldwide. The LOSA Collaborative is formally partnered with The University of Texas through Intellectual Property Agreement #01-034. In return for exclusive license, The University of Texas receives a percentage of LOSA Collaborative gross revenue and data, de-identified by airline, to support research.

an effective, although more time-consuming alternative to the software. The opening screen of the LOSA Data Collection Tool is in Illustration 3.1.



Current Observation 1 From AUS To IAH Fleet 737-300

Observation Navigator

Number	Observer	From	To	Fleet
1	Obs1	AUS	IAH	737-300
2	Obs1	IAH	AUS	737-300
3	Obs1	AUS	LAX	757
4	Obs1	LAX	JFK	777
5	Obs1	JFK	AUS	757

1. Departure Airport (3 letter city code)
AUS

2. Arrival Airport (3 letter city code)
IAH

3. Fleet
737-300

4. Pilot Flying Takeoff
● A. Captain ○ B. First Officer ○ C. Other pilot

5. Pilot Flying Landing
● A. Captain ○ B. First Officer ○ C. Other pilot

6. Flight Time between T/O and Landing (HRS:MIN)
00:32

7. Was the flight delayed more than 10 minutes?
○ A. Yes ● B. No

8. Did you observe this crew on the previous observation?
○ A. Yes ● B. No

9. Were there any threats on this observation?
● A. Yes ○ B. No

10. Were there any errors on this observation?
● A. Yes ○ B. No

Illustration 3.1 – LOSA Data Collection Tool: Opening Screen

There are added benefits to using a custom software application for LOSA data collection. The first is better data security. Entry into the LOSA Data Collection Tool is password protected and all data are saved as encrypted files that can only be read by those with special decryption keys. This not only strengthens confidentiality assurances to pilots but also helps keep the data safe from improper distribution. Possibly the most important benefit of the LOSA Data Collection Tool is better data accuracy. With software applications, designers can program checks that remove out-of-range data entries, and spell-check narrative data. Software also allows for response-determined guidance for observers in their coding of TEM events. For example, when an observer codes “undetected” as a flight crew error response, the software will automatically code "nobody" for the variable that asks for those responsible in detecting the error. All of these features help keep data clean and observers standardized.

LOSA DATA MEASURES

Data collected with LOSA instrumentation can be divided into four types: (1) demographic, (2) narrative, (3) crew resource management (CRM) behavioral markers, and (4) threat and error management measures. As discussed in Chapter Two, the intense focus on CRM performance in aviation teamwork research has yielded much information on behavioral marker evaluation. As such, the CRM markers used in LOSA will not be discussed here; rather, the dissertation will focus on the new demographic, narrative, and threat and error management measures. For more information on these markers, readers are referred to Helmreich, Butler, Taggart and Wilhelm (1994), Law and Sherman (1995), and Hines (1998). They are also shown on the LOSA Observation Form in Appendix A.

Demographic Measures

Demographic measures capture several flight characteristics. As shown in Figure 2.1, they are entered on the opening screen and include:

- City pairs (departure and arrival city)
- Aircraft type and series (e.g., Boeing 737-400)
- Pilot flying for takeoff and landing (Captain or First Officer)
- Flight time (length of time between takeoff and landing)
- Delayed flight (was the flight delayed more than 10 minutes? Yes/No)[15]

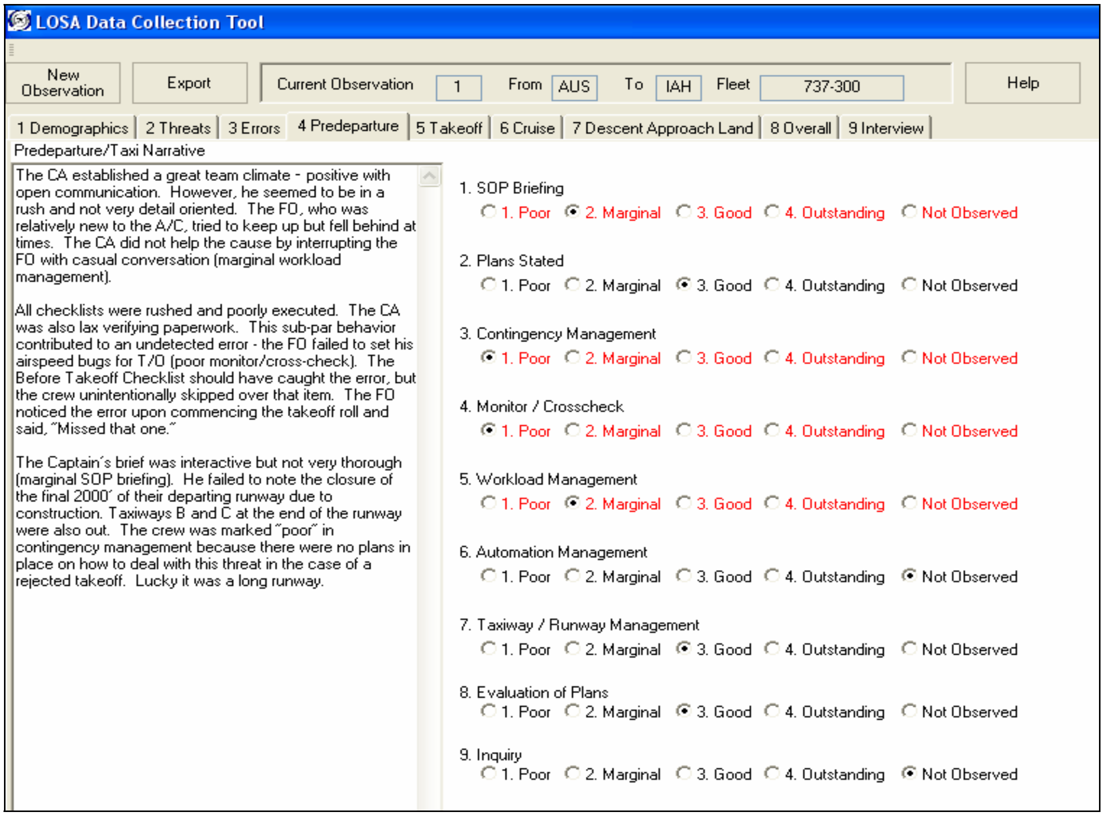
---

[15] The delayed departure variable is included in the LOSA Data Collection Tool for a future study to compare threat and error management performance during predeparture between on-time and delayed flights.

Narrative Measures

LOSA narrative measures have several purposes. Recognizing that novice observers might struggle with coding, narratives encourage observers to “write the story” of the flight. The objective nature of the TEM framework makes it possible to retrospectively code an event provided the narrative is thorough and clearly written. It must be realized that even the most comprehensive of coding schemes cannot fully capture the complexities between context and flight crew performance (Campbell, 1961; Hawkins, 1982). Narrative measures get a little closer because they are free form and unrestricted. As a result, LOSA observations can provide insight that is not possible with an exclusive reliance on categorical coding.

There are four flight phase narratives (predeparture/taxi-out, takeoff/climb, cruise, and descent/approach/land) and an overall narrative that provides an opportunity for observers to provide impressions and interpretations of performance. Illustration 3.2 provides a screen shot of the Predeparture tab in the LOSA Data Collection Tool with a sample narrative and CRM behavioral marker measures.



Current Observation 1 From AUS To IAH Fleet 737-300

Predeparture/Taxi Narrative

The CA established a great team climate - positive with open communication. However, he seemed to be in a rush and not very detail oriented. The FO, who was relatively new to the A/C, tried to keep up but fell behind at times. The CA did not help the cause by interrupting the FO with casual conversation (marginal workload management).

All checklists were rushed and poorly executed. The CA was also lax verifying paperwork. This sub-par behavior contributed to an undetected error - the FO failed to set his airspeed bugs for T/O (poor monitor/cross-check). The Before Takeoff Checklist should have caught the error, but the crew unintentionally skipped over that item. The FO noticed the error upon commencing the takeoff roll and said, "Missed that one."

The Captain's brief was interactive but not very thorough (marginal SOP briefing). He failed to note the closure of the final 2000' of their departing runway due to construction. Taxiways B and C at the end of the runway were also out. The crew was marked "poor" in contingency management because there were no plans in place on how to deal with this threat in the case of a rejected takeoff. Lucky it was a long runway.

1. SOP Briefing
○ 1. Poor ● 2. Marginal ○ 3. Good ○ 4. Outstanding ○ Not Observed

2. Plans Stated
○ 1. Poor ○ 2. Marginal ● 3. Good ○ 4. Outstanding ○ Not Observed

3. Contingency Management
● 1. Poor ○ 2. Marginal ○ 3. Good ○ 4. Outstanding ○ Not Observed

4. Monitor / Crosscheck
● 1. Poor ○ 2. Marginal ○ 3. Good ○ 4. Outstanding ○ Not Observed

5. Workload Management
○ 1. Poor ● 2. Marginal ○ 3. Good ○ 4. Outstanding ○ Not Observed

6. Automation Management
○ 1. Poor ○ 2. Marginal ○ 3. Good ○ 4. Outstanding ● Not Observed

7. Taxiway / Runway Management
○ 1. Poor ○ 2. Marginal ● 3. Good ○ 4. Outstanding ○ Not Observed

8. Evaluation of Plans
○ 1. Poor ○ 2. Marginal ● 3. Good ○ 4. Outstanding ○ Not Observed

9. Inquiry
○ 1. Poor ○ 2. Marginal ○ 3. Good ○ 4. Outstanding ● Not Observed

Illustration 3.2 – LOSA Data Collection Tool: Predeparture Tab

Threat and Error Management (TEM) Measures

TEM performance in LOSA is captured with qualitative descriptions and quantitative coding. The next three sections of this chapter present the variables used in recording threats, errors, undesired aircraft states, and their management.

Threat Management: Type, Response, and Outcome

There are several types of threats that a flight crew can face during a flight. Some of these can be anticipated, such as high terrain at an airport, and others can happen unexpectedly, such as ATC controller error. In LOSA data collection, the different kinds of threats are captured with a variable called *Threat Type*. However, management only occurs when a flight crew responds to a threat. This is measured by *Threat Response*. Finally, the effectiveness of threat management is measured by *Threat Outcome*. All three of these variables together define the primary units of threat management performance in LOSA.

Threat Types – A threat is defined as an event, error, or aircraft state that occurs outside the control of the flight crew but still requires their management to maintain safety. There are two types of threat: *environmental threats* that occur outside an airline's direct control, and *airline threats* that originate within flight operations. Table 3.1 provides descriptions and examples of environmental and airline threats collected during LOSA observations.

Table 3.1 – LOSA Threat Types and Examples

Environmental Threats	Examples
Adverse Weather	Thunderstorms, turbulence, poor visibility, wind shear, icing conditions, IMC
Airport	Poor signage, faint markings, runway/taxiway closures, INOP navigational aids, poor braking action, contaminated runways/taxiways
ATC	Tough-to-meet clearances/restrictions, reroutes, language difficulties, controller errors
Environmental Operational Pressure	Terrain, traffic, TCAS TA / RA, radio congestion
Airline Threats	**Examples**
Aircraft Malfunction	Systems, engines, flight controls, or automation anomalies detected by the flight crew, or MEL items with operational implications
Airline Operational Pressure	Operational time pressure, diversion, late-arriving airplane
Cabin	Cabin events, flight attendant errors, distractions, interruptions
Dispatch/Paperwork	Load sheet errors, crew scheduling events, late paperwork changes or errors
Ground/Ramp	Aircraft loading events, fueling errors, agent interruptions, improper ground support, de-icing
Ground Maintenance	Aircraft repairs on ground, maintenance log problems, maintenance errors
Manuals/Charts	Missing information or documentation errors

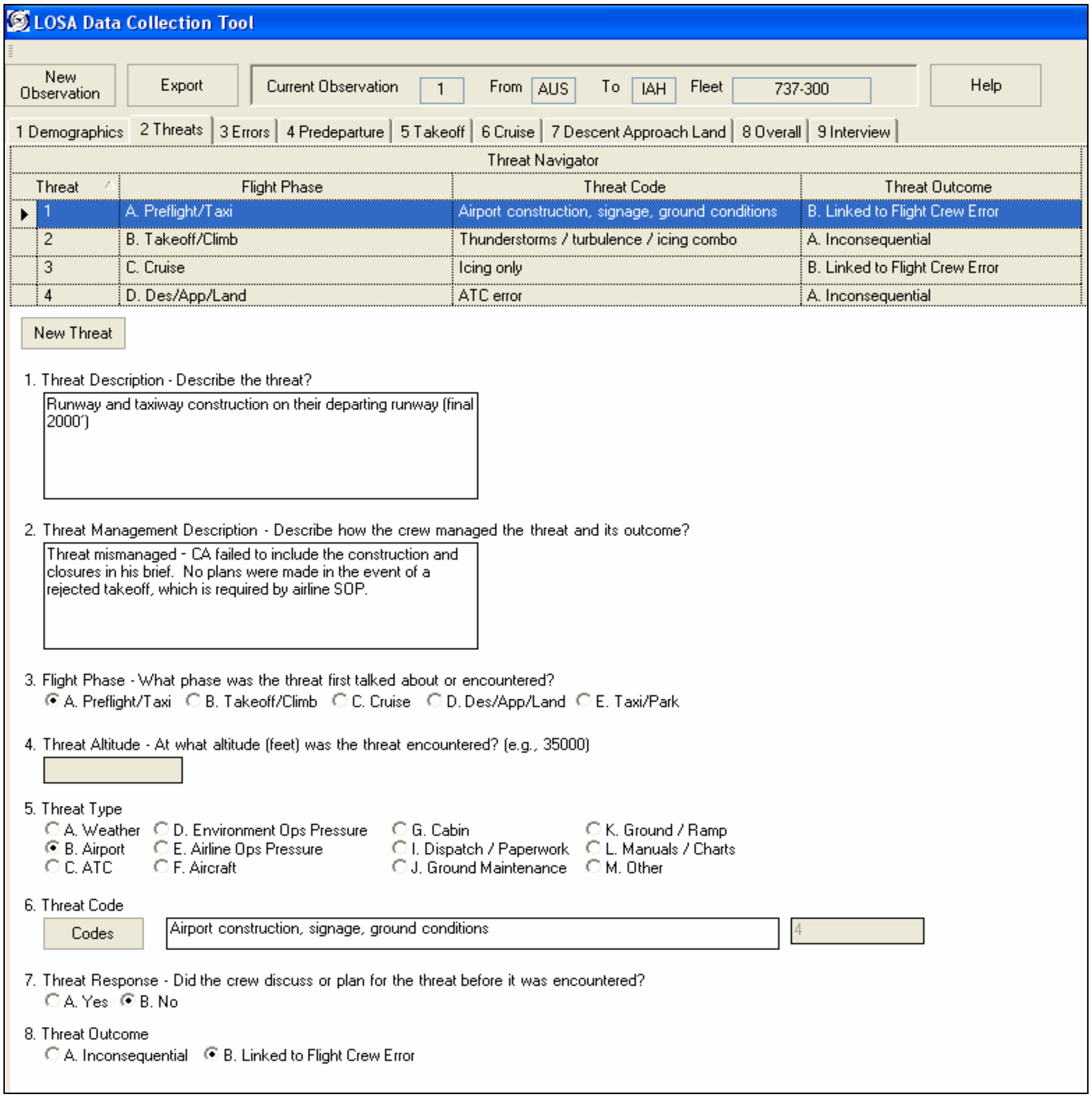
In all, there are four categories of environmental threat types and seven categories of airline threat types that can be selected for coding. Once an observer selects a threat type in the LOSA Data Collection Tool, a set of sub-categories is provided. For example, if an observer chooses to code an environmental threat of adverse weather, sub-categories appear in the software that further describes adverse weather, such as thunderstorms, icing, and wind shear. These sub-categories with their parent threat types are in Appendix B. In providing further detail, observers also record the phase of flight and altitude when the threat is first encountered or discussed.

Threat Responses – The first step in threat management is the flight crew response or lack of it. This is measured by a "yes/no" question that asks whether the pilots discussed or planned for the threat before it was encountered. A "yes" answer shows the flight crew anticipated a threat before having to manage it. An example is a flight crew that expects weather at their destination and asks the ground crew to load an extra 3000 pounds fuel for a possible diversion. A "no" answer shows a threat arose unexpectedly or the flight crew failed to plan for it before it was managed. For example, an unexpected threat could be an in-flight aircraft malfunction, such as a loss of pressurization, or ground crew members failing to close an aircraft cargo door before pushback.

Threat Outcomes – The final quantitative measure of threat management is outcome, for which there are two choices. The first possible threat outcome is *inconsequential*. This is defined as a threat that is either successfully managed or fails to contribute to a crew error or undesired aircraft state. An example of an inconsequential threat is ATC giving the flight crew a wrong heading toward another aircraft, with a Captain subsequently detecting the controller error because of a possible traffic conflict. For threat response, the threat was unexpected but effectively managed by the Captain.

The second threat outcome is *linked to flight crew error,* which also means a threat is mismanaged. An example of a mismanaged threat is a flight attendant who interrupts the flight crew while completing a checklist. As a result, crew members miss an item and forget to set their altimeters for descent. This is a mismanaged cabin threat linked to an outcome of flight crew error. A common management procedure for such a situation is for pilots to start a checklist over when interrupted or distracted.

Additional Threat Variables – Besides coding threats and their management, observers also write descriptions of the threat and how it was managed by the flight crew. They provide a written account of threats and their management, often giving more details on contextual influences that are hard to capture with a categorical coding scheme. As with the phase of flight narratives, these descriptions are useful during data verification roundtables to check observer coding. A screenshot of the Threats tab in the LOSA Data Collection Tool is in Illustration 3.3. Table 3.2 follows, which presents a review of primary threat management variables and their coding categories.



Current Observation 1 From AUS To IAH Fleet 737-300

Threat Navigator

Threat	Flight Phase	Threat Code	Threat Outcome
1	A. Preflight/Taxi	Airport construction, signage, ground conditions	B. Linked to Flight Crew Error
2	B. Takeoff/Climb	Thunderstorms / turbulence / icing combo	A. Inconsequential
3	C. Cruise	Icing only	B. Linked to Flight Crew Error
4	D. Des/App/Land	ATC error	A. Inconsequential

1. Threat Description - Describe the threat?

Runway and taxiway construction on their departing runway (final 2000')

2. Threat Management Description - Describe how the crew managed the threat and its outcome?

Threat mismanaged - CA failed to include the construction and closures in his brief. No plans were made in the event of a rejected takeoff, which is required by airline SOP.

3. Flight Phase - What phase was the threat first talked about or encountered?

● A. Preflight/Taxi ○ B. Takeoff/Climb ○ C. Cruise ○ D. Des/App/Land ○ E. Taxi/Park

4. Threat Altitude - At what altitude (feet) was the threat encountered? (e.g., 35000)

5. Threat Type

○ A. Weather ○ D. Environment Ops Pressure ○ G. Cabin ○ K. Ground / Ramp
● B. Airport ○ E. Airline Ops Pressure ○ I. Dispatch / Paperwork ○ L. Manuals / Charts
○ C. ATC ○ F. Aircraft ○ J. Ground Maintenance ○ M. Other

6. Threat Code

Airport construction, signage, ground conditions 4

7. Threat Response - Did the crew discuss or plan for the threat before it was encountered?

○ A. Yes ● B. No

8. Threat Outcome

○ A. Inconsequential ● B. Linked to Flight Crew Error

Illustration 3.3 – LOSA Data Collection Tool: Threats Tab

Table 3.2 – Primary Threat Management Variables

Variable Name	Categorical Levels
Threat Type	Environmental Threats Adverse weather Airport ATC Environmental operational pressure Airline Threats Aircraft malfunction Airline operational pressure Cabin Dispatch/paperwork Ground/ramp Ground maintenance Manuals/charts
Threat Response	Was the threat discussed or planned for before it was encountered? Yes / No
Threat Outcome	Inconsequential Linked to flight crew error

Error Management: Type, Response, and Outcome

In LOSA, flight crew errors can be a momentary slip or lapse and they can also be a by-product of the threat environment. For example, a challenging ATC clearance might encourage a procedural shortcut that results in error, or a gate agent might interrupt the crew with a passenger issue during predeparture when on-time departure pressure demands efficiency. Like threats and their management, error management begins with error commission, followed by a flight crew response and outcome.

Error Types – An error within the TEM framework is defined as crew action or inaction that leads to a deviation from crew or organizational intentions or expectations.

Errors can be divided into three categories: *aircraft handling*, *procedural* and *communication* errors. Aircraft handling errors involve those that are directly linked to the flying, direction, speed and configuration of the aircraft. They can involve automation keypunch errors or pilots failing to control speed during an approach for landing. Procedural errors are defined as pilot deviations from government regulations or airline standard operating procedures. These can include errors with performing checklists, forgetting to make procedural cross-verification of automation or conducting an incorrect takeoff briefing. Finally, communication errors involve poor or absent communication between pilots or from pilots to external agents, such as flight attendants, ground personnel, or ATC controllers. Table 3.3 shows LOSA error types and examples used in data collection.

Table 3.3 – LOSA Flight Crew Error Types and Examples

Aircraft Handling Errors	Examples
Automation	Incorrect altitude, speed, heading, autothrottle settings, mode executed, or entries
Flight Control	Incorrect flaps, speed brake, autobrake, thrust reverser or power settings
Ground Navigation	Attempting to turn down wrong taxiway/runway Missed taxiway/runway/gate
Manual Handling	Hand flying vertical, lateral, or speed deviations Missed runway/taxiway failure to hold short, or taxi above speed limit
Systems/Radio/Instruments	Incorrect packs, altimeter, fuel switch settings, or radio frequency dialed
Procedural Errors	Examples
Briefings	Missed items in the brief, omitted departure, takeoff, approach, or handover briefing
Callouts	Omitted takeoff, descent, or approach callouts
Checklist	Performed checklist from memory or omitted checklist Missed items, wrong challenge and response, performed late or at wrong time
Documentation	Wrong weight and balance, fuel information, ATIS, or clearance recorded Misinterpreted items on paperwork
Pilot Flying (PF)/Pilot Not Flying (PNF) Duty	PF makes own automation changes, PNF doing PF duties, PF doing PNF duties
SOP Cross-verification	Intentional and unintentional failure to cross-verify automation inputs
Other Procedural	Other deviations from government regulations or standard operating procedures
Communication Errors	Examples
Crew to External	Missed calls, misinterpretation of instructions, or incorrect read-backs to ATC Wrong clearance, taxiway, gate or runway communicated
Pilot to Pilot	Within-crew miscommunication or misinterpretation

In all, there are five categories of aircraft handling error types, seven categories of procedural error types, and two categories of communication error types. As with threat types, error types have a set of sub-categories that appear in Appendix C.

Error Responses – In LOSA and under the TEM framework, there are two types of error responses. The first is *detected and action*, which is recorded when a flight crew detects an error and actively tries to manage it. An example is an aircraft climbing from 24,000 to 25,000 feet and the Captain forgetting to make an airline mandated callout of 1000 feet to level. However, the error is detected and corrected by the First Officer, who makes the callout for the Captain.

The other error response in LOSA data collection is *failing to respond*, which is recorded when a flight crew fails to detect or ignores an error, leaving it unmanaged. An example of ignored error with no error management is when a First Officer decides to complete a checklist by memory instead of visually referring to the checklist as required by airline standard operating procedures. The Captain notices but opts not to correct the error. An example of an undetected error is when a Captain accidentally flies 20 degrees off heading during a Standard Instrument Departure (SID) and the deviation goes unnoticed by the flight crew. However, it is detected by ATC, which alerts the crew of their mistake.

Error Outcomes – The third piece of information needed to define error management is the error outcome. There are three types of outcomes: *inconsequential, additional error,* and *undesired aircraft state*. An inconsequential error is similar to an inconsequential threat in that it shows an error is either successfully managed or fails to lead to an additional crew error or undesired aircraft state. An example of an inconsequential error is when a Captain notices and tells a First Officer that he failed to turn on the aircraft's pressurization packs before takeoff.

The next type of outcome is an error that is linked to an additional error, which in aviation is often referred to as an error chain. An example of an error chain is when a

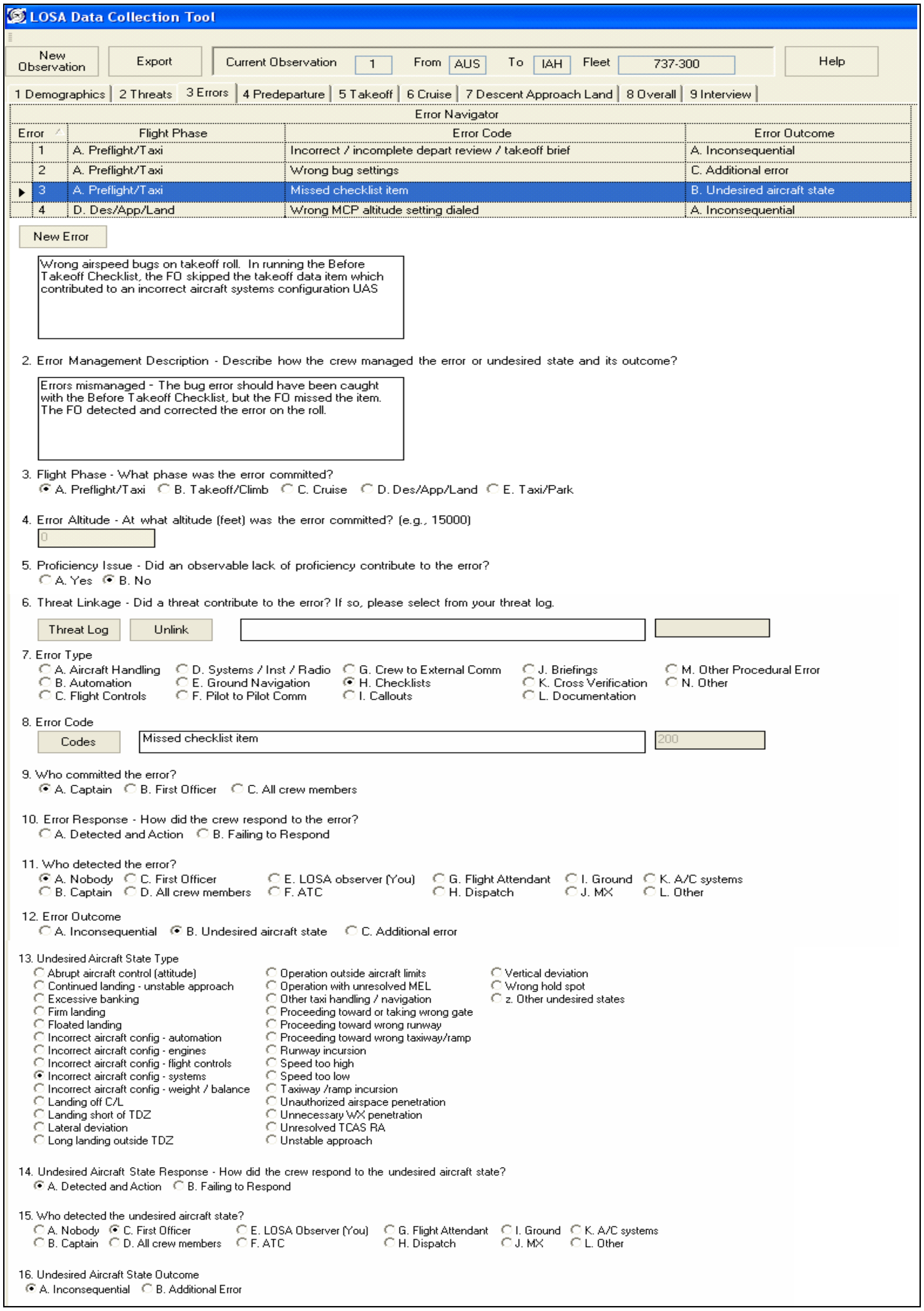
Captain, on receiving clearance from ATC to descend from 21,000 feet to 10,000 feet, dials a wrong altitude of 8,000 feet into the autopilot, the first error. This is followed by a second error by the First Officer, who fails to cross-verify the setting as required by airline standard operating procedures. This scenario is an automation error linked to a procedural error of failing to cross-verify the setting.

The final and perhaps most dangerous error outcome is the undesired aircraft state (UAS). An example of an error leading to an undesired aircraft state is a First Officer manually flying the aircraft and accidentally exceeding an ATC speed restriction, which, after a few moments, is detected by the Captain.[16] More details on undesired aircraft states and their management are provided in the next section.

Additional Error Management Variables – As they do for threats, observers write a description of each error and its management. Other variables collected are phase of flight and altitude when the error was committed, who caused and detected the error, and whether the error linked to a threat or revealed an observable lack of proficiency. Illustration 3.4 provides a screenshot of the Errors tab in the LOSA Data Collection Tool and Table 3.4 is a summary of the primary error management variables and their coding categories.

---

[16] ATC controllers will often instruct flight crews to fly at a certain speed before crossing a particular navigational aid to maintain separation from other aircraft traffic.



Current Observation 1 From AUS To IAH Fleet 737-300

Error Navigator

Error	Flight Phase	Error Code	Error Outcome
1	A. Preflight/Taxi	Incorrect / incomplete depart review / takeoff brief	A. Inconsequential
2	A. Preflight/Taxi	Wrong bug settings	C. Additional error
3	A. Preflight/Taxi	Missed checklist item	B. Undesired aircraft state
4	D. Des/App/Land	Wrong MCP altitude setting dialed	A. Inconsequential

Wrong airspeed bugs on takeoff roll. In running the Before Takeoff Checklist, the FO skipped the takeoff data item which contributed to an incorrect aircraft systems configuration UAS

2. Error Management Description - Describe how the crew managed the error or undesired state and its outcome?

Errors mismanaged - The bug error should have been caught with the Before Takeoff Checklist, but the FO missed the item. The FO detected and corrected the error on the roll.

3. Flight Phase - What phase was the error committed?
● A. Preflight/Taxi ○ B. Takeoff/Climb ○ C. Cruise ○ D. Des/App/Land ○ E. Taxi/Park

4. Error Altitude - At what altitude (feet) was the error committed? (e.g., 15000)
0

5. Proficiency Issue - Did an observable lack of proficiency contribute to the error?
○ A. Yes ● B. No

6. Threat Linkage - Did a threat contribute to the error? If so, please select from your threat log.

7. Error Type
○ A. Aircraft Handling
○ B. Automation
○ C. Flight Controls
○ D. Systems / Inst / Radio
○ E. Ground Navigation
○ F. Pilot to Pilot Comm
○ G. Crew to External Comm
● H. Checklists
○ I. Callouts
○ J. Briefings
○ K. Cross Verification
○ L. Documentation
○ M. Other Procedural Error
○ N. Other

8. Error Code
Missed checklist item 200

9. Who committed the error?
● A. Captain ○ B. First Officer ○ C. All crew members

10. Error Response - How did the crew respond to the error?
○ A. Detected and Action ○ B. Failing to Respond

11. Who detected the error?
● A. Nobody
○ B. Captain
○ C. First Officer
○ D. All crew members
○ E. LOSA observer (You)
○ F. ATC
○ G. Flight Attendant
○ H. Dispatch
○ I. Ground
○ J. MX
○ K. A/C systems
○ L. Other

12. Error Outcome
○ A. Inconsequential ● B. Undesired aircraft state ○ C. Additional error

13. Undesired Aircraft State Type
○ Abrupt aircraft control (attitude)
○ Continued landing - unstable approach
○ Excessive banking
○ Firm landing
○ Floated landing
○ Incorrect aircraft config - automation
○ Incorrect aircraft config - engines
○ Incorrect aircraft config - flight controls
● Incorrect aircraft config - systems
○ Incorrect aircraft config - weight / balance
○ Landing off C/L
○ Landing short of TDZ
○ Lateral deviation
○ Long landing outside TDZ
○ Operation outside aircraft limits
○ Operation with unresolved MEL
○ Other taxi handling / navigation
○ Proceeding toward or taking wrong gate
○ Proceeding toward wrong runway
○ Proceeding toward wrong taxiway/ramp
○ Runway incursion
○ Speed too high
○ Speed too low
○ Taxiway /ramp incursion
○ Unauthorized airspace penetration
○ Unnecessary WX penetration
○ Unresolved TCAS RA
○ Unstable approach
○ Vertical deviation
○ Wrong hold spot
○ z. Other undesired states

14. Undesired Aircraft State Response - How did the crew respond to the undesired aircraft state?
● A. Detected and Action ○ B. Failing to Respond

15. Who detected the undesired aircraft state?
○ A. Nobody
○ B. Captain
● C. First Officer
○ D. All crew members
○ E. LOSA Observer (You)
○ F. ATC
○ G. Flight Attendant
○ H. Dispatch
○ I. Ground
○ J. MX
○ K. A/C systems
○ L. Other

16. Undesired Aircraft State Outcome
● A. Inconsequential ○ B. Additional Error

Illustration 3.4 – LOSA Data Collection Tool: Errors Tab

Table 3.4 – Primary Error Management Variables

Variable Name	Categorical Levels
Error Type	Aircraft Handling Errors Automation Flight control Ground navigation Manual handling Systems/Radio/Instruments Procedural Errors Briefings Callout Checklist Documentation PF/PNF duty SOP cross-verification Other procedural Communication Errors Crew to External Pilot to Pilot
Error Response	Detected and action Failing to respond
Error Outcome	Inconsequential Additional error Undesired aircraft state

Undesired Aircraft State Management: Type, Response, and Outcome

When an error leads to an undesired aircraft state (UAS), flight crews are no longer managing the error that contributed to the state but managing the state itself. This is obvious in the UAS example in the preceding error management section when the First Officer exceeded an ATC speed restriction. The Captain in this case detected the high-speed UAS, not the manual flying error that induced the state. The manual flying error could have been managed if either the Captain or First Officer had noticed and made the correction before exceeding the speed limit. This means that undesired aircraft states have their own form of flight crew management that is independent of error management performance.

Undesired Aircraft State Types – An undesired aircraft state (UAS) is defined as a crew-error-induced aircraft state that clearly reduces safety margins. There are three types of UAS: *aircraft handling, ground navigation states,* and *incorrect aircraft configurations*. Table 3.5 shows the three types of undesired aircraft states with examples. A full listing of undesired aircraft state types and sub-categories is in Appendix D.

Table 3.5 – LOSA Undesired Aircraft State Types and Examples

Undesired Aircraft State Types	Examples
Aircraft Handling	Vertical, lateral or speed deviations Unnecessary weather penetration Unstable approach Long, floated, firm or off-centerline landings
Ground Navigation	Runway/taxiway incursions Wrong taxiway, ramp, gate, or hold spot Taxi above speed limit
Incorrect Aircraft Configuration	Automation, engine, flight control, systems, or weight/balance events

Undesired Aircraft State Responses – UAS responses are the same as error management: *detected and action* and *failing to respond.* An example of a detected UAS is when a Captain enters icing conditions for over ten minutes, and then realizes he forgot to turn on the engine anti-ice. An example of a crew failing to respond to an undesired aircraft state is a First Officer exceeding glide slope parameters for an unstable approach at 400 feet. The Captain calls out the deviation but the First Officer decides to continue the approach and land.

Undesired Aircraft State Outcomes – UAS outcomes are similar to those used in error management performance, with one difference. Since undesired aircraft states are defined as crew-error-induced situations, there cannot be an undesired aircraft state that links to another state without a mediating crew error. Therefore, the two outcomes for undesired aircraft states are *inconsequential* and *linked to additional error*. An example of an inconsequential UAS is a final approach during which a Captain allows the sink rate to exceed the permissible limit of 1,000 feet per minute. The Captain detects the deviation and immediately corrects it. An example of UAS linked to an additional flight crew error is when a Captain exceeds an ATC speed limit on final approach and uses the speed brake to slow the airplane, but leaves the brake extended. The error is caught by the First Officer who immediately stows the brake and prevents a possible speed too low UAS late in the approach.

Additional Undesired Aircraft State Variables – The written descriptions of undesired aircraft states are folded in with what observers write about errors and their management. The only extra variable is asking the observer to record the person responsible for detecting the UAS.[17] All UAS management variables are coded on the Errors Tab in the LOSA Data Collection Tool shown before in Illustration 3.4. Table 3.6 presents the primary UAS management variables and their coding categories.

---

[17] Some readers might wonder why there is no "who detected" variable for threat management performance. It was dropped early in LOSA development because it was thought not to add much to the diagnostic value of LOSA data and was kept out to conserve on the amount of observer coding. However, future research on the "who detected" variables in error and UAS management might lead to a reconsideration.

Table 3.6 – Primary Undesired Aircraft State Management Variables

Variable Name	Categorical Levels
UAS Type	Aircraft handling Ground navigation Incorrect aircraft configurations
UAS Response	Detected and action Failing to respond
UAS Outcome	Inconsequential Additional error

The next section of this chapter shifts from LOSA instrumentation to the observers. As mentioned in the LOSA operating characteristics, observers need to be trusted and trained. Thus, effective selection and quality observer training are key to a successful LOSA.

LOSA OBSERVER SELECTION AND TRAINING

Perhaps the most critical factor in any observational method is the observer. Poor or indifferent observers can spoil the best of field methodology and instrumentation. Therefore, much effort is spent on selecting and training the right corps of LOSA observers.

Observer Selection

Previous research suggests the best and most reliable observers are unobtrusive (Kazdin, 1982), analytic (Yarrow & Waxler, 1979), motivated (Dancer et al., 1978) and trustworthy (J. M. Johnson, 1975). These important personal characteristics, as well as team composition, size, and method of selection, are all factors to consider in planning a LOSA. The first consideration is the composition of the LOSA observer team.

Observer Team Composition

Usually, regular line pilots make up most of the observation team. All pilot ranks are represented, from Captains to lower ranked pilots, such as First Officers or Flight Engineers. While such varied representation has minimal impact on data quality (since all observers regardless of rank are standardized), it can nevertheless have an important effect on the perceived credibility of LOSA results. This perception also extends to other potential LOSA observers, such as trusted line checkers, training instructors and safety personnel. Participation in LOSA observation by varied pilot ranks and airline departments provides everyone with a sense of ownership, which arguably increases organizational acceptance of LOSA.

Observers not affiliated with the airline can also be included on the LOSA observer team. Such observers are typically retired pilots or pilots from other airlines who are experienced LOSA observers. External observers serve as a control group for the rest of the observation team. For instance, observational differences between external and airline observers could signal potential measurement issues. Perhaps pilots were less reactive to external observers, which allowed these observers to detect more variation in performance than airline observers.[18] Depending on the qualifications and experience of external observers, they can provide a valuable outsider perspective of operations.

---

[18] Another data quality check for this issue is quantitative comparison, to see if airline observers are recording as many threats, errors, or undesired aircraft states as the externals.

Observer Team Size

The number of LOSA observers can vary depending on scope, resources, and size of the airline. However, team size is often determined by the total number of LOSA observations to be collected and the maximum number of observations each observer is allowed to collect.[19] From previous LOSA experience, a general rule of thumb is to limit each observer to no more than 15 observations. Since it is not uncommon for an observer to spend three to four hours writing up one LOSA observation, the 15 observation limit lessens observer burnout, which could affect data quality.

Observer Selection Method

The process for LOSA observer selection involves airline management and the pilot association. It entails tasking airline managers and pilot association representatives to create separate lists of people they would like to have as LOSA observers. Once the lists are combined, matching names automatically qualify as candidates to be observers. Unmatched names are denied consideration. Keeping in mind the needed makeup for an observer team (a majority of regular line pilots, representation of all pilot ranks, and so on), an elected representative approaches and asks the matched candidates to be part of the observer team. After a LOSA observer team is selected and everyone agrees with the team's membership, training in the LOSA methodology begins.

Observer Training

LOSA observer training is a two-part approach. The first part involves educating observers in procedural protocol and the second involves teaching TEM concepts and

---

[19] The determination of sample size is discussed in the next chapter in the section titled, "A Demonstration of LOSA Data Analysis."

classifications. Orienting observers to procedural protocol is an often overlooked but important first step to make sure observations are conducted in a consistent fashion (Reid, 1982). This orientation includes presenting the thinking behind LOSA operating characteristics, elaborating on anonymity and confidentiality, and discussing ways to brief flight crews and gain voluntary consent. It also includes training on observer etiquette, such as discussing when to speak up regarding a safety event that the flight crew fails to detect. After presenting LOSA protocol, teaching observers how to *recognize*, *record* and *code* TEM performance is the basis for the rest of the training course.

TEM Recognition

Educating observers on ways to recognize threats and errors in normal flight operations first involves presenting the theoretical rationale behind the TEM framework (Reid, 1982). Training observers is more efficient when they not only know what to look for but understand the rationale behind their observations (Hawkins, 1982). TEM recognition training involves a presentation and discussion of definitions with examples. On competition, observer training shifts to TEM *recording*.

TEM Recording

When inside the cockpit, LOSA observers only have a notepad to record their observations. No laptop computers or other electronic devices are allowed since they can create an unnecessary distraction to the flight crew. For example, imagine the degree of observation reactivity if an observer is seen or heard punching keys on a computer or tapping on a personal digital assistant (PDA) throughout a flight. It is much more discreet having observers write their comments in a notepad. Therefore, TEM recording training is spent discussing techniques on how to collect observations inconspicuously in the cockpit (for example, delaying the recording of an event to some time after its happening or showing ways to shorthand notes).

Perhaps the most important element of TEM recording training is getting observers to focus on capturing the observation first and worrying about event classification and coding, when time is less of an issue. Such post-observation coding lowers the likelihood of measurement error, which improves overall data reliability (Riesman & Watson, 1964). Therefore, while inside the cockpit, LOSA observers should only concern themselves with recording enough information to write a “story” of the flight based on TEM performance. Outside the cockpit, observers enter their narratives and code TEM events in the LOSA Data Collection Tool.

TEM Coding

Training in TEM coding first involves a presentation and discussion of the event types, responses, and outcomes. After the instructor addresses questions, observers practice identifying and coding TEM events with text exercises. Table 3.7 shows an example of an exercise with its solution.

Table 3.7 – Sample TEM Training Exercise

EXERCISE: Record / code all threats, errors and undesired aircraft states and their management.
Exercise #1 – While taxiing to the gate on the Bravo taxiway, the crew was instructed to take a right on the Hotel taxiway to the gate. As they approached Hotel, the CA, who was looking down at the airport chart, started to turn left instead of right. The FO spoke up and the Captain made the correct turn.
SOLUTION: One error, no threats or undesired aircraft states
1. Error Description: Captain started to make an incorrect turn down a taxiway. *2. Error Management Description: The Captain wasn't paying attention while taxiing and the First Officer detected the error.* *3. Flight Phase: Taxi-in* *4. Error Altitude: Not applicable* *5. Proficiency-based Error: No* *6. Threat Linkage: No* *7: Error Type: Ground Navigation* *8: Error Code: Attempting to turn down wrong taxiway* *9. Who committed the error? Captain* *10: Error Response: Detected and Action* *11. Who detected the error? First Officer* *12. Error Outcome: Inconsequential*

The observer training entails observers completing text exercises, such as the one in Table 3.7, individually in the LOSA Data Collection Tool. When observers complete the exercise, the instructor leads a class discussion on the correct number and coding of threats, errors, and undesired aircraft states, and answers any observer questions. After several exercises, observers become more consistent and accurate in their coding. Finally, at the end of the training, a knowledge test is given to measure their understanding of TEM concepts. More discussion of this test, with sample airline results, is presented later in this chapter in the section on data reliability.

A Note on Video versus Text Exercises in TEM Training

Some readers might question the use of text exercises over flight crew performance videos in observer training. Doesn't video provide a more realistic simulation of what it takes to do a LOSA observation? The answer is yes and no. While videos are arguably the standard in estimating observer reliability, they are not as practical as one would think in training TEM data collection. For example, during a regular observation, observers get to experience the entire flight and at any time can look at different aircraft instruments, switch settings, automation inputs, or listen to radio communications. With video, this is not possible. Either the camera angle prevents student observers from seeing something specific, or the instrumentation readings are too small to read. While some newer cockpit videos have enlarged the primary flight instruments on the bottom of the screen, it still inhibits the observer's ability to see other gauges not included, such as engine readouts or navigational raw data. Therefore, instead of trying to simulate a LOSA observation, training stresses the knowledge of TEM. As long as observers understand TEM concepts, deciding whether to record or not record is dependent on the information processing abilities of the observer and not on their know-how to recognize and code events.

Observer Training Summary

LOSA observer training takes place across five days. The first two days involve classroom training in LOSA protocol and TEM data collection. For the next two days, observers conduct one or two observations. The final day of training is an individual calibration check with the training instructor that provides observers with an opportunity to ask questions and receive feedback. After the instructor is satisfied with the calibration session, observers are certified to continue data collection. Observers are also informed at this time that instructors will monitor their observations throughout the data collection period. Researchers recommend this as a strategy to lessen the effects of observer drift in coding accuracy (Reid, 1970; Romanczyk, Kent, Diament, & O'Leary, 1973).

LOSA OBSERVATION PROCEDURES

This section provides an outline of procedures followed in conducting a LOSA observation. The process starts with meeting the flight crew either in crew dispatch or at the gate no later than 30 minutes before the flight is scheduled to leave.[20] Once the observer sets up contact, he introduces himself as a LOSA observer and briefs the pilots as follows:

- Explains the rationale of LOSA as a safety data gathering project, not individual evaluations of performance. Shows the signed letter of agreement between management and the pilot association if necessary.
- Discusses how the observational data will be kept anonymous and confidential.
- Asks whether any of the crew members have been previously observed during the LOSA. The procedure is to cancel an observation if any one of the crew members has been observed two times for larger airlines and three times for smaller airlines. This limit eases pilot worries of being targeted and insures a representative coverage of pilots at the airline.
- Asks for permission to conduct a LOSA observation. At this point, the observer steps aside for the crew members to discuss in private. Previous experience has often shown this to be unnecessary as many flight crews will often grant

---

[20] Crew dispatch is a location, usually at an airport, where pilots go to pick up flight documentation, such as route planning, weather, and alerts on hazards within the aviation infrastructure (e.g., an out-of-service navigation aid).

> permission on the spot. However, if any pilot or flight crew denies the observation for any reason, the LOSA observer walks away with no questions asked. If permission is granted, the observer takes his or her place in the jump seat and begins conducting the observation.

As for pilot-observer interaction, LOSA protocol calls for a cooperative approach that allows observers to engage in conversation when approached and to answer questions when asked. This in turn creates a more relaxed, non-evaluative atmosphere that reduces reactivity and improves data quality (Douglas, 1976; Weick, 1985). Otherwise, observers are trained to be as unobtrusive as possible, like a "fly on a wall" whose sole purpose is to buzz into the cockpit, collect safety data, and buzz out.

LOSA DATA RELIABILITY CHECKS

A major threat to the validity of observational findings is unreliable data (Cook & Campbell, 1979). This means various quality control checks are essential to improve data reliability. In LOSA, there are two primary checks. The first check, which occurs after observer training, assesses student accuracy in applying TEM concepts. The second happens after data collection with an external verification of observer recording and coding of TEM events.

Observer Accuracy in Applying TEM Concepts

As previously discussed, the primary objective of observer training is to teach students to recognize, record, and code TEM events. This emphasis is supported by the principles of signal detection theory (Lord, 1985; Stanislaw & Todorov, 1999). In LOSA, observers must be able to recognize threats, errors, and undesired aircraft states from other bits of information because it is these signals that prompt data recording.[21] If observers fail to notice TEM signals, no data collection takes place. Therefore, to be an effective data recorder in LOSA, it is crucial for observers to have a good working knowledge of TEM concepts. However, having this knowledge is only a precondition. It is the observer's application of this knowledge to identify and code TEM events *accurately* that improves LOSA data reliability.

Assessing Observer Accuracy: LOSA Observer Feedback Form (LOFF)

Observer accuracy in applying TEM knowledge is measured at the end of training with a test called the LOSA Observer Feedback Form (LOFF). It has 22 items across two coding exercises. The first exercise provides one-line descriptions of TEM events and asks observers to decide whether each is a threat, error, undesired aircraft state, or nothing. The second exercise asks observers to read a descent/approach/land flight narrative and record all possible TEM events. After everyone finishes both exercises, the instructor scores the tests and goes over the answers with each observer individually to clarify areas of misunderstanding. The higher the total score on LOFF, the more accurate observers are gauged to be in recognizing and coding TEM events. A full version of the LOFF with answers is in Appendix E.

LOFF: Test Reliability and Airline Results

To verify LOFF as a trustworthy measure of observer accuracy, its test reliability must be estimated. Perhaps the most commonly used index to estimate test reliability is

---

[21] This does not mean that observers will capture every TEM signal that occurs on a flight. The reality is they will not. Observers are humans and just like other humans, they are limited in their ability to process information. This natural limitation of humans results in LOSA findings being underestimates of TEM events that actually occur during a flight.

internal consistency (John & Benet-Martinez, 2000). Internal consistency assesses the extent to which all items on a test measure a single construct, or in this case, TEM concepts (Gilner & Morgan, 2000). Estimating internal consistency is a necessary first step before LOFF results can be interpreted as a reliable indicator of observer accuracy.

LOFF Internal Consistency – The following analysis is based on the most recent version of LOFF completed by 116 observers across six airlines in the last two years. These data provide the basis for the following internal consistency estimates. Since all items on the LOFF are dichotomously scored right or wrong, the Kuder-Richardson 20 (KR-20) coefficient is the most suitable estimator of internal consistency (Kuder & Richardson, 1937). The KR-20 ranges from 0 to 1 and is the mean of all possible split-half correlation coefficients. Much like the Cronbach Alpha, which estimates internal consistency for response-based items, such as a Likert scale, the KR-20 represents the lower limit of test reliability. The KR-20 for LOFF was .70.

Internal consistency was also estimated by dividing LOFF into two halves by its odd and even numbered items across both exercises. The Spearman-Brown split-half estimate of equal length was used for this estimation, which yielded a coefficient of .88. Both the KR-20 and Spearman Brown results suggest that LOFF is a reasonably homogeneous measure of TEM concepts. However, there is still some room for improvement with future research, such as lengthening the test, increasing item difficulty or estimating test-retest reliability after training and again once data collection is completed to test observer retention of TEM knowledge.

LOFF Airline Results – With LOFF showing good estimates of internal consistency, this section turns to presenting some sample observer accuracy scores.

Table 3.8 shows LOFF data from six airlines based in different geographic regions. With a perfect LOFF score being 22, the table shows the average score and standard deviation for the observer group at each airline and the total sample of 116 observers. The bottom of the table presents internal consistency estimates along with the standard error of measurement for LOFF.

Table 3.8 – Sample LOFF Results (Six Airlines)

Airline	Geographic Region	Observer Count	Observer Accuracy	
			Average Score (Perfect Score = 22)	Standard Deviation
1	North America	28	19.75	1.38
2	North America	20	19.25	2.34
3	North America	13	19.00	3.39
4	North America	13	19.15	2.08
5	Asia Pacific	27	17.89	3.30
6	Asia Pacific	15	20.00	1.25
TOTAL		116	19.11	2.49

LOFF Test Reliability Results
KR-20 = .70
Spearman Split-half coefficient (odd/even and equal length) = .88
Standard Error of Measurement = 1.36 (based on KR-20 coefficient)

Observer accuracy scores are consistently high across all airlines with each observer group averaging above 19 of 22 items, with the exception of Airline Five. The lower average for Airline Five was somewhat expected because many observers were non-native speakers of English. However, Airline Six, the other Asia-Pacific airline, had the highest average LOFF score and smallest degree of variability, suggesting that TEM concepts transfer to airlines in other geographic regions.

A chi-square analysis was used to test the consistency of observer TEM training across different airlines. Observers at each of the six airlines were divided into two groups–those with a perfect score or 1 item missed on the LOFF and those who missed 2 or more items. The resulting 2x6 chi-square analysis showed no significant differences, $\chi^2(5, N = 116) = 9.23$, $p = .10$, that is, the proportion of observers scoring high on the LOFF did not vary across airlines.[22] Combined with the uniformly high scores noted in Table 2.8, these results suggest that TEM concepts can be successfully trained and learned by observers independent of airline affiliation or geographic region.

This demonstration of successfully trained observer accuracy is only a starting point in improving LOSA data reliability. Even the best observers can sometimes make mistakes, some of which can be attributed to failing to grasp certain concepts, while others can be simple errors of commission or omission during data entry. Therefore, observer reliability estimates calculated during training do not always generalize to the data collected, which calls for extra data quality checks (Hartman, 1982).

External Data Verification

External data verification is another major check in improving LOSA data reliability. It occurs in two sequential stages. The first is an independent analyst review of the data. The second stage is the data verification roundtable that involves a group of managers and pilots from the airline to provide one last check of the data before it is analyzed.

---

[22] A chi-square was also completed by dividing observer LOFF scores into three groups. The first group was 0 to 1 item incorrect, the second was 2 or 3 items incorrect, and the third was observers with 4 or more items incorrect. This analysis also showed non-significant differences across the six airlines, providing further evidence that LOSA observer training is consistent, $\chi^2(10, N = 116) = 17.52$, $p = .062$.

Initial Data Analyst Verification

Initial data analyst verification is an opening review of LOSA data to reconcile narratives with logged TEM events. Sometimes observers will write about a threat, error, or undesired aircraft state in a phase of flight narrative but mistakenly code or fail to code the event in the software. This is not uncommon since most of the airline observers are novices. Therefore, the LOSA data analyst reviews the narratives, follows up with observers on any missing or confusing observations when necessary, and completes the missing coding. However, before any changes are made to the database, they are confirmed at the data verification roundtables.

Data Verification Roundtables

Data verification roundtables happen at the end of LOSA data collection. Their primary objective is to examine each threat, error, and undesired aircraft state, and how each was managed. Coding consistency and accuracy are also verified for each event. For example, an observer might record an error, such as not cross-verifying an automation entry, for which is there is no written procedure in the airline flight operations manual. In this case, the logged error would be removed from the database. Another example is an observer who miscodes an adverse weather threat as an airport threat. This would be corrected and retained in the database as a threat only on unanimous approval of those attending the roundtable. In those rare instances where unanimous agreement is not possible, the event is automatically deleted from the LOSA database. This is also the case if roundtable members cannot verify the accuracy of the coding with narratives or text descriptions. While this process of verification might be burdensome and time-consuming, the payoff improves the “trustworthiness” of LOSA findings (Lincoln & Guba, 1985).

Data verification roundtables typically involve four to five pilots who represent each fleet of the airline. Fleet-specific pilots are best at providing their expertise on technical and procedural issues specific to certain types of airplanes. For example, the operating philosophy and procedures for flying an Airbus airplane are different from those for flying a Boeing airplane. Therefore, fleet representation ensures these differences are properly understood and handled. Also included at every roundtable is a pilot association representative, who is there to make certain all data are confidential and free of information that could potentially identify a pilot. Once roundtable members have signed off on the dataset as being an accurate record of TEM performance, data analysis can begin. This step has the added benefit of building ownership in the results and dispelling any later criticism that the coding was not an accurate representation of the airline's operations.

CHAPTER SUMMARY

LOSA uses systematic cockpit observation of flight crews to study TEM performance in normal flight operations. The primary advantage of LOSA is that it allows one to examine performance as it naturally occurs. However, there are several measurement concerns that need to be addressed before implementation. Perhaps the biggest of the concerns is observation reactivity. This is a particular problem in commercial airlines because an observer in the cockpit usually means a performance evaluation under jeopardy conditions. While such evaluations are needed to assess flight crew proficiency, there is also a strong likelihood of altered performance to impress the evaluator. This is not the case with LOSA. It is a non-jeopardy method that aims to capture performance as close to operational reality as possible.

LOSA is built on ten operating characteristics that address pilot trust, observation reactivity, data reliability, and validity concerns. The characteristics are:

1. Jump seat observations of regularly scheduled flights
2. Voluntary flight crew participation
3. Anonymous, confidential, and non-punitive data collection
4. Joint management/union sponsorship
5. Secure data collection repository
6. Trusted and trained observers
7. Systematic observation instrument
8. Data verification roundtables
9. Data-derived targets for enhancement
10. Feedback of results to line pilots

The LOSA operating characteristics serve as an important reminder that even the best designed data collection protocols and instruments are rendered useless without participant trust. On the flip side, pilot trust is inconsequential if LOSA does not yield reliable and accurate data. Both pilot trust and trust in data quality are needed to ensure LOSA success.

Chapter Four: LOSA Data Analysis and Interpretation

After data collection and verification, the next phase of LOSA begins with data analysis. Since LOSA can create large volumes of data, choosing a starting point for analysis can sometimes be difficult.[23] This is where the Threat and Error Management (TEM) framework plays an important role. The equal emphasis the framework places on the existence and management of threats, errors, and undesired aircraft states enables analysts to mine data with a sense of direction. Based on TEM, this chapter presents a three-stage approach to LOSA data analysis and interpretation. Also presented is a demonstration of this approach using actual LOSA data.

A THREE-STAGE APPROACH TO LOSA DATA ANALYSIS

Data from flight recorders, training evaluations, incidents and accidents are only a few sources that airlines can tap into to gain knowledge about their safety performance. With any data, there must be a transition from raw data into coherent pieces of information that airline managers can act on. LOSA is no different. A database of LOSA observations only holds the potential to uncover factors that positively or negatively shape airline safety. Unless there is analysis and interpretation, these factors will go undiscovered, which defeats the purpose of data collection.

Organizational research shows that many organizations are *overdeveloped* in data collection and *underdeveloped* in their "capability to aggregate, analyze, and use data to make informed decisions that lead to action" (Davenport, Harris, De Long, & Jacobson, 2001, p. 117). Since airlines are inundated with operational performance data, this

---

[23] See the "LOSA Data Collection Tool Measures" in Chapter Three for a complete listing of LOSA narratives, text descriptions, and quantitative variables.

finding made it important for LOSA methodology to include an analytical strategy. The result is a three-stage approach that guides analysts from LOSA raw data to safety targets that airlines can act on to drive safety change. The three stages are:

1. LOSA Indices and Organizational Profiles
2. Drill-Down Analyses
3. Targets for Enhancement

These three stages represent one of several ways to structure LOSA data analysis. Regardless of approach, the purpose of analysis remains constant: to transform raw data into information that is diagnostic of systemic and flight crew performance safety issues.

STAGE ONE – LOSA INDICES AND ORGANIZATIONAL PROFILES

The LOSA indices and organizational profiles represent the output from the first stage of data analysis. This stage is critical in sending warning signals to an airline of system performance issues that merit further investigation. The first step for this effort is the calculation of threat and error management (TEM) performance indicators called LOSA indices.

LOSA Indices

The intent of the LOSA indices is to provide general indicators of system and flight crew performance. Their rationale is similar to other indices used in the business world designed to provide snapshots of large amounts of data. A good example is the Standard & Poor’s 500 Index (S&P 500). The S&P 500 measures the average performance of the 500 most actively traded stocks on the New York Stock Exchange. The number it generates is not intended to represent every facet of the stock market but to provide investors with a rough sense of its overall performance. The same is true of the LOSA indices. Their purpose is to offer airline managers a set of data points that summarize flight crew TEM performance strengths and weaknesses. The information they provide does not communicate anything about causes. Instead, the data act as pointers to vulnerabilities in organizational safety defenses (for example, latent conditions known as the "holes" in Reason's Swiss cheese model).

LOSA Indices Defined: Prevalence and Mismanagement

There are two types of LOSA indices. The first type is known as the prevalence indices. These provide information on the most frequent threats, errors, and undesired aircraft states in an airline's flight operation. The other type, called mismanagement indices, provides results on the most frequently mishandled threats, errors, and undesired aircraft states. This measurement is perhaps the more diagnostic of the two indices because it shows where flight crews, who are often the last line of defense for an airline, are having the most difficulty maintaining safety margins. Table 4.1 presents definitions and examples of the LOSA prevalence and mismanagement indices for threats, errors, and undesired aircraft states.

Table 4.1 – LOSA Prevalence and Mismanagement Indices

Prevalence Indices		
Index	Definition	Example
Threat Prevalence	Percent of LOSA observations with one or more threats	*45% of flights experienced an aircraft malfunction threat*
Error Prevalence	Percent of LOSA observations with one or more errors	*40% of flights had an automation error*
Undesired Aircraft State Prevalence	Percent of LOSA observations with one or more undesired aircraft states	*10% of flights had an unstable approach*
Mismanagement Indices		
Index	Definition	Example
Threat Mismanagement	Percent of threats that were linked to a flight crew error or undesired aircraft state	*Of the ATC threats observed, 20% were linked to flight crew errors*
Error Mismanagement	Percent of errors that were linked to an additional error or undesired aircraft state	*Of the checklist errors observed, 5% were linked to undesired aircraft states*
Undesired Aircraft State Mismanagement	Percent of undesired aircraft states that were linked to an additional error or undesired aircraft state	*Of the vertical deviation undesired aircraft states observed, 2% led to additional error*

The LOSA prevalence and mismanagement indices in Table 4.1 are aggregated for all threats, errors, and undesired aircraft states observed during a LOSA. While this information is useful to gain an overall sense of TEM performance at an airline, they are limited in diagnosticity. To solve this issue, the prevalence and mismanagement indices are also calculated for each sub-category of threat, error, and undesired aircraft state (see Table 4.2).[24] For example, the prevalence index showing 75% of flights having a

---

[24] The threat, error, and undesired aircraft state sub-categories were first presented in the “Threat and Error Management Measures” section of Chapter Three.

checklist error could indicate a specific TEM weakness. On the other hand, a low mismanagement index of 1% for airport threats shows an area of TEM performance strength. Information such as this allows for greater precision in identifying those issues that an airline needs to focus on for improvement versus those not needing any organizational action at all.

Table 4.2 – Threat, Error, and Undesired Aircraft State Types

Threats	Errors	Undesired Aircraft States
Environmental Adverse weather Airport ATC Environmental operational pressure Airline Aircraft malfunction Airline operational pressure Cabin Dispatch/paperwork Ground/ramp Ground maintenance Manuals/charts	Aircraft Handling Automation Flight controls Ground navigation Manual handling Systems/Radio/Instruments Procedural Briefings Callout Checklist Documentation PF/PNF duty SOP cross-verification Other procedural Communication Crew to External Pilot to Pilot	Aircraft Handling Landing deviations Lateral deviations Speed deviations Unstable approach Vertical deviations Ground Navigation Navigation deviations Speed deviations Incorrect Aircraft Configuration Automation Engine Flight control Systems Weight and balance

LOSA Organizational Profiles

The second half of stage one focuses on the presentation of index results to airline management. Organizational researchers have identified an increasing problem of information overload, which is characterized as the amount of information available exceeding an organization's ability to process it (see Edmunds & Morris, 2000 for a review of this literature). The people most affected by information overload are managers who receive more performance data than anyone else in an organization (Katzer & Fletcher, 1996). As waves of information come in, managers tend to act on bits that are rapidly accessible due to their presentation, and structured in a way that plainly points to a need for organizational change (Koniger & Janowitz, 1995; Simpson & Prusak, 1995). Therefore, the more difficult it becomes for airline managers to process LOSA results, the less likely the results affect an organizational change. It was this concern that triggered the development of LOSA organizational profiles.

The LOSA organizational profiles present the prevalence and mismanagement indices together in one place for threats, errors, and undesired aircraft states. Take for example the error organizational profile of a major airline in North America (Table 4.3). All the major error results are presented in one page so that managers can perform a quick assessment and decide which areas need more focus.

Table 4.3 – Sample Error Organizational Profile

Error Type	Error Prevalence Index	Error Mismanagement Index	Mismanaged Error / Error Count
All Errors	76% of observations	25% of errors	294/1159
Aircraft Handling Errors			
System/Instrument/Radio	31%	43%	62/144
Automation	32%	38%	51/134
Manual Handling/Flight Control*	27%	72%	73/102
Ground Navigation	7%	81%	17/21
Procedural Errors			
Checklist	33%	15%	27/182
SOP Cross-verification	39%	7%	10/144
Other Procedural	24%	32%	28/88
Callout	22%	1%	1/78
Briefing	18%	12%	8/67
Documentation	10%	25%	9/36
PF/PNF Duty	9%	6%	2/31
Communication Errors			
Crew to External	29%	5%	6/123
Pilot to Pilot	3%	0%	0/9

Notes:

N=308 LOSA observations

* Manual flying and flight control error categories are often combined since they both involve the physical manipulation of the aircraft.

In summary, the LOSA indices provide data-driven indicators of safety performance strengths and weaknesses in normal operations. For example, some readers might have noticed in Table 4.3 the high rate of mismanaged ground navigation and aircraft handling/flight control errors. Results such as these could be taken as warning signs of systemic weaknesses that need to be addressed by the airline. However, this information alone is not enough. Further description of the contributing factors that led to these high rates of error mismanagement is needed. Such explanation is achieved with the next stage of LOSA data analysis called drill-down analyses.

STAGE TWO – DRILL-DOWN ANALYSES

The term "drill-down" describes the process of moving from the prevalence and mismanagement indices to deeper layers of LOSA data. Drill-down analyses provide the explanation that underlies index results. This stage of LOSA data analysis consists of four primary types of drill-downs that provide airlines with unique pieces of information. They are: (1) event description, (2) demographic, (3) TEM process, and (4) operational context drill-downs.

Event Description Drill-Downs

As mentioned in Chapter Three, each of the threat, error, and undesired aircraft state types listed in Table 3.3 has another layer of categorical coding that describes the event in more detail. For example, the callout error type has several subtype codes, such as omitted altitude callout or nonstandard approach callouts.[25] Such drill-downs are possible with any of the LOSA prevalence and mismanagement indices. All sub-categories of threats, errors, and undesired aircraft states are in Appendixes B, C, and D.

---

[25] An example of nonstandard callout is a pilot calling out "let the wheels dangle" instead of saying the procedurally mandated "gear down."

Demographic Drill-Downs

LOSA collects various demographic measures that identify particular characteristics of an observation. They include city pairs (departure and arrival city), aircraft type, pilot flying for takeoff and landing, and flight time. Similar to the event description drill-downs, analysts can check for demographic differences based on the signals provided by the prevalence and mismanagement indices.

An example of a demographic drill-down analysis is researching pilot flying effects underlying error prevalence indices. For instance, let's say 30% of an airline's LOSA observations had "high speed" undesired aircraft states. On drilling down into demographic data, the results show that First Officers were pilot flying in 85% of these states. The more detailed nature of this result allows airlines to become more specific in their response. Instead of simply focusing on reducing "high speed" undesired aircraft states, an airline can tailor their safety fix to First Officers and speed management issues during training.

TEM Process Drill-Downs

The third drill-down analysis in LOSA focuses on understanding flight crew processes involved in managing threats, errors, or undesired aircraft states. This activity is captured with a measure called flight crew response. For threats, the flight crew response is whether a threat was discussed or planned for by the pilots before it was encountered. For errors and undesired aircraft states, observers code flight crew responses as "detected and action" or "failing to respond." Understanding these processes can dictate different organizational responses, such as designing safety fixes to improve flight crew detection.

An example of a TEM process drill-down might begin with a high mismanagement index for automation errors. Drilling down using the flight crew response measure, an analyst might find that 90% of the mismanaged errors were undetected or ignored by the flight crew (that is, failing to respond). In response to the finding, an airline might try to find ways to not only lower the prevalence of automation errors but increase their likelihood of being detected with better pilot awareness, training, or standard operating procedures.

Operational Context Drill-Downs

A distinct advantage of LOSA is that all data collection occurs within an airline's operational context. While the categorical coding of threats, errors, and undesired aircraft states captures pieces of this context, they are nevertheless limited. They can sometimes leave gaps in understanding that are critical in explaining findings. To fill these gaps and gain insight on the context surrounding TEM performance, airlines must drill-down to the observer narratives and text descriptions, which are considered the deepest layer of LOSA data.

The following example of an operational context drill-down occurred at a LOSA airline that found a high prevalence of checklist errors occurring during the taxi-out phase of flight. On drilling down into the narrative data, the airline found that most of these errors occurred at airports with short taxi-outs to the runway. In these instances, flight crews would rush and try to save time by hurrying through the taxi and before-takeoff checklists. The problem was the pilots would sometimes accidentally skip items leaving previously committed system and flight control errors in predeparture undetected.

In response to these results, the airline decided the taxi and before takeoff checklists were accidents waiting to happen because there was not enough time to properly complete them during short taxis. Their solution was to move items from the taxi checklist to predeparture and shorten the before takeoff checklist to a few critical items such as checking flaps. This would then give crews enough time to perform the before takeoff checklist and improve their detection of errors committed in predeparture. At the time of writing this chapter, the airline is planning a follow-up LOSA that should provide data that indicate whether their safety solutions were effective. Without LOSA, one is left to wonder whether this issue would have been identified before an incident or accident. Fortunately, this airline was proactive and did not wait to find out.

STAGE THREE – TARGETS FOR ENHANCEMENT

The final stage in LOSA data analysis is developing data-derived targets for enhancement that steer an airline's safety change process.[26] There is no limit to the number or specificity of these targets. An example target might focus on reducing threats in the predeparture/taxi-out while another could be more specific, such as wanting to achieve a 50% improvement in the management of ground navigation errors. Whatever the targets are, they are essential in marking a clear end-point for the LOSA project and the beginning of an airline's responding to the results with safety change. The next section of this chapter guides readers through all stages of data analysis using observations collected from a major commercial airline.

A DEMONSTRATION OF LOSA DATA ANALYSIS

Perhaps nothing has been more satisfying in LOSA development than taking the methodology from a collection of ideas to a practical application used by commercial

---

[26] The generation of data-derived targets for enhancement is one of the ten LOSA Operating Characteristics discussed in Chapter Three.

airlines. This section uses actual LOSA observations collected from an airline to guide readers through the three stages of LOSA data analysis. The purpose of this demonstration is to provide readers with a sense of the tactics and decisions involved in making the transition from LOSA data to knowledge that airlines can use to improve safety.

The airline chosen for the demonstration is a major carrier with five fleets that fly domestic and international routes all over the world. For confidentiality reasons, airline name, fleet types, number of departures flown and routes are kept anonymous. For this demonstration, the airline will be referred to as Zed Airlines.

Zed Airlines performed their LOSA with adherence to the ten operating characteristics and procedural protocol outlined in Chapter Three. The result was a 309 observation sample collected using a stratified random sample that was weighted by the departures flown by each fleet. For instance, if a Boeing 737 fleet accounted for 30% of an airline's daily departures, then roughly 30% of the LOSA sample consisted of 737 observations. This sampling technique allows the organizational profiles to represent TEM performance on any given day of operations.

With background details of Zed Airline's dataset in place, let's continue with the airline's LOSA indices and their presentation in threats, error, and UAS organizational profiles. For simplicity, only the threat organizational profile will be analyzed and discussed.

Threat Indices and Organizational Profile

Table 4.4 presents the threat organizational profile for Zed Airlines. The prevalence and mismanagement indices are calculated for all threats combined as well as their subtypes. The far right column of the profile compares the number of threats observed with the number mismanaged.

Table 4.4 – Zed Airline's Threat Organizational Profile

Threat Type	Threat Prevalence Index	Threat Mismanagement Index	Mismanaged Threats / Threat Count
All Threats	100%	6%	86/1454
Environmental Threats			
ATC	61%	8%	27/342
Adverse Weather	61%	6%	18/304
Environmental Operational Pressure	51%	2%	4/228
Airport	13%	5%	2/42
Airline Threats			
Aircraft Malfunction	52%	9%	27/290
Cabin	18%	5%	2/65
Airline Operational Pressure	18%	5%	3/60
Dispatch/Paperwork	12%	2%	1/44
Ground Maintenance	8%	3%	1/32
Ground/Ramp	8%	4%	1/27
Manual/Charts	7%	0%	0/20

Definitions:

Threat Prevalence Index – Percent of LOSA observations with one or more threat.

Threat Mismanagement Index – Percent of threats that were linked to a flight crew error or undesired aircraft state.

Notes:

N = 309 observations (Zed Airlines)

The purpose of the threat organizational profile in Table 4.4 is to highlight operational strengths and weaknesses associated with threats and their management at Zed Airlines. One way to distinguish systemic performance strengths from weaknesses is to examine the prevalence and mismanagement indices together. For instance, those threats with low prevalence and low mismanagement index results, such as airport, dispatch/paperwork, ground/ramp, ground maintenance, and manual/chart threats, could signify safety performance strengths. On the other hand, threats, such as ATC, adverse weather, and aircraft malfunctions, with high prevalence and high mismanagement indices, suggest safety performance weaknesses and possible holes in systemic defenses. Those threats with high prevalence/low mismanagement or low prevalence/high mismanagement could also signal safety performance weaknesses. For the sake of simplicity though, let's limit the discussion to those threats with high prevalence/high mismanagement rates, because they undoubtedly show areas of concern for Zed Airlines.

ATC, adverse weather, and aircraft malfunction threats occurred on over half of the flights observed and accounted for 84% of all mismanaged threats at Zed Airlines (72 of 86 mismanaged threats observed). Perhaps the most interesting of these three threats are aircraft malfunctions. Unlike the environmental threats of ATC and adverse weather, airline decision makers have direct influence on the *prevalence* and *management* of aircraft malfunctions because it is their implemented policies and practices that largely determine organizational maintenance standards. With this in mind, let's examine these threats in more detail using drill-down analyses.

Drill-Down Analysis (Event Description): Aircraft Malfunction Threats

The first drill-down analysis usually performed on finding a threat, error, or UAS with high prevalence and mismanagement indices is an event description drill-down.

Table 4.5 presents the results of this drill-down by showing the different kinds of aircraft malfunction threats that Zed Airlines' flight crews face and have to manage.

Table 4.5 – Event Description Drill-Down Analysis: Aircraft Malfunction Threats

Aircraft Malfunction Threats	Threat Prevalence Index	Threat Mismanagement Index	Mismanaged Threats / Threat Count
MEL with operational implications	36%	8%	14/178
Malfunction unexpected by the flight crew	20%	13%	12/96
Automation event/anomaly	5%	6%	1/16
Total	52%	9%	27/290

Judging from the results in Table 4.5, Minimum Equipment List (MEL) threats with operational implications were the most frequent aircraft malfunction threats observed at Zed Airlines (36% of flights). An MEL is a type of malfunction that airline maintenance can legally label as needing repair but is not considered serious enough to keep an airplane from being airworthy. For instance, an erratic thrust reverser, which is a piece of equipment often used to aid braking on landing, is an MEL with operational implications. In response, maintenance workers would tag the reverser as being "INOPERATIVE," preventing its use by pilots during landing. The mismanagement index for MEL threats was 8% (14 of 178 threats observed). These threats were mostly linked to minor procedural errors of not discussing the MEL during briefings.

Potentially the most serious of aircraft malfunction threats are those unidentified by airline maintenance and occurring during flight. These are known as aircraft malfunction threats unexpected by flight crew. They occurred on 20% of the observations and were mismanaged at a rate of 13% (12 of 96 threats observed). After reviewing the observer narratives and error data linked to these threats, it became clear that most involved failing to use the Quick Reference Handbook (QRH), which is a checklist that gives step–by–step guidance to resolve certain types of malfunctions. In those instances where crews failed to use the QRH, pilots were found to carry out their own actions or no actions at all. Table 4.6 and Table 4.7 provide qualitative data of one such instance when a flight crew failed to use the QRH and continued the flight with an unresolved Auxiliary Power Unit (APU) malfunction.[27]

Table 4.6 – Sample Narrative Data

LOSA Observation #21 **Fleet Three** **Pilot Flying: First Officer**

Predeparture/Taxi–Out Narrative

It was a very early morning departure, i.e., even the observer's pick-up was at 3:30 AM local time. All respective duties were done with everyone 'working–in–sync'. While the Flight Engineer (FE) was still busy with his work/scan, the Captain (CA) offered to make a round of drinks, which was accepted by First Officer (FO). This showed good team building by the CA, which was observed throughout the flight with good rapport among crew.

The ground crew completed the pushback but did not call for the CA to set the parking brakes. After some time, the CA asked the ground crew if he wanted the brakes on, which was quickly acknowledged with an affirmative answer. All checklists were read with the correct procedural protocol with everyone verifying and cross-checking each switch position.

On taxi out, the FE noticed the Auxiliary Power Unit (APU) Fault light was ON (Aircraft malfunction threat unexpected by the crew). The FE informed the rest of the crew. The CA ordered the FE to check the ops manual and to recycle the switch. However, there was no luck as the light was still on. Consequently, this led to the crew discussing this fault and it was diagnosed as being caused by a slightly opened or not flushed APU air inlet door. Although all crew had a consensus on this, the FE was still very much bothered with it and waited for it to disappear all the way to entering the runway. No abnormal checklist was performed (flight crew error) and the light remained on throughout the flight.

---

[27] The Auxiliary Power Unit (APU) is a small turbine engine often located in the tail of an airplane that serves as an independent source of electrical power. During flight, running engines provide electricity to the airplane. Therefore, APU's are especially important in providing backup power in the event of an emergency such as an engine failure.

Table 4.7 – Sample Threat Description Data

LOSA Observation #21 Fleet Three Pilot Flying: First Officer

Threat #1

Threat Description
APU amber fault light came on during taxi-out.

Threat Management Description
FE checked the operational manual and attempted to recycle the APU switch but the light was still on. This was linked to a flight crew error of failing to run the abnormal checklist (Quick Reference Handbook) and resulted in an undesired aircraft state of operation with an unresolved MEL. Consequently, the crew discussed the fault and diagnosed it as being caused by a slightly opened or not flushed APU air inlet door. The crew took off with the light on where it remained on until landing.

Phase of Flight Preflight/Taxi

Threat Type Aircraft Malfunction **Threat Code** Aircraft malfunction unexpected by crew

Threat Outcome: Linked to Flight Crew Error (Procedural error – Failure to execute an abnormal checklist)

After reading the text data in the previous tables, some readers might wonder why the observer didn't speak up and ask the crew to review the QRH. The answer is the observer, who was a pilot at the airline, felt the Captain was correct in his diagnosis and decision to continue the flight. Nevertheless, he recorded what he observed. It was only at the data verification roundtables that the error in Table 4.7 was picked from the observer's narrative and coded. This shows the importance of data verification in order to capture the most complete snapshot of TEM performance in normal operations, and why data verification is considered one of the ten operating characteristics of LOSA.

Threat Drill-Down Analysis (Fleet Differences): Aircraft Malfunction Threats

Another possible drill-down analysis that sheds more light on aircraft malfunction threats at Zed Airlines is examining fleet differences. The information gained determines whether the high prevalence of these threats is limited to particular fleets or a systemic issue that spans across all fleets. Table 4.8 presents the indices for aircraft malfunctions for each fleet at Zed Airlines.

Table 4.8 – Fleet Drill-Down Analysis: Aircraft Malfunction Threats

Fleet	Threat Prevalence Index	Threat Mismanagement Index	Mismanaged Threats / Threat Count
1	56%	6%	3/52
2	49%	9%	7/77
3	61%	3%	2/70
4	52%	6%	2/32
5	47%	22%	13/59

The results from Table 4.8 show a difference among fleets in the prevalence of aircraft malfunction threats, especially with Fleet Three, which happens to be one of the older fleets at Zed Airlines. There is also a difference in the mismanagement indices, with Fleet Five showing the highest rate (22% mismanagement). On further drill-down analysis into the threat sub-categories (not shown), it was found that Fleet Five faced the most unexpected aircraft malfunctions which, based on previous findings, were mostly linked to omitting the QRH.

Targets for Enhancement: Aircraft Malfunction Threats

The results of Zed Airlines' threat analysis showed several operational strengths and weaknesses associated with threats and their management. Of particular importance for this demonstration were aircraft malfunction threats. Based on the findings provided by the threat organizational profile and drill-down analyses, a few examples of targets for enhancement include:

- Lower the system-wide prevalence of aircraft malfunction threats (MEL threats as well as those malfunctions unexpected by the crew). The high prevalence of these threats across all fleets might indicate a latent condition of slipping maintenance standards at Zed Airlines.
- Improve the management of aircraft malfunction threats that were unexpected by the flight crew. Drill-down analyses show many of these threats were mismanaged because flight crews failed to properly reference the QRH. Possible organizational action is to further explore the reasons behind QRH noncompliance. Possible explanations might be found at the flight crew level (i.e., sub-group of complacent pilots resulting in procedural noncompliance), or at the system level, such as poor QRH design that discourages its use. This is especially relevant for Fleet Five since they had the highest prevalence and mismanagement rates for unexpected malfunctions at Zed Airlines.

After solutions are implemented and time has passed for the operation to settle, Zed Airlines plans to re-measure with a follow-up LOSA. It is this "measure, make a change, re-measure" approach that represents the essence of LOSA and all proactive safety efforts in preventing incidents and accidents in commercial aviation.

This demonstration analysis only represents one possible stream of LOSA data analysis. The same analytic approach can be taken with other types of threats. For those interested, the error and undesired aircraft state organizational profiles for Zed Airlines are in Appendixes F and G.

LOSA DATA ANALYSIS IN RESEARCH

The discussion to this point has focused on an analytical approach to help airlines process LOSA data. With little adjustment, the same approach is possible for the research purpose of modeling threat and error management (TEM) performance in the field. Again, this is best demonstrated with data. Instead of using one airline, this section presents results from an archive of 2612 observations collected from 10 airlines since January 2000.[28] While there were several LOSA projects conducted before 2000, the methodology was constantly refined from project to project. The data from the most recent LOSA implementations can be aggregated together since they were all collected with consistent protocol and instrumentation. Table 4.9 lists the ten airlines in the archive by their geographic base and observation count.

---

[28] Each airline in the archive formally agreed to donate their LOSA data at no cost to The University of Texas for research purposes.

Table 4.9 – Geographic Base and Observation Count: Ten Airline Archive

Airline	Geographic Base	Observations Count
1	Asia Pacific	238
2	Asia Pacific	265
3	Asia Pacific	309
4	Asia Pacific	221
5	Europe	174
6	Latin America	132
7	North America	308
8	North America	372
9	North America	390
10	South Pacific	203
	TOTAL	2612

While the archive presented in Table 4.9 contains a diverse cross-section of airlines, it should not be inferred as providing a representative sample of the global aviation industry. For example, there is under-representation of LOSA carriers from Europe and Latin America. On the other hand, the archive does provide insights into TEM performance issues that are common across a large sample of airlines. The next three sections provide readers with a sense of these issues by presenting the most general threat, error, and undesired aircraft state results from the archive.

General Threat Results: Ten Airline Archive

Just as the LOSA prevalence and mismanagement indices serve as warning signals for airlines, they are also signals that could prioritize research efforts. Table 4.10 shows the average threat prevalence and mismanagement index rates with their range across the ten airline archive. Also shown is the most frequent and most frequently mismanaged threat by type in the archive.

Table 4.10 – LOSA Threat Indices: Ten Airline Archive

N=2612 observations	Threat Prevalence Index (Average)	Index Range (10 airlines)	Threat Count	Most Prevalent by Type (All Airlines Combined)
Threat Prevalence	96%	81%–100%	10444	Adverse weather (61% of flights) ATC (56%) Environmental operational pressure (36%) Aircraft malfunctions (33%) Airline operational pressure (18%)
	Threat Mismanagement Index (Average)	Index Range (10 airlines)	Mismanaged Threat Count	Most Often Mismanaged by Type (All Airlines Combined)
Threat Mismanagement	9%	6%–13%	932	ATC (12% mismanaged) Aircraft malfunctions (12%) Adverse weather (9%) Dispatch/paperwork (9%) Airline operational pressure (7%)

For example, across the ten airlines, the most prevalent threats were adverse weather, ATC, and environmental operational pressure (for example, terrain and traffic). The ATC threats are noteworthy for researchers because they also had the highest average rate of mismanagement. Another interesting result is shown by the 9% threat mismanagement rate, which is the percent of threats that directly contributed to a flight crew error. There was also little variability in threat mismanagement, shown by the small range from 6% to 13%. These results indicate that flight crews are good threat managers in an operating environment full of threats, as shown by the high range of threat prevalence [(81% to 100% of flights observed with at least one threat; average number of threats per flight = 4.00 (10444 threats / 2612 flights)].

General Error Results: Ten Airline Archive

Similar to the threat results presented above, Table 4.11 provides some interesting insights into flight crew errors and their management.

Table 4.11 – LOSA Error Indices: Ten Airline Archive

N=2612 observations	Prevalence Index Average	Index Range (10 airlines)	Error Count	Most Prevalent by Type
Error Prevalence	80%	62%–95%	7257	Automation (25% of flights) Systems/instruments/radio (24%) Checklist (23%) Manual handling (22%) Crew to external communication (22%)
	Mismanagement Index Average	Index Range (10 airlines)	Mismanaged Error Count	Most Often Mismanaged by Type (All Airlines Combined)
Error Mismanagement	27%	18%–47%	1825	Manual handling (79% mismanaged) Ground navigation (61%) Automation (37%) Systems/instruments/radio (37%) Checklist (15%)

As discussed in the literature review in Chapter Two, much of the research on pilot error has focused on error commission with little attention paid to error management processes. The results in Table 4.11 show why both deserve equal attention from researchers. The most prevalent error across the ten airlines was automation, but the errors most often mismanaged were those committed when pilots turned the automation off and were manually flying the airplane. These general findings suggest that automation is a pervasive source of flight crew error. However, the mismanagement results imply that possibly more automation is needed to support manual flying tasks. This paradox is a particular issue for aircraft designers, who want to provide the most technologically advanced airplanes without degrading pilot performance.

General Undesired Aircraft State Results: Ten Airline Archive

Table 4.12 presents the prevalence and mismanagement index results for undesired aircraft states.

Table 4.12 – LOSA Undesired Aircraft State Indices: Ten Airline Archive

N=2612 observations	Prevalence Index Average	Index Range (10 airlines)	UAS Count	Most Prevalent by Type (All Airlines Combined)
Undesired Aircraft State Prevalence	34%	24% – 51%	1347	Incorrect systems configuration (9% of flights) Incorrect automation configuration (6%) Speed deviations (high speed) (6%) Unstable approach (5%) Vertical deviations (3%)
	Mismanagement Index Average	Index Range (10 airlines)	Mismanaged UAS Count	Most Often Mismanaged by Type (All Airlines Combined)
Undesired Aircraft State Mismanagement	13%	5% – 20%	175	Unstable approach/No go-around (98% mismanaged) Incorrect automation configuration (8%) Incorrect systems configuration (8%) Incorrect flight controls configuration (8%) Lateral deviation (7%)

Since undesired aircraft states are considered on the cusp of becoming an incident or accident, they are arguably the most urgent topics for LOSA research. For example, research could assess the quality of procedural countermeasures (such as, procedurally mandated checklists or deviation callouts) and their role in undesired aircraft states. Due to the high rate of mismanagement, unstable approaches provide another avenue for future research. The primary reason for this mismanagement was not carrying out a mandated missed approach, which is also known as a go-around. Instead, most of the flight crews continued the unstable approach and landed in the runway touchdown zone without incident. Unfortunately, the Flight Safety Foundation reports that unstable approaches that should have resulted in a missed approach have been implicated in many of the aviation industry's approach and landing accidents (FSF, 1998). The TEM perspective provided by LOSA research provides a unique opportunity to study factors contributing to unstable approaches independent of flight outcome.

CHAPTER SUMMARY

This chapter presented a multistage approach to analyzing and interpreting LOSA data. With this approach, airlines are able to create indicators of TEM performance strengths and weaknesses presented in organizational profiles. These results are the "vital signs" of LOSA data analysis. Much like the vital signs recorded during a medical visit (for example, temperature, heart rate, and blood pressure), they act as warning signals of deeper health concerns. Using these signals, the second stage begins with drill-down analyses that allow airlines to more precisely explore the contributing factors behind the index results. After the results are processed and prioritized, the final stage is developing the targets for enhancement that provide airlines with empirical justification for organizational safety changes.

Chapter Five: Discussion

This chapter presents a discussion of the Line Operations Safety Audit (LOSA) with an outlook towards the future. The first half integrates the material presented in previous chapters to provide a final review of LOSA and its role in research and airline safety management. This section also includes a listing of the research tasks completed in this dissertation that form the basis of LOSA. The second half of the chapter examines the current and potential implications of LOSA for aviation safety. Finally, the last section discusses future research directions for LOSA and the concept of Threat and Error Management (TEM).

REVIEW: LOSA RATIONALE AND DISSERTATION RESEARCH TASKS COMPLETED

LOSA uses trained observers to gather TEM data from inside the cockpit during routine flights. As discussed throughout this dissertation, this has relevance for both aviation safety researchers and airline safety managers.

LOSA as a Field Research Method

From the research standpoint, LOSA is an opportunity to observe flight crew performance in its natural context. Unfortunately, as discussed in Chapter Two, such research has been rare in aviation safety. Much of the current work is either theoretically driven with little empirical evidence, simulator experiments lacking true operational context, or incident or accident studies skewed to negative pilot performance outcomes. The opportunity provided by LOSA complements this research by providing insight into the simultaneous interrelationships among contextual influences, flight crew processes, and safety outcomes. Guided by the TEM framework presented in Chapter Two, LOSA records how flight crews anticipate, detect, and recover from threats, errors, and undesired aircraft states during normal operations. With the high cost of conducting field observations, and increased security measures that have made researcher access to cockpits difficult, such knowledge has been mostly unavailable or fragmented. LOSA is able to fill this gap in research knowledge because it is a method used directly by airlines.

LOSA as an Airline Safety Data Collection Tool

From the airline perspective, LOSA is a proactive safety measure that complements existing data sources, such as line evaluations, quick access recorders, voluntary incident reports and accident investigations. As discussed in Chapter Two, all of these data sources are specific in what they capture. For example, quick access recorders collect data on flight parameters (for example, altitude and speed), but nothing on flight crew interaction in the cockpit. LOSA fills this knowledge gap with TEM performance data collected from inside the cockpit during regular operations.

The rationale behind LOSA for airlines is perhaps best explained with an analogy to an annual health exam. A health exam can signal excellent fitness or a person's risk of a life-threatening condition, such as cancer or a heart attack. The risks for these conditions are determined by a series of diagnostic measures, such as X-rays, blood pressure readings, and laboratory tests. Depending on the results of these measures, a doctor might advise treatment, changes in lifestyle, or nothing at all. LOSA is based on the same preventative notion; the diagnostic indicators represent strengths and weaknesses in TEM performance before an incident or accident. As the commercial aviation environment becomes increasingly more demanding with the congestion and capacity issues discussed in the introduction, such proactive monitoring is becoming more and more essential to maintain current levels of safety. However, monitoring normal operations is not enough. Much like a patient who gets his results from an annual health exam, an airline can decide to do nothing, or can respond to LOSA findings with corrective action.[29]

Dissertation Research Tasks Completed

The primary research objectives of this dissertation were to demonstrate the rationale, methods, instrumentation and analytic strategies for LOSA. In doing so, there were several tasks completed throughout the dissertation, mostly in Chapters Three and Four. Table 5.1 provides a listing of these tasks divided by methodology and data analysis/interpretation.

---

[29] It is important to note that some people fail to get health checkups, preferring to wait until they are sick before going to the doctor. They often consider themselves healthy until something happens that proves them otherwise. The same can be true for airlines. While a patient who opts out of a health checkup can only endanger himself, an airline that chooses not to have them can endanger many people at one time.

Table 5.1 – Dissertation Research Tasks Completed

LOSA Methodology (Chapter Three)

Protocol – Outlined ten operating characteristics that define LOSA and are designed to minimize observation reactivity.

Instrumentation – Presented various LOSA measures with emphasis on the categorical and narrative measures of TEM performance.

Observer Selection and Training – Developed guidelines to select quality observers and train them to recognize, record, and code TEM performance.

Observation Procedures – Provided step–by–step procedures for observation standardization starting from the point of gaining flight crew permission to observe and ending with the proper etiquette in conducting an observation.

Data Reliability – Discussed various data quality control checks such as observer accuracy testing, initial data analyst verification, and data verification roundtables.

LOSA Data Analysis/Interpretation (Chapter Four)

Data Analysis – Presented a three-stage approach to LOSA data analysis and interpretation.

1. LOSA Indices and Organizational Profiles – A technique that offers descriptive indicators of threat, error, and undesired aircraft state prevalence and mismanagement in normal flight operations.
2. Drill-Down Analyses – A series of analyses that provide explanation of underlying TEM performance with text event descriptions, demographic effects, flight crew processes and contextual influences.
3. Targets for Enhancement – A process for deriving targets that airlines can use to benchmark organizational safety change.

Data Analysis Demonstration – Conducted a sample data analysis using LOSA data collected from a major airline. Also presented results from an archive of over 2600 observations across ten airlines to demonstrate the research possibilities of a LOSA archive.

LOSA IMPLICATIONS FOR AVIATION SAFETY

At the time of writing this dissertation, over 20 airlines from 10 different countries have conducted a LOSA.[30] Many of these implementations notably occurred

---

[30] These only include airlines that have worked with The University of Texas and LOSA Collaborative. There are some airlines that have implemented LOSA by themselves without external assistance.

during periods of financial instability within the aviation industry. Intense cost cutting, competition, and rising jet fuel prices after the 2001 World Trade Center attacks have and continue to put a stranglehold on airline resources. The willingness of these airlines to carry out LOSA during stressed times and to do it without being mandated by government regulations is evidence of its sensed utility. However, this level of commitment would not be possible without the early airline pioneers that reported LOSA findings as driving improvements in their flight operations. Examples of these improvements include positive changes in training curriculum, simulator evaluations, standard operating procedures and on a wider scale, their safety culture (see Craig, 2003; Gunther, 2002; Tesmer, 2002).[31] These public success stories have caught the attention of other organizations and domains within the commercial flying environment, some of which are discussed next.

Joint Organizational Support for LOSA

Several key aviation organizations have lent their support to LOSA. First, the International Civil Aviation Organization (ICAO), which is the United Nations' regulatory agency for commercial aviation worldwide, made LOSA the central focus of its flight safety program between 2000 and 2004 (Maurino, 2002). They also plan to formally recommend it as an industry best practice to all member nations in the next few years for the monitoring of safety in normal flight operations. The Federal Aviation Administration (FAA) has followed the ICAO lead. An LOSA Advisory Circular that provides FAA guidance to all airline stakeholders in the United States has been drafted and is in final review. This support is noteworthy because of the influence that these

---

[31] These examples only represent the changes that airlines have publicly reported. For a more complete case study of LOSA at an airline, readers are referred to Gunther (2002) and Tesmer (2002).

organizations have in being able to communicate LOSA, TEM and a proactive safety approach to airlines of all types and sizes.

Airline pilot associations have also embraced LOSA. The Air Line Pilots Association (ALPA), which is the largest pilot union in the United States, is providing its members with policy on LOSA and its ten operating characteristics (Scott Scheiffler, personal communication, March 2, 2004).[32] The International Federation of Air Line Pilot's Association (IFALPA), which represents more than 100,000 pilots worldwide, has also shown support for LOSA (Greeves, 2002). Both of these organizations are consistent in their concerns to preserve pilot confidentiality and the non-disciplinary nature of LOSA. Their support has already proved useful in gaining joint management/association sponsorship at several airlines wanting to implement LOSA. As discussed in Chapter Three, this sponsorship is essential in easing pilot suspicions of LOSA. The less threatened pilots feel in front of an observer, the less likely they will react by "faking good" in their performance, which improves the validity of LOSA findings. Much like the regulatory support discussed above, pilot association support for LOSA encourages airlines to become more proactive in their safety management practices. The difference here is that support for LOSA is coming from the pilots themselves, which is another sign of its ability to positively impact airline safety.

Finally, LOSA has potential for aircraft manufacturers in its ability to gather data on how pilots fly their airplanes in the real world versus how designers expect them to be flown. This notion motivated an effort already underway at The Boeing Company. They are using airline de-identified LOSA/TEM data to help drive future aircraft design, procedures, and training (Tesmer, 2002). The following is an excerpt from Castano and

---

[32] Captain Scott Scheiffler is the Group Chairman, Human Factors and Training, ALPA Safety. The formal guidance on LOSA appears in the ALPA Administrative Manual, Section 80, Part 5, I. Line Operations Safety Audit.

Graeber (2002) that best summarizes their perspective of LOSA and its influence on aircraft design.

> As a manufacturer, Boeing is particularly eager to have access to a reliable source of information about normal operations. Insight into how the company's products are being operated in today's constantly changing operational environment is invaluable. While always seeking to learn lessons from incidents and accidents, it is equally, if not more important to anticipate safety issues related to human error before they lead to mishaps. LOSA data can provide such insights into both aeroplane designs and procedures. (p. 27)

In summary, regulator, pilot association, and aircraft manufacturer support of LOSA show that its implications for safety are not limited to what occurs inside an airline. LOSA influence pervades the industry because these organizations have the power to affect safety at many levels within the aviation system. In another sense, their common support for LOSA creates a chance for these diverse organizations to work together and share information more efficiently in their common goal of proactively improving safety.

Extension of LOSA to Other Aviation Domains

The LOSA methods presented in this dissertation can potentially extend to other aviation domains, such as maintenance, cabin, dispatch, and ramp operations. In fact, a feasibility study in progress is adapting LOSA to air traffic control (ATC). The project is the Normal Operations Safety Survey (NOSS) and similar to LOSA, bases its measurement on TEM (Maurino & Ruitenberg, 2004). Its purpose is to collect observations of controllers during their regular shifts to provide a perspective of performance not currently being captured by ATC safety management systems. While

NOSS is in its early stages of development, the plans are to link its data to LOSA/TEM and provide a two-way perspective of pilot-controller performance during normal operations. The ability to cross-reference these data sources with TEM should prove useful in designing better aircraft control procedures that both pilots and controllers can accept since both of their perspectives are considered.

LOSA AND ITS FUTURE RESEARCH DIRECTIONS

With LOSA development in place, there are three general directions for future research. The first is the continued assessment and refinement of LOSA methods, measures, procedures and analytic strategies. Second, more research is needed to advance the current knowledge of TEM performance using LOSA. Finally, the third direction involves examining the effectiveness of formally linking LOSA to other airline safety data sources, such as voluntary incident reporting systems.

Continued Assessment and Refinement of LOSA Methodology

One of the major goals of this dissertation was to demonstrate the potential of LOSA to produce valid results. Without validity, any conclusions made from findings are suspect, which is arguably the death knell of any data collection effort. This is why most of Chapter Three not only presented methodology, but discussed decisions made in addressing such threats to validity as pilot trust, flight crew performance reactivity to the presence of an observer, observer bias, and data reliability. However, establishing validity is not an end unto itself. It should be a continuous effort that involves the constant assessment of internal and external validity. Therefore, future research could involve additional third-party assessment of LOSA methods, instrumentation and applicability, particularly from different research disciplines. Insight from other parties and disciplines would refine and improve LOSA as a research method and airline safety process.

Advancing TEM Performance Knowledge with LOSA

A much needed step in research is to advance the current knowledge of TEM performance and its role in aviation safety. This is perhaps best gained through LOSA data. A promising area is mining the LOSA archive, described in Chapter Four, that provides an opportunity to study different TEM performance issues. Judging from the early results presented in Chapter Four, ATC threats, manual handling errors, automation errors, and unstable approaches stand out by their high prevalence and mismanagement rates. However, for TEM issues to be fully understood, multiple methods and data sources are needed to complement LOSA findings. An example of this work is provided by Thomas (2004), who used LOSA-like observations during training flights to study the effect of flight crew behaviors, such as contingency planning and workload management, on TEM processes and outcomes. More research such as this adds to the knowledge base of TEM performance, which can be used by airlines to improve pilot training and standard operating procedures.

Linking LOSA to Other Airline Safety Data Sources

Another area of future research is formally linking LOSA to other airline safety data sources. An example of this would be to study the relationship between LOSA and the voluntary incident reporting system called the Aviation Safety Action Partnership (ASAP). This would be possible since some airlines have begun to capture ASAP information using TEM-based measurement (Harper & Helmreich, 2003; Jones & Tesmer, 1999). Where LOSA collects data from an observer's perspective, ASAP captures it from the reported perspective of the flight crew. Clarifying the similarities and differences between these two data sources could go a long way in helping airlines more efficiently diagnose precursors to incidents and accidents. It would also allow a broader perspective of what airlines are doing well in managing safety risk.

CONCLUSION

LOSA is the product of many years' collaboration between airlines, pilots, researchers, safety managers, and other aviation safety professionals. The different perspectives and contributions they have provided are gratefully acknowledged and reflected throughout this dissertation. Today, the sense is that developing LOSA is only the beginning of what should be an interesting flight.

Appendix A: LOSA Observation Form

Observer Information

Observer ID		Observation Number	

Crew Observation Number (e.g., "1 of 2" indicates segment one for a crew that you observed across two segments)		of	

Flight Demographics

City Pairs (e.g., PIT–CLT)			
A/C Type (e.g., 737–300)			
Pilot flying (Check one)	CA	FO	

Time from Pushback to Gate Arrival (Hours:Minutes)		Local Arrival Time (Use 24 hour time)	
Late Departure? (Yes or No)			

Predeparture / Taxi

Narrative	Your narrative should provide a contextual story for the flight. How did the crew manage threats, crew errors, and significant events? Also, be sure to justify your behavioral ratings.

Takeoff / Climb

Narrative	Your narrative should provide a contextual story for the flight. How did the crew manage threats, crew errors, and significant events? Also, be sure to justify your behavioral ratings.

Cruise

Narrative	Your narrative should provide a contextual story for the flight. How did the crew manage threats, crew errors, and significant events? Also, be sure to justify your behavioral ratings.

Descent / Approach / Land / Taxi

Narrative	Your narrative should provide a contextual story for the flight. How did the crew manage threats, crew errors, and significant events? Also, be sure to justify your behavioral ratings.

Overall Flight

Narrative	This narrative should include your overall impressions of the crew.

Threat Management Worksheet

Threat ID	Threat Description					Threat Management		
	Describe the threat.	Phase of Flight	Altitude	Threat Type	Threat Code	Threat Response	Threat Outcome	Describe how the crew managed or mismanaged the threat with its outcome.
T1								
T2								
T3								
T4								
T5								

Phase of Flight	Threat Type		Threat Code (Sub-Category)	Threat Response	Threat Outcome
	Environmental	Airline			
1 Predeparture / taxi 2 Takeoff/climb 3 Cruise 4 Des/App/Land 5 Taxi/Park	1 Adverse weather 2 Airport 3 ATC 4 Environmental operational Pressure	5 Aircraft malfunction 6 Airlines operational pressure 7 Cabin 8 Dispatch/paperwork 9 Ground/ramp 10 Ground maintenance 11 Manuals/charts	Sub-category codes are found in Appendix B	Was the threat discussed or planned for before it encountered? Yes (1) or No (2)	1 Inconsequential 2 Linked to flight crew error

Error Management Worksheet

Error ID	Error Description							Error Management				
	Describe the error.	Phase of Flight	Altitude	Error Type	Error Code	Proficiency Based?	Linking Threat ID?	Who caused the error?	Error Response	Who detected the error?	Error Outcome	Describe how the crew managed or mismanaged the error with its outcome.
E1												
E2												
E3												
E4												
E5												

Phase of Flight	Error Type			Error Code (Sub-Category)	Proficiency Based?	Who caused? / Who detect?	Error Response	Error Outcome
	Aircraft Handling	Procedural	Communication					
1 Predeparture / taxi 2 Takeoff/climb 3 Cruise 4 Des/App/Land 5 Taxi/Park	1 Automation 2 Flight control 3 Ground navigation 4 Manual handling 5 Systems/Radio/Inst.	6 Briefings 7 Callout 8 Checklist 9 Documentation 10 PF / PNF 11 SOP cross-verif. 12 Other procedural	13 Pilot-to-Pilot 14 Crew-to-External	Sub-category codes are found in Appendix C	1 Yes 2 No	1 Nobody 2 Captain 3 First Officer 4 All crew members 5 LOSA observer 6 ATC 7 Flight attendant 8 Dispatch 9 Ground 10 Maintenance 11 Aircraft systems 12 Other	1 Detected and action 2 Failing to respond	1 Inconsequential 2 Undesired aircraft state 3 Additional error

Undesired Aircraft State Management Worksheet

UAS ID	Undesired Aircraft State (UAS) Description						Undesired Aircraft State (UAS) Management			
	Describe the UAS.	Linking Error ID?	Phase of Flight	Altitude	UAS Type	UAS Code	UAS Response	Who detected the UAS?	UAS Outcome	Describe how the crew managed or mismanaged the UAS with its outcome.
UAS1										
UAS2										
UAS3										
UAS4										
UAS5										

Phase of Flight	UAS Type	UAS Code (Sub-Category)	Who caused? / Who detect?	UAS Response	UAS Outcome
1 Predeparture / taxi 2 Takeoff/climb 3 Cruise 4 Des/App/Land 5 Taxi/Park	1 Aircraft handling 2 Ground navigation 3 Incorrect aircraft configurations	Sub-category codes are found in Appendix D	1 Nobody 2 Captain 3 First Officer 4 All crew members 5 LOSA observer 6 ATC 7 Flight attendant 8 Dispatch 9 Ground 10 Maintenance 11 Aircraft systems 12 Other	1 Detected and action 2 Failing to respond	1 Inconsequential 2 Additional error

Crew Performance Marker Worksheet

1	2	3	4
Poor Observed performance had safety implications	Marginal Observed performance was adequate but needs improvement	Good Observed performance was effective	Outstanding Observed performance was truly noteworthy.

		Phase of Flight Ratings		
Planning Performance Markers		Predeparture / Taxi	Takeoff / Climb	Descent / Approach / Land / Taxi / Park
SOP BRIEFING	The required briefing was interactive and operationally thorough			
PLANS STATED	Operational plans and decisions were communicated and acknowledged			
CONTINGENCY MANAGEMENT	Crew members developed effective strategies to manage threats to safety			
Execution Performance Markers				
MONITOR / CROSS-CHECK	Crew members actively monitored and cross-checked systems and other crew members			
WORKLOAD MANAGEMENT	Operational tasks were prioritized and properly managed to handle primary flight duties			
VIGILANCE	Crew members remained alert of the environment and position of the aircraft			
AUTOMATION MANAGEMENT	Automation was properly managed to balance situational and/or workload requirements			
TAXIWAY / RUNWAY MANAGEMENT	Crew members used caution and kept watch outside when navigating taxiways and runways			
Review / Modify Performance Markers				
EVALUATION OF PLANS	Existing plans were reviewed and modified when necessary			
INQUIRY	Crew members not afraid to ask questions to investigate and/or clarify current plans of action			

Overall Performance Markers		Ratings
COMMUNICATION ENVIRONMENT	Environment for open communication was established and maintained	
LEADERSHIP	Captain showed leadership and coordinated flight deck activities	

Appendix B: Threat Sub-Categories

Environmental Threat Sub-Categories

Adverse Weather

1. Crosswind, tailwind, gusty or high winds aloft
2. Icing only
3. IMC only
4. Thunderstorms/turbulence
5. Turbulence only
6. Windshear
7. z – Other adverse weather threat

Airport

8. Airport construction
9. Contaminated taxiway/runway
10. Lack of or faded signage/markings
11. Complex takeoff/approach requirements (difficult SID/STAR, non-radar environment or noise abatement procedures)
12. Out–of–service/malfunctioning NAVAID/ILS/PAPI
13. Other runway threats
14. Other taxiway/ramp threats
15. z – Other airport threat

ATC

16. ATC command – challenging clearances, late changes
17. ATC error
18. ATC language difficulty
19. ATC non–standard phraseology
20. ATC radio congestion
21. ATC runway change
22. Similar call signs
23. z – Other ATC threat

Environment Operational Pressure

24. ATIS or ACARS communication
25. GPWS warning
26. TCAS RA
27. Terrain
28. Traffic (air or ground congestion)
29. z –Other environmental operational pressure threat

Airline Threat Sub-Categories

Aircraft Malfunction

30. Aircraft malfunction unexpected by crew
31. Automation event or anomaly
32. MEL/CDL with operational implications
33. z – Other aircraft malfunction threat

Airline Operational Pressure

34. Crew scheduling event
35. Operational time pressure (delays, OTP, late arriving pilot or aircraft)
36. z – Other airline operational pressure threat

Cabin

37. Cabin event/distraction/interruption
38. Flight attendant error
39. z – Other cabin threat

Dispatch/Paperwork

40. Dispatch/paperwork error
41. Dispatch/paperwork event
42. z – Other dispatch/paperwork threat

Ground/Ramp

43. Ground crew error
44. Ground handling event
45. z – Other Ground/Ramp threat

Ground Maintenance

46. Maintenance error
47. Maintenance event
48. z – Other ground maintenance threat

Manuals/Charts

49. Chart error
50. Manual error
51. z – Other manuals/chart threat

Appendix C: Error Sub-Categories

Aircraft Handling Error Sub-Categories

Automation

1. Discretionary omission of FMC/FMGC data
2. Failure to execute a FMC/FMGC mode when needed
3. Failure to execute a MCP/FCU mode when needed
4. Failure to use flight directors
5. Manual aircraft control with MCP/FCU mode engaged
6. Omitted/wrong waypoint or route settings entered into the FMC/FMGC
7. Other MCP/FCU error
8. Other wrong FMC/FMGC entries
9. Wrong altitude entered into the FMC/FMGC
10. Wrong approach selected in FMC/FMGC
11. Wrong FMC/FMGC format for input
12. Wrong FMC/FMGC page displayed
13. Wrong MCP/FCU altitude setting dialed
14. Wrong MCP/FCU course setting dialed
15. Wrong MCP/FCU heading set or dialed
16. Wrong MCP/FCU mode executed
17. Wrong MCP/FCU mode left engaged
18. Wrong MCP/FCU navigation select setting (NAV,GPS,ILS,VOR switch)
19. Wrong MCP/FCU setting on autothrottle switch
20. Wrong MCP/FCU speed setting dialed
21. Wrong MCP/FCU vertical speed/flight path angle setting
22. Wrong mode executed in the FMC/FMGC
23. Wrong mode left engaged in the FMC/FMGC
24. Wrong present position entered into the FMC/FMGC
25. Wrong setting on the MCP/FCU autopilot or flight director switch
26. Wrong speed setting entered into the FMC/FMGC
27. Wrong weight and balance calculations entered into FMC/FMGC

28. Intentional Noncompliance – Nonstandard or wrong MCP/FCU settings
29. Intentional Noncompliance – Nonstandard automation usage
30. Intentional Noncompliance – Other automation error
31. z – Other automation error

Flight Control

32. Attempting to use INOP controls
33. Decision to use wrong thrust/power
34. Failure to engage thrust reversers on landing
35. Failure to raise or lower landing gear on schedule
36. Wrong autothrottle setting
37. Wrong autobrake setting
38. Wrong flaps setting
39. Wrong thrust/power settings
40. Wrong speed brakes setting
41. Wrong spoilers setting
42. Wrong stab trim settings
43. Wrong thrust reversers setting
44. Intentional Noncompliance – Failure to arm spoilers
45. Intentional Noncompliance – Use of excessive power on pushback
46. z – Intentional Noncompliance – Other flight control error
47. z – Other flight control error

Ground Navigation

48. Attempting or turning down wrong gate/taxiway/ramp/hold spot
49. Attempting or turning down wrong runway
50. Attempting to taxi off C/L
51. Failure to hold short
52. Missed gate
53. Missed runway
54. Missed taxiway
55. Taxi on taxiway/runway with oncoming traffic
56. Taxi too close to other aircraft
57. Taxi too fast

58. Intentional Noncompliance – Taxi above speed limit
59. z.– Intentional Noncompliance – Other ground navigation error
60. z – Other ground navigation error

Manual Handling

61. Attempting or lining up for incorrect runway for landing
62. Unnecessary low maneuver on approach
63. Excessive brake use
64. Landing deviation by choice
65. Lateral or vertical deviation by choice
66. Decision to navigate through known bad weather that increased risk
67. Speed deviation by choice
68. Unintentional bank deviation
69. Unintentional crosswind technique
70. Unintentional landing deviation
71. Unintentional lateral deviation
72. Unintentional pitch deviation
73. Unintentional speed deviation
74. Unintentional vertical deviation
75. Unintentional vertical speed deviation
76. Unintentional weather penetration
77. Unintentional yaw deviation
78. Intentional Noncompliance – Intentionally flying below the G/S
79. Intentional Noncompliance – Intentionally flying a nonstandard visual approach
80. Intentional Noncompliance – Intentionally not following published Jepp procedures
81. Intentional Noncompliance – Unauthorized speed deviation
82. z – Intentional Noncompliance – Other manual flying error
83. z – Other manual flying error

Systems/Instrument/Radio

84. Failure to respond to GPWS warnings
85. Failure to respond to TCAS warnings
86. Failure to turn on A/C packs (no pressurization)
87. Incorrect nav display setting

88. Wrong ACARS entries
89. Wrong altimeter settings
90. Wrong anti–ice setting
91. Wrong ATC frequency dialed/selected
92. Wrong ATIS frequency dialed
93. Wrong bug settings
94. Wrong display switch setting
95. Wrong fuel switch setting
96. Wrong nav radio frequency dialed
97. Wrong panel setup for engine start
98. Wrong radar settings
99. Wrong squawk
100. Wrong TCAS setting
101. Intentional Noncompliance – Failure to respond to GPWS warnings
102. Intentional Noncompliance – Failure to respond to TCAS warnings
103. Intentional Noncompliance – Setting altimeters before the transition altitude or trans level
104. Intentional Noncompliance – Using equipment placarded as INOP
105. Intentional Noncompliance – Failure to respond to overspeed warning
106. Intentional Noncompliance – Unauthorized response to aircraft warning
107. Intentional Noncompliance – Wrong bug settings
108. z – Intentional Noncompliance – Other systems/inst/radio error
109. z – Other systems/inst/radio error

Procedural Error Sub-Categories

Briefing

110. Brief performed late
111. Incorrect/incomplete approach brief
112. Incorrect/incomplete depart review/takeoff brief
113. Incorrect/Incomplete F/A brief
114. Omitted approach briefing
115. Omitted departure review/takeoff briefing
116. Omitted required engine–out briefing
117. Omitted required F/A briefing

118. Intentional Noncompliance – Incorrect/incomplete approach briefing
119. Intentional Noncompliance – Omitted depart review/T/O briefing
120. Intentional Noncompliance – Omitted handover briefing
121. Intentional Noncompliance – Omitted required engine–out briefing
122. Intentional Noncompliance – Incorrect/incomplete depart review/T/O briefing
123. Intentional Noncompliance – Incorrect/Incomplete F/A brief
124. Intentional Noncompliance – Intentional late brief
125. Intentional Noncompliance – Omitted approach briefing
126. Intentional Noncompliance – Omitted required flight attendant briefing
127. z – Intentional Noncompliance Other briefing error
128. z – Other briefing error

Callout

129. Incorrect approach callouts
130. Incorrect climb or descent callouts
131. Incorrect V–speed callouts
132. Omitted approach callouts
133. Omitted climb or descent callouts
134. Omitted landing callouts
135. Omitted altitude callouts
136. Omitted V–speed callouts
137. Intentional Noncompliance – Omitted altitude callouts
138. Intentional Noncompliance – Nonstandard altitude callouts
139. Intentional Noncompliance – Nonstandard approach callouts
140. Intentional Noncompliance – Omitted approach callouts
141. Intentional Noncompliance – Omitted climb or descent callouts
142. Intentional Noncompliance – Nonstandard climb or descent callouts
143. Intentional Noncompliance – Nonstandard V–speed callouts
144. Intentional Noncompliance – Omitted V–speed callouts
145. z – Intentional Noncompliance – Other callout error
146. z – Other callout error

Checklist

147. Checklist not performed to completion

148. Checklist performed late or at the wrong time

149. Completed checklist not called 'complete'

150. Missed checklist item

151. Omitted abnormal checklist

152. Omitted checklist

153. Wrong checklist performed

154. Wrong response to a challenge on a checklist

155. Intentional Noncompliance – Checklist performed from memory

156. Intentional Noncompliance – Completed checklist not called 'complete'

157. Intentional Noncompliance – Omitted abnormal checklist

158. Intentional Noncompliance – Self–initiated checklist (not called for by CA)

159. Intentional Noncompliance – Use of nonstandard checklist protocol

160. Intentional Noncompliance – Checklist not performed to completion

161. Intentional Noncompliance – Checklist performed as "to–do" checklist

162. Intentional Noncompliance – Checklist performed late or at wrong time

163. Intentional Noncompliance – Omitted checklist

164. Intentional Noncompliance – Self–initiated checklist (not called for by PF)

165. Intentional Noncompliance – Self–performed checklist – no challenge and response

166. z – Intentional Noncompliance – Other checklist error

167. z – Other checklist error

Documentation

168. Incorrect or failing to make an entry into the logbook

169. Miscalculation of hold times

170. Misinterpreted items on flight documentation

171. Missed items on flight documentation (flight plan, NOTAM, dispatch release)

172. Wrong clearance recorded

173. Wrong fuel information recorded

174. Wrong or no ATIS information recorded

175. Wrong or no Jepp pages out (approach plates, 10–7 page, etc.)

176. Wrong performance chart used

177. Wrong runway information recorded

178. Wrong times calculated in flight plan

179. Wrong V–speeds recorded

180. Wrong weight and balance information recorded

181. Intentional Noncompliance – Failure to make logbook entry

182. Intentional Noncompliance – No Jepp pages out (approach charts, 10–7 page, etc.)

183. Intentional Noncompliance – T/O without proper weight & balance figures

184. z – Intentional Noncompliance – Other documentation error

185. z – Other documentation error

PF/PNF Duty (PF – Pilot Flying and PNF – Pilot Not Flying)

186. PF makes own FMC/FMGC changes

187. PF makes own MCP/FCU changes

188. PNF performs PF automation duties

189. Intentional Noncompliance – PF makes own FMC/FMGC changes

190. Intentional Noncompliance – PF sets own flight controls or switches

191. Intentional Noncompliance – PF makes own MCP/FCU changes

192. Intentional Noncompliance – PNF carried out PF duties

193. z – Intentional Noncompliance – Other PF/PNF duty error

194. z – Other PF/PNF duty error

SOP Cross Verification

195. Failure to clarify MEL or logbook entry

196. Failure to cross–verify altimeter settings

197. Failure to cross–verify automation navigation with raw data

198. Failure to cross–verify clearance

199. Failure to cross–verify documentation or paperwork

200. Failure to cross–verify FMC/FMGC inputs

201. Failure to cross–verify MCP/FCU/altitude alerter changes

202. Failure to cross–verify speed before flap selection

203. Failure to monitor engine start

204. Omitted flight mode annunciation

205. Intentional Noncompliance – Failure to cross–verify paperwork

206. Intentional Noncompliance – Failure to cross–verify altimeter settings

207. Intentional Noncompliance – Failure to cross–verify FMC/FMGC/CDU changes

208. Intentional Noncompliance – Failure to cross–verify manual with paperwork

209. Intentional Noncompliance – Failure to cross–verify MCP/FCU/altitude alerter changes
210. Intentional Noncompliance – Nonstandard cross–verification
211. Intentional Noncompliance – Omitted flight mode annunciation
212. z – Intentional Noncompliance – Other SOP cross verification error
213. z – Other SOP cross verification error

Other Procedural

214. Admin duties performed at inappropriate time
215. Decision not to ask FA to stay seated for bad weather
216. Decision not to turn on seat belt sign in bad WX
217. Failure to G/A after stabilized approach window
218. Omitted RVSM procedure
219. Pushback without clearing right or left
220. Taxi in position and hold with unready cabin
221. Unintentional operation with MEL
222. Wrong MEL action performed
223. Crew omitted cabin/flight attendant call
224. Intentional Noncompliance – Failure to use proper WX SOP
225. Intentional Noncompliance – Operation with unresolved aircraft malfunction
226. Intentional Noncompliance – Operation with unresolved MEL
227. Intentional Noncompliance – Admin duties performed at inappropriate times
228. Intentional Noncompliance – Failure to G/A after stabilized approach window
229. Intentional Noncompliance – Nonstandard "ready to push" procedures
230. Intentional Noncompliance – Taxi duties performed before leaving runway
231. Intentional Noncompliance – Taxi–in duties performed before crossing an active runway
232. Intentional Noncompliance – Taxi–in or out without wing walkers
233. z – Intentional Noncompliance – Other procedural error
234. z – Other procedural error

Communication Error Sub-Categories

Pilot to Pilot

235. Crew miscommunication of information
236. Failure to communicate approach information
237. Misinterpretation of ATIS

238. Missed command within crew

239. Wrong airport communicated

240. Wrong nav aid communicated

241. Wrong runway communicated

242. Wrong taxiway/ramp/gate/hold spot communicated

243. Intentional Noncompliance – Sterile cockpit violation

244. z – Intentional Noncompliance– Other pilot to pilot communication error

245. z – Other pilot to pilot communication error

Crew to External

246. Crew did not repeat ATC clearance

247. Crew omitted ATC call

248. Failure to give readbacks or callbacks to ATC

249. Incomplete clearance readback

250. Misinterpretation of ATC instructions

251. Misinterpretation of ground instructions

252. Misinterpretation of tower instructions

253. Missed ATC calls

254. Omitted call signs to ATC

255. Omitted non–radar environment report to ATC

256. Omitted position report to ATC

257. Use of nonstandard ATC phraseology

258. Wrong position report

259. Wrong readbacks or callbacks to ATC

260. Intentional Noncompliance – Accepting a landing clearance 10+ knot tailwind

261. Intentional Noncompliance – Omitted ATC calls

262. Intentional Noncompliance – Omitted call signs to ATC

263. Intentional Noncompliance – Omitted non–radar environment report to ATC

264. Intentional Noncompliance – Use of nonstandard ATC phraseology

265. Intentional Noncompliance – Omitted position report to ATC

266. z – Intentional Noncompliance – Other crew to external communication error

267. z – Other crew to external communication error

Appendix D: Undesired Aircraft State Sub-Categories

Aircraft Handling

1. Abrupt aircraft control (attitude)
2. Vertical deviation
3. Lateral deviation
4. Low speed deviation
5. High speed deviation
6. Unnecessary weather penetration
7. Unauthorized airspace penetration
8. Operation outside aircraft limitations
9. Unstable approach
10. Long landing
11. Short landing
12. Floated landing
13. Firm landing
14. Off centerline landing
15. z – Other aircraft handling undesired aircraft state

Ground Navigation

16. Runway incursion
17. Taxiway incursion
18. Wrong taxiway/ramp
19. Wrong hold spot
20. Taxi too fast
21. z – Other taxi undesired aircraft state

Incorrect Aircraft Configuration

22. Incorrect automation configuration
23. Incorrect engine configuration
24. Incorrect flight control configuration
25. Incorrect systems/instruments/radio configuration
26. Incorrect weight and balance configuration
27. z – Other configuration undesired aircraft state

Appendix E: LOSA Observer Feedback Form

Observer ID: ________________________________

Airline: _____________________________

LOSA TEM Exercise #1 (7 Items)

1. Select the best option that captures each event. (ANSWERS ARE IN BOLD)

A. FO mis-set the speed bugs for takeoff – immediately caught on cross–verification.

○ Threat ○ **Error** ○ Error and undesired aircraft state ○ No event

B. Crew had difficulty understanding ATC controller accents during the approach.

○ **Threat** ○ Error ○ Error and undesired aircraft state ○ No event

C. Dispatch flight plan had the incorrect zero fuel weight – detected by the crew.

○ **Threat** ○ Error ○ Error and undesired aircraft state ○ No event

D. Wrong V2 takeoff speed entered into the FMS – detected by the crew after takeoff.

○ Threat ○ Error ○ **Error and undesired aircraft state** ○ No event

E. FO extended flaps at 200 knots and exceeded flaps 15 limit speed by 5 knots.

○ Threat ○ Error ○ **Error and undesired aircraft state** ○ No event

F. Captain intentionally did not conduct a required takeoff briefing to save time.

○ Threat ○ **Error** ○ Error and undesired aircraft state ○ No event

G. Crew failed to make level–off callouts as required by SOP throughout climb.

○ Threat ○ **Error** ○ Error and undesired aircraft state ○ No event

LOSA TEM Exercise #2 (15 Items)

2. Read the narrative and complete the exercise that follows.

Descent/Approach/Land

Briefing to TOD – Briefing for the arrival was thorough, interactive and took place well before TOD. The weather was CAVOK and briefed the arrival over SABINE to 27 ILS.

Transition – 10000 – The approach controller kept dropping aircraft callsigns from his transmissions throughout the descent. The crew confirmed with the controller whether the transmissions in question were for them or someone else. All checklists and cross–verifications per SOP.

Final Approach – On vectors to RWY 27, ATC gave the crew a late turn to an intercept heading for the RWY 27 final approach. The First Officer, hand–flying, used insufficient bank to capture the localizer and exceeded two dots. ATC observed the overshoot and clears the crew to land 26L. The ILS 26L was intercepted at 1500 feet. Flaps were selected at 1300 ft with the checklist completed. The aircraft was stable by 1000 ft. per SOP and the landing was within the touchdown zone.

Taxi-In – The FO misinterpreted ATC and believed that they had been cleared to cross runway 28R when they were not. The Captain was unsure and asked the FO to confirm before they moved any further. The Tower controller confirmed that they needed to hold for a company 737 that was cleared to takeoff.

Instructions – From the narrative, list all threats, errors and undesired aircraft states along with their crew responses and outcomes.

Threat Type and Description	Threat Response	Threat Outcome
ATC – ATC dropping callsigns	Not discussed or planned for before the threat was encountered	Inconsequential
ATC – Late turn to an intercept	Not discussed or planned for before the threat was encountered	Linked to flight crew error
Error Type and Description	Error Reponse	Error Outcome
Manual Handling – Insuffcient bank by FO	Failing to respond	Undesired aircraft state
Crew to External Communication – FO misintepreted ATC	Detected and action	Inconsequential
UAS Type and Description	UAS Response	UAS Outcome
Aircraft handling state – Lateral deviation – FO overshot localizer by 2 dots	Detected and action	Inconsequential

Note: Observers can use the LOSA Data Collection Tool to complete the exercise.

Appendix F: Sample Error Organizational Profile (Zed Airlines)

Error Type	Error Prevalence Index	Error Mismanagement Index	Mismanaged Error/Error Count
All Errors	77%	20%	139/710
Aircraft Handling Errors			
System/Instrument/Radio	18%	34%	22/65
Automation	18%	24%	16/66
Manual Handling/Flight Control**	15%	77%	41/53
Ground Navigation	3%	NC*	6/9
Procedural Errors			
Callout	32%	5%	8/149
Checklist	20%	15%	13/88
SOP Cross-Verification	18%	14%	9/66
Briefing	15%	4%	2/55
PF/PNF Duty	10%	3%	1/35
Other Procedural	9%	42%	13 /31
Documentation	4%	NC*	2/11
Communication Errors			
Crew to External	20%	7%	5/72
Pilot to Pilot	3%	NC*	1/9

Definitions:

Error Prevalence – Percentage of LOSA observations with one or more errors

Error Mismanagement Index – Percent of errors linked to an additional error or undesired aircraft state

Notes:

N=309 LOSA observations (Zed Airlines)

* Not calculated (NC) due to the potentially misleading nature of percentages at lower counts. A minimum of 20 errors observed is needed for percentages to be calculated.

** Manual flying and flight controls error categories are often combined since they both involve the physical manipulation of the aircraft.

Appendix F: Sample UAS Organizational Profile (Zed Airlines)

UAS Type	UAS Prevalence Index	UAS Mismanagement Index	Mismanaged UAS /7 UAS Count
All Undesired Aircraft States	24%	13%	13/87
Aircraft Handling States			
Vertical Deviations	4%	Fewer than 20 states*	1/12
Landing Deviations	4%		1/11
Lateral Deviations	3%		1/9
Unstable Approach	3%		9/9
Speed Deviations	2%		1/6
Ground Navigation States			
Wrong taxiway or Taxi too Fast	2%	Fewer than 20 states*	0/6
Configuration States			
Systems	4%	Fewer than 20 states*	0/13
Automation	3%		0/9
Flight Control	2%		0/6
Engine	2%		0/6

Definitions:

UAS Prevalence – Percentage of LOSA observations with one or more undesired aircraft states

UAS Mismanagement Index – Percent of undesired aircraft states linked to an additional error or undesired aircraft state

Notes:

N=309 LOSA observations (Zed Airlines)

* Not calculated (NC) due to the potentially misleading nature of percentages at lower counts. A minimum of 20 states observed is needed for percentages to be calculated.

Bibliography

Bagnara, S., & Rizzo, A. (1989). A methodology for analysis of error processes in human-computer interaction. In M. J. Smith & G. Salvendy (Eds.), *Work with computers: Organizational, management, stress, and health aspects* (pp. 605-612). Amsterdam: Elsevier Science.

Bakeman, R. (2000). Behavioral observation and coding. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 138-159). Cambridge, UK: Cambridge University Press.

Boeing. (2003). Current market outlook 2003. 30 June 2003 <http://www.boeing.com>.

Boeing. (2004). Statistical summary of commercial jet airplane accidents: Worldwide operations 1959-2003. Seattle, Washington: Boeing.

Burke, C. S., Wilson, K. A., & Salas, E. (2003). Teamwork at 35,000 feet: Enhancing safety through team training. *Human Factors and Aerospace Safety, 3*(4), 287-312.

Campbell, D. T. (1961). The mutual methodological relevance of anthropology and psychology. In F. L. K. Hsu (Ed.), *Psychological anthropology* (pp. 333-352). Homewood, Illinois: Dorsey.

Cannon-Bowers, J. A., Tannebaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies and establishing team training requirements. In R. A. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 330-380). San Francisco, CA: Jossey-Bass.

Castano, D. J., & Graeber, R. C. (2002). Aircraft manufacturer eager to learn from reliable data on normal flight operations. *ICAO Journal, 57*(4), 10-11, 27.

Champanis, A., Gardner, W. R., & Morgan, C. T. (1947). *Lectures on men and machines: An introduction to human engineering.* Baltimore, Maryland: Systems Research Laboratory.

Chidester, T. R. (2003). Understanding normal and atypical operations through analysis of flight data. In R.S. Jensen (Ed.), *Proceedings of the Twelfth International Symposium on Aviation Psychology* (pp. 239-242). Dayton, OH: Wright State University.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand-McNally.

Cooper, G. E., White, M. D., & Lauber, J. K. (Eds.). (1980). *Resource management on the flight deck: Proceedings of the NASA/Industry Workshop (NASA CP-120).* Moffett Field, CA: NASA-Ames Research Center.

Costa, G. (1999). Fatigue and biological rhythms. In D. J. Garland, J. A. Wise & V. D. Hopkin (Eds.), *Handbook of aviation human factors* (pp. 233-255). Manwah, New Jersey: Lawrence Erlbaum Associates.

Craig, H. (2003). Cathay Pacific LOSA. *Proceedings of the First ICAO-IATA LOSA and TEM Conference* (pp. 32-34). Dublin, Ireland: International Civil Aviation Organization (ICAO).

Dancer, D. D., Braukmann, C. J., Schumaker, J. B., Kirigin, K. A., Willner, A. G., & Wolf, M. M. (1978). The training and validation of behavior observation and description skills. *Behavior Modifications, 2*, 113-134.

Davenport, T. H., Harris, J. G., De Long, D. W., & Jacobson, A. L. (2001). Data to knowledge to results: Building an analytic capability. *California Management Review, 43*(2), 117-138.

Diehl, A. E. (1991). Human performance and systems safety considerations in aviation mishaps. *International Journal of Aviation Psychology, 1*(2), 97-106.

Douglas, J. D. (1976). *Investigative social research*. Beverly Hills, CA: Sage Publications.

Edmunds, A., & Morris, A. (2000). The problem of information overload in business organizations: A review of the literature. *International Journal of Information Management, 20*, 17-28.

Edwards, E. (1988). Introductory overview. In E. L. Weiner & D. Nagel (Eds.), *Human factors in aviation* (pp.3-25). San Diego, CA: Academic Press.

Federal Aviation Administration (FAA). (2001). *FAA long range aerospace forecasts: Fiscal years 2015, 2020, and 2025*. Washington DC: FAA Office of Aviation Policy and Plans.

Federal Aviation Administration (FAA). (2004). *Capacity needs in the national airspace system: An analysis of airport and metropolitan area demand and operational capacity in the future*. Washington DC: FAA.

Fischer, J. W. (2001). *Aviation congestion: Proposed non-air traffic control remedies* (No. RS20914). Washington DC: The Library of Congress: Congressional Research Service (CRS).

Fitts, P. M., & Jones, R. E. (1947). *Psychological aspects of instrument display: Analysis of 270 "pilot error" experiences in reading and interpreting aircraft instruments* (No. TSEAA-694-12A). Dayton, OH: Aeromedical Laboratory, Air Material Command.

Flanagan, J. C. (1947). *The aviation psychology program in the Army Air Force* (No. 17). Washington, DC: Army Air Forces Aviation Psychology Program.

Foushee, H. C., Lauber, J. K., Baetge, M. M., & Acomb, D. B. (1986). *Crew performance as a function of exposure to high density, short haul duty cycles* (No. NASA Technical Memo 88322). Moffett Field, CA: NASA-Ames Research Center.

Flight Safety Foundation (FSF). (1998). Killers in aviation: FSF task force presents facts about approach and landing and controlled flight into terrain accidents. *Flight Safety Digest, 17*, 1-121.

General Accounting Office (GAO). (2002). *FAA needs to better prepare for impending wave of controller attrition* (No. GAO-02-591). Washington DC: United States GAO.

Geertz, C. (1973). Thick description: Toward an interpretative theory of culture. In C. Geertz (Ed.), *The interpretation of cultures* (pp. 3-30). New York: Basic Books.

Gilner, J. A., & Morgan, G. A. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Manwah, New Jersey: Lawrence Erlbaum Associates.

Greeves, B. (2002). LOSA: What's in it for the Line Pilot? *Proceedings of the Third LOSA Week* (pp. 61-65), Dubai, United Arab Emirates: International Civil Aviation Organization.

Gunther, D. (2002). Threat and error management training counters complacency, fosters operational excellence. *ICAO Journal, 57*(4), 12-13.

Gunther, D. (2003). The value of non-jeopardy reporting to a safety management system. In R.S. Jensen (Ed.), *Proceedings of the Twelfth International Symposium on Aviation Psychology* (pp. 453-456). Dayton, OH: Wright State University.

Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, *47*, 307-338.

Hale, A., Baram, M., & Hovden, J. (1998). Perspectives on safety management and change. In A. Hale & M. Baram (Eds.), *Safety management: The challenge of change* (pp. 1-18). Oxford: Pergamon.

Harper, M. L., & Helmreich, R. L. (2003). Applying the threat and error management model to an Aviation Safety Action Program. In R.S. Jensen (Ed.), *Proceedings of the Twelfth International Symposium on Aviation Psychology* (pp. 491-495). Dayton, OH: Wright State University.

Hartman, D. P. (1982). Assessing dependability of observational data. In D. P. Hartman (Ed.), *Using observers to study behavior* (pp. 51-65). San Francisco: Jossey-Bass.

Hawkins, R. P. (1982). Developing a behavior code. In D. P. Hartman (Ed.), *Using observers to study behavior* (pp. 21-35). San Francisco: Jossey-Bass.

Haynes, S. N., & Horn, W. F. (1982). Reactivity in behavioral observations: A methodological and conceptual critique. *Behavioral assessment, 4*, 369-385.

Heinrich, H., Petersen, D., & Roos, N. (1980). *Industrial accident prevention: A safety management approach* (5th ed.). New York: McGraw-Hill.

Helmreich, R. L. (1997). Managing human error in aviation. *Scientific American, 276*, 40-45.

Helmreich, R. L., Butler, R. E., Taggart, W. R., & Wilhelm, J. A. (1994). *The NASA/ University of Texas/ FAA Line/LOS Checklist: A behavioral marker-based checklist for CRM skills assessment*. Austin, Texas: The University of Texas at Austin.

Helmreich, R. L., & Foushee, H. C. (1993). Why Crew Resource Management? Empirical and theoretical bases of human factors training in aviation. In E. L. Weiner, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 3-45). New York: Academic Press.

Helmreich, R. L., & Merritt, A. C. (1998). *Culture at work in aviation and medicine: National, organizational, and professional influences*. Aldershot: Ashgate.

Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999). The evolution of Crew Resource Management training in commercial aviation. *International Journal of Aviation Psychology, 9*(1), 19-32.

Helmreich, R. L., & Taggart, W. R. (1995). CRM: Where are we today? In *Proceedings of the CRM Industry Workshop*, Seattle, Washington: September 12-13, 1995.

Helmreich, R. L., Wilhelm, J. A., Klinect, J. R., & Merritt, A. C. (2001). Culture, error and crew resource management. In E. Salas, C. A. Bowers & E. Edens (Eds.), *Improving teamwork in organizations: Applications of resource management training* (pp. 305-331). Mahwah, NJ: Lawrence Erlbaum Associates.

Herbert, J., & Attridge, C. A. (1975). A guide for developers and users of observation systems and manuals. *American Educational Research Journal, 12*(1), 1-20.

Heyns, R. W., & Lippitt, R. (1954). Systematic observational techniques. In G. Lindzey (Ed.), *Handbook of social psychology: Vol. 1* (pp. 370-404). Cambridge, MA: Addison-Wesley.

Hines, W. E. (1998). *Teams and technology: Flightcrew performance in standard and automated aircraft*. Unpublished doctoral dissertation, The University of Texas at Austin.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method (CREAM)*. Oxford, England: Elsevier Science.

Hopkins, A. (2001). Was Three Mile Island a 'normal accident'? *Journal of Contingencies and Crisis Management, 9*(2), 65-72.

International Civil Aviation Organization (ICAO). (2001). *Outlook for air transport to the year 2010* (ICAO Circular 281). Montreal, Canada.

International Civil Aviation Organization (ICAO). (2002). *ICAO human factors training manual* (No. ICAO Document 9683). Montreal, Canada: The International Civil Aviation Organization.

International Civil Aviation Organization (ICAO). (2004). *The world of civil aviation 2001-2004* (ICAO Circular 291). Montreal, Canada.

John, O. P., & Benet-Martinez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339-369). Cambridge, UK: Cambridge University Press.

Johnson, J. M. (1975). *Doing field research*. New York: Free Press.

Johnson, S., & Bolstad, O. D. (1973). Methodological issues in naturalistic observation: Some problems and solutions for field research. In L. A. Hamerlynck, L. C. Handy & E. J. Mash (Eds.), *Behavior change: Methodology, concepts, and practice*. Champaign, IL: Research Press.

Jones, S. G., & Tesmer, B. A. (1999). A new tool for investigating and tracking human factors issues in incidents. In R.S. Jensen (Ed.), *Proceedings of the Tenth International Symposium on Aviation Psychology* (pp. 696-700). Columbus, OH: The Ohio State University.

Kantowitz, B. H., & Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships*. New York: Wiley.

Katzer, J., & Fletcher, P. (1996). *The information environment of managers*. New York: Neal-Schuman.

Kazdin, A. E. (1982). Observer effects: Reactivity of direct observation. In D. P. Hartman (Ed.), *Using observers to study behavior* (pp. 5-19). San Francisco: Jossey-Bass.

Kirmeyer, S. L. (1985). Employee reactivity and adaptation to observation on the job. *Journal of Environmental Psychology, 5*(4), 355-372.

Koniger, P., & Janowitz, K. (1995). Drowning in information, but thirsty for knowledge. *International Journal of Information Management, 15*(1), 5-16.

Koonce, J. M. (1984). A brief history of aviation psychology. *Human Factors, 26*(5), 499-508.

Koonce, J. M. (1999). A historical overview of human factors in aviation. In D. J. Garland, J. A. Wise & V. D. Hopkin (Eds.), *Handbook of aviation human factors*. Manwah, New Jersey: Lawrence Erlbaum Associates.

Kotarba, J. A. (1980). Discovering amorphous social experience: The case of chronic pain. In W. B. Shaffir, R. A. Stebbins & A. Turowetz (Eds.), *Fieldwork experience: Qualitative approaches to social research* (pp. 57-67). New York: St. Martin's Press.

Kuder, G. F., & Richardson, M. W. (1937). The theory of estimation of test reliability. *Psychometrika, 2*, 151-160.

La Porte, T. R., & Consolini, P. (1991). Working in practice but not in theory: Theoretical challenges of high reliability organizations. *Journal of Public Administration Research and Theory, 1*(1), 19-47.

Lauber, J. K. (1993). Forward. In E. L. Weiner, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. xv-xviii). New York: Academic Press.

Law, J. R., & Sherman, P. J. (1995). Do raters agree? Assessing inter-rater agreement in the evaluation of air Crew Resource Management skills. In R.S. Jensen (Ed.), *Proceedings of the Eighth International Symposium on Aviation Psychology* (pp. 608-612). Columbus, OH: The Ohio State University.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications.

Lord, R. C. (1985). Accuracy in behavioral measurement: An alternative definition based on rater errors in the observation of behavior. *Journal of Applied Psychology, 70*, 66-71.

March, J. G., Sproull, L. S., & Tamuz, M. (1991). Learning from samples of one or fewer. *Organization Science, 2*(1), 1-14.

Maurino, D. E. (1999). Safety prejudices, training practices and CRM: A mid-point perspective. *International Journal of Aviation Psychology, 9*(3), 413-427.

Maurino, D. E. (2001). Proactive safety culture: Do we need human factors? In B. J. Hayward & A. R. Lowe (Eds.), *Aviation resource management, Vol. 1* (pp. 1-10). Aldershot: Ashgate.

Maurino, D. E. (2002). Line Operations Safety Audits have become the central focus of ICAO's Human Factors Programme. *ICAO Journal, 57*(4), 5.

Maurino, D. E., Reason, J. T., Johnston, N., & Lee, R. B. (1995). *Beyond aviation human factors*. Aldershot: Ashgate.

Maurino, D. E., & Ruitenberg, B. (2004). ICAO examining ways to monitor safety during normal ATS operations. *ICAO Journal, 59*(3), 14-16.

McGrath, J. E. (1964). *Social psychology: A brief introduction*. New York: Holt, Rinehart, and Winston.

Merritt, A. C. (2000). Culture in the cockpit: Do Hofstede's dimensions replicate? *Journal of Cross-Cultural Psychology, 31*(3), 283-301.

Nagel, D. (1988). Human error in aviation operations. In E. L. Weiner & D. Nagel (Eds.), *Human factors in aviation* (pp. 263-303). San Diego, CA: Academic Press.

National Safety Council (NSC). (2003). *Injury facts*. Itasca, IL: NSC.

National Transportation Safety Board (NTSB). (2000). *In-flight Breakup over The Atlantic Ocean, Trans World Airlines Flight 800, Boeing 747-131, N931119, Near East Monriches, New York, July 17, 1996* (No. NTSB/AAR-00/03). Washington, DC: NTSB.

National Transportation Safety Board (NTSB). (2002a). *Annual Review of Aircraft Accident Data: U.S. Air Carrier Operations, Calendar Year 1999* (No. NTSB/ARC-02-03). Washington, DC: NTSB.

National Transportation Safety Board (NTSB). (2002b). *Aircraft Accident Report: Loss of control and Impact with Pacific Ocean, Alaska Airlines Flight 261, McDonnell Douglas MD-83, N963AS, about 2.7 Miles North of Anacapa Island, California, January 31, 2000* (No. NTSB/AAR-02/01). Washington, DC: NTSB.

Patterson, D. R., & Sechrest, L. (1983). Nonreactive measures in psychotherapy outcome research. *Clinical Psychology Review, 3*, 391-416.

Perrow, C. (1981). The President's commission and the normal accident. In V. B. Shelanski, D. L. Sills & C. P. Wolf (Eds.), *The accident at Three Mile Island: The human dimensions* (pp. 173-184). Boulder, Colorado: Westview Press.

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies* (1st ed.). New York: Basic Books.

Perrow, C. (1999). *Normal accidents: Living with high-risk technologies* (2nd ed.). New York: Basic Books.

Pidgeon, N. (1997). The limits to safety? Culture, politics, learning and man-made disasters. *Journal of Contingencies and Crisis Management, 5*(1), 1-14.

Previc, F. H., & Ercoline, W. R. (2004). *Spatial disorientation in aviation*. Reston, VA: American Institute of Aeronautics and Astronautics.

Prince, C., & Salas, E. (1999). Team processes and their training in aviation. In D. J. Garland, J. A. Wise & V. D. Hopkin (Eds.), *Handbook of aviation human factors* (pp. 193-213). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Rasmussen, J., Duncan, K., & Leplat, J. (1987). *New technology and human error*. Cambridge, England: Wiley.

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

Reason, J. T. (1997). *Managing the risks of organizational accidents*. Aldershot: Ashgate.

Rees, J. (1994). *Hostages of each other: The transformation of nuclear safety since Three Mile Island*. Chicago: University of Chicago Press.

Reid, J. B. (1970). Reliability assessment of observational data: A possible methodological problem. *Child Development, 41*, 1143-1150.

Reid, J. B. (1982). Observer training in naturalistic research. In D. P. Hartman (Ed.), *Using observers to study behavior* (pp. 37-50). San Francisco: Jossey-Bass.

Riesman, D., & Watson, J. (1964). The sociability project: A chronicle of frustration and achievement. In P. E. Hammond (Ed.), *Sociologists at work* (pp. 235-321). New York: Basic Books.

Rizzo, A., Bagnara, S., & Visciola, M. (1987). Human error detection processes. *International Journal of Man-Machine Studies, 27*(6), 555-570.

Roberts, K. H. (1993). *New challenges to understanding organizations*. New York: Macmillan.

Rochlin, G. I. (1993). Defining "high reliability" organizations in practice: A taxonomic prologue. In K. H. Roberts (Ed.), *New challenges to understanding organizations* (pp. 11-32). New York: Macmillan.

Romanczyk, R. G., Kent, R. N., Diament, C., & O'Leary, K. D. (1973). Measuring the reliability of observational data: A reactive process. *Journal of Applied Behavior Analysis, 6*, 175-186.

Roscoe, S. N. (1980). *Aviation psychology*. Ames, Iowa: State University Press.

Roscoe, S. N. (1997). *The adolescence of engineering psychology*. Santa Monica, California: Human Factors and Ergonomics Society.

Rosenberg, M. J. (1969). The conditions and consequences of evaluation apprehension. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 279-349). New York: Academic Press.

Ruffell Smith, H. P. (1979). *A simulator study of the interaction of pilot workload with errors, vigilance and decisions* (No. TM-78482). Moffett Field, CA: NASA-Ames Research Center.

Sagan, S. D. (1993). *The limits of safety: Organizations, accidents and nuclear weapons*. Princeton, New Jersey: Princeton Press.

Salas, E., Burke, S. C., Bowers, C. A., & Wilson, K. A. (2001). Team training in the skies: Does Crew Resource Management (CRM) training work? *Human Factors, 43*(4), 641-674.

Sellen, A. J. (1994). Detection of everyday errors. *Applied Psychology: An International Review, 43*(4), 475-498.

Senders, J. W., & Moray, N. P. (1991). *Human error: Cause, prediction, and reduction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Shappell, S. A., & Wiegmann, D. A. (1997). A human error approach to accident investigation: The taxonomy of unsafe operations. International Journal of Aviation Psychology, 7, 269-291.

Simpson, C. W., & Prusak, L. (1995). Troubles with information overload: Moving from quantity to quality in information provision. International Journal of Information Management, 15(6), 413-425.

Singleton, R. A., Straits, B. C., & Straits, M. M. (1993). Approaches to social research (2nd ed.). New York: Oxford University Press.

Snow, D. A., Benford, R. D., & Anderson, L. (1986). Fieldwork roles and information yield: A comparison of alternative settings and roles. *Urban Life, 14*, 377-408.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers, 31*(1), 137-149.

Stout, R. J., Salas, E., & Carson, R. (1994). Individual task proficiency and team process: What's important for team functioning? *Military Psychology, 6*, 177-192.

Tamuz, M. (2001). Learning disabilities for regulators: The perils of organizational learning in the air transportation industry. *Administration & Society, 33*(3), 276-302.

Tesmer, B. A. (2002). LOSA programme stimulates change that improves safety in line operations. *ICAO Journal, 57*(4), 13, 29.

Thomas, M. J. W. (2004). Predictors of threat and error management: Identification of core nontechnical skills and implications for training systems design. *International Journal of Aviation Psychology, 14*(2), 207-231.

Turner, B. (1978). *Man-made disasters*. London: Wykeham.

Turner, B., & Pidgeon, N. (1997). *Man-made disasters*. London: Butterworth Heinemann.

Weick, K. E. (1968). Systematic observational methods. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., pp. 357-451). Reading, MA: Addison-Wesley.

Weick, K. E. (1985). Systematic observational methods. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., pp. 567-634). New York: Random House.

Weick, K. E. (1987). Organizational culture as a source of high reliability. California Management Review, 2, 112-127.

Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the unexpected*. San Francisco: Jossey-Bass.

Weiner, E. L. (1993). CRM in high-tech cockpits. In E. L. Weiner, B. G. Kanki & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 199-226). New York: Academic Press.

Weiner, E. L., Kanki, B. G., & Helmreich, R. L. (Eds.) (1993). *Cockpit resource management*. New York: Academic Press.

Wells, A. T. (2001). *Commercial aviation safety*. New York: McGraw-Hill.

Wiegmann, D. A., & Shappell, S. A. (2002). Human error perspectives in aviation. *International Journal of Aviation Psychology, 11*(4), 341-357.

Wiegmann, D. A., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis*. Aldershot: Ashgate.

Woods, D. D., Johannesen, L. J., Cook, R. I., & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers, and hindsight*. Wright-Patterson Air Force Base, OH: Crew Systems Ergonomics Information Analysis Center.

Yarrow, M. R., & Waxler, C. Z. (1979). Observing interaction: A confrontation with methodology. In R. B. Cairns (Ed.), *The analysis of social interactions: Methods, issues and illustrations* (pp. 119-162). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Zapf, D., Maier, G. W., Rappensperger, G., & Irmer, C. (1994). Error detection, task characteristics, and some consequences for software design. *Applied Psychology: An International Review, 43*(4), 499-520.

Zapf, D., & Reason, J. T. (1994). Introduction: Human errors and error handling. *Applied Psychology: An International Review, 43*(4), 427-432.

Vita

James Klinect was born in Lake Jackson, Texas on May 21st, 1972, the son of Glenn Carl and Anna Lee Klinect. He attended Brazoswood High School from 1986 to 1990 and upon graduation, enrolled in Brazosport Junior College for two semesters. In September 1991, he transferred to The University of Texas at Austin completing a Bachelor of Arts, Cum Laude in December 1996. During the following nine months, James entered the business world and worked at Dow Chemical as a mechanical engineering contractor. In September 1997, he entered the Graduate School of The University of Texas at Austin working as a graduate researcher with Dr. Robert Helmreich.

Permanent address: 312 Oak Drive, Lake Jackson, Texas 77566

This dissertation was typed by the author.