

Institute of Transportation Studies
University of California at Berkeley

Understanding and Evaluating the Federal Aviation Administration Safety Oversight System

**Mark Hansen, Carolyn McAndrews, Emily Berkeley, Joana Gribko, David
Berkey, Shahab Hasan**

RESEARCH REPORT

July 2006

The National Center of Excellence



For Aviation Operations Research

Preface and Acknowledgements

This report documents research undertaken by the National Center of Excellence for Aviation Operations Research and funded by Federal Aviation Administration contract DTFA01-01-C-00029. In its current form, this document has not been reviewed by the Federal Aviation Administration (FAA). Any opinions expressed herein do not necessarily reflect those of the FAA or the U.S. Department of Transportation.

Throughout this project, the authors have appreciated the support of John Lapointe, Jennelle Derrickson, and Kathy Fazen of the FAA.

Table of Contents

Preface and Acknowledgements	1
Table of Contents	2
List of Tables	4
List of Figures	4
List of Abbreviations	5
Executive Summary	7
1. Introduction.....	7
2. Understanding of the Safety Oversight System.....	13
2.1 Elements of Safety Oversight	14
2.1.1 Air Carrier Certification and Certificate Management	17
2.1.2 Monitoring	17
2.1.3 Enforcement.....	26
2.1.4 Supporting Systems and Databases	26
2.1.5 Regulations and Standards.....	28
2.2 National Program Guidelines and Non-ATOS Surveillance	28
2.2.1 History of NPG	28
2.2.2 How the NPG Surveillance Program Works	30
2.3 Air Transportation Oversight System (ATOS).....	37
2.3.1 What ATOS Generates	41
2.3.2 ATOS Inspections.....	42
2.4 Practitioner and Air Carrier Perspectives on Safety Oversight	44
2.4.1 Standardization of Safety Oversight	45
2.4.2 Role of Trust in Surveillance	46
2.4.3 Improving Safety Oversight.....	47
2.4.4. Training.....	48
3. External Evaluation of the Aviation Safety Oversight System.....	50
3.1 Background.....	50
3.2 Evaluation Methodology.....	52
3.3 Frequency of Evaluation.....	53
3.4 Focus of Evaluation	53

3.5 Indicators of Effectiveness.....	55
4. FAA Safety Oversight Evaluation Challenge	60
4.1 Background and motivation.....	60
4.2 Why Evaluate the Oversight System?.....	61
4.3 Challenges.....	64
4.3.1 Experimentation Is Infeasible	64
4.3.2 Accident Numbers Are Small	65
4.3.3 Safety Production Is Not Well Understood	67
4.3.4 Inspection Results Are Difficult to Interpret	68
4.4 Experience in Other Sectors.....	68
4.5 Lessons for a Safety Oversight Evaluation Methodology	72
4.6 FAA Safety Oversight Evaluation and Monitoring	73
5. Generic Methodology to Evaluate the Effectiveness and Efficiency of System-Based Safety Oversight.....	74
5.1 Preliminary Conceptual Design of an Evaluation Methodology	74
5.2 How the Preliminary Conceptual Design Was Developed.....	75
5.3 Description of the Oversight System Evaluation Methodology	76
5.4 Features of the Oversight System Evaluation System	77
5.4.1 Framework.....	77
5.4.2 Analysis Tool.....	77
5.4.3 Interpretation and Feedback.....	78
5.4.4 Primary Data Collection	80
5.4.5 Analysis Methods.....	80
5.5 Example Analyses.....	84
5.5.1 Introduction.....	84
5.5.2 AIDS Database.....	84
5.5.3 Dataset.....	86
5.5.4 Analysis.....	87
6. Summary of Findings.....	93
REFERENCES	96

List of Tables

Table 1: Sources of Information about Safety and Security Violations and Enforcement Cases	19
Table 2: Summary of Methods Used by GAO to Evaluate the FAA Safety Oversight System.....	56
Table 3: Indicators of Effectiveness Used by GAO to Evaluate the Performance of the FAA Safety Oversight System.....	59
Table 4: Sample Criteria to Evaluate the Safety Oversight System	81
Table 5: Inputs to Evaluation Criteria.....	82
Table 6: Accidents, Flights, and Accident Rates, Major and National Airlines, by Period	91
Table 7: Incidents, Flights, and Accident Rates, Major and National Airlines, by Period	91

List of Figures

Figure 1: The Federal Aviation Administration Oversight Program	14
Figure 2: Organization of the Safety Oversight System's Elements	16
Figure 3: Number of GAO Evaluations of the FAA Safety Oversight System Published per Year, 1970-2005	54
Figure 4: Cumulative Number of GAO Evaluations of the FAA Safety Oversight System, 1970-2005	54
Figure 5: The OSEM Concept	74
Figure 6: Preliminary Conceptual Design.....	76
Figure 7: Components of the Oversight System Evaluation Methodology	79
Figure 8: Number of Airlines in Data Set, by Category and Month.....	88
Figure 9: Trends in Accidents and Accident Rates, Major and National Carriers, 1990-2004.....	89
Figure 10: Trends in Incidents and Incident Rates, Major and National Carriers, 1990-2004.....	90

List of Abbreviations

AFS	Aviation Flight Standards Service
ATO-	Air Traffic Organization Operations Planning
P/R&D	Research and Development
AIDS	Accident/Incident Data System
ASAP	Aviation Safety Action Program
ASAS	Aviation Safety Analysis System
ASI	aviation safety inspector
ASRP	Aviation Safety Reporting Program
ASRS	Aviation Safety Reporting System
ATO	Air Traffic Organization
ATOS	Air Transportation Oversight System
CAA	Civil Aviation Authority
CASS	Continuing Analysis and Surveillance System
CFR	Code of Federal Regulation
CMT	certificate management team
CSET	Certification Standardization and Evaluation Team
DCT	data collection tool
DER	designated engineering representative
DOT	Department of Transportation
EIS	Enforcement Information System
EPI	element performance inspection
FAA	Federal Aviation Administration
FAR	Federal Aviation Regulation
FBI	Federal Bureau of Investigation
FOQA	Flight Operations Quality Assurance
FRA	Federal Railroad Administration
GA	general aviation
GAO	Government Accountability Office
IASA	International Aviation Safety Assessment
ICAO	International Civil Aviation Organization
IG	Inspector General
NASA	National Aeronautics and Space Administration
NASIP	National Aviation Safety Inspection Program
NATI	National Air Transportation Inspection National Flight Standards Work Program
NPG	Guidelines
NPR	National Performance Review
NRC	Nuclear Regulatory Commission
NTSB	National Transportation Safety Board
OMB	Office of Management and Budget
ORA	operation research analyst
OSHA	Organizational Safety and Health Administration
OTA	Office of Technology Assessment
PART	Program Assessment Rating Tool
PASS	Professional Airways Systems Specialists
PBO	Performance Based Organization
PCD	preliminary conceptual design
PI	Principal Inspector

PTRS	Program Tracking Reporting System
RMDS	Risk Management Decision Support
SAGA	System Safety Approach for General Aviation
SAI	safety attribute inspection
SASO	System Approach to Safety Oversight
SEP	Surveillance and Evaluation Program
SME	subject matter expert
SMS	safety management system
SPAS	Safety Performance Analysis System
SRR	specific regulatory requirement
TSA	Transportation Security Administration
VIS	Vital Information Subsystem
VDRP	Voluntary Disclosure Reporting Program
WPMS	Work Program Management Subsystem

Executive Summary

This report presents an overview of the Federal Aviation Administration's (FAA) current safety oversight practices performed on commercial carriers operating under Title 14, Code of Federal Regulations, Part 121, *Operating Requirements: Domestic, Flag, and Supplemental Operations*. Also presented is a preliminary methodology for evaluating the effectiveness, efficiency, and sustainability of FAA safety oversight on these carriers.

The FAA is responsible for protecting the safety of the U.S. airspace system. Accordingly, it has established a safety oversight system to help fulfill this responsibility. Using this system, the FAA inspects all aspects of commercial carrier operations, including aircraft, training, equipment, maintenance, and operations for purposes of assessing compliance with mandatory safety requirements.

At its core, an evaluation methodology for the FAA safety oversight system asks what the safety oversight system contributes to the overall safety record of the aviation industry. It is difficult to establish exactly what the safety oversight system contributes to this safety record using statistical methods because the aviation industry has powerful incentives to be reliable and safe. Understanding and documenting the FAA safety oversight system, and the mechanisms through which it affects the behavior of the aviation industry, is key to developing a methodology for evaluating its effectiveness.

The FAA National Program Guidelines (NPG), initiated in 1985, and the Air Transportation Oversight System (ATOS), established in 1998, are the principal systems currently in use by the FAA for overseeing the safety of commercial air carriers. While both systems have positive attributes and some weaknesses, ATOS is considered to be more comprehensive, and accordingly will eventually be applied to all commercial carriers and possibly general aviation. ATOS also embodies one of the FAA's first steps in implementing a system safety approach to safety oversight rather than focusing strictly on compliance-based oversight.

Information obtained from interviews with key FAA personnel, subject matter experts, and secondary sources, including Government Accountability Office assessments of FAA inspection activities, was analyzed and used to develop a preliminary methodology for evaluating the FAA safety oversight system. The principal characteristics of this system include a top-level framework, an analysis tool, and a

review process for stakeholders. The framework poses the overall questions regarding the efficiency, effectiveness, and sustainability of the oversight system. The analysis tool performs processes such as monitoring and evaluating incoming data, preparing pre-defined analysis reports or charts, performing comparative analyses, and identifying opportunities for improvement that are evident based on data. The review process includes stakeholders who review evaluation reports and other output from the analysis branch of the methodology. Our preliminary methodology requires multiple inputs including data regarding surveillance and other oversight activities, communication, trends in the aviation environment, and internal and external audit reports. In turn, the methodology would produce metrics such as the responsiveness of air carriers to corrective and preventive actions, the effect of safety oversight on safety precursors, inspection output (including safety recommendations), and inspector workload and readiness.

The users of this methodology may be both inspectors and their managers; operations research analysts who help identify targeted inspection opportunities or effective sampling inspection techniques; as well as FAA headquarters staff and policy makers who require an overarching view of health of the entire aviation safety system.

1. Introduction

The Federal Aviation Administration (FAA) safety oversight system promotes flight safety by “assuring airworthiness of aircraft, competency of airmen, and adequacy of flight procedures and air operations.”¹ It performs this mission by setting and enforcing regulations and standards, certifying individuals, organizations, and equipment, monitoring airlines and aviation-related activities, and investigating accidents, incidents, and unsafe conditions. The FAA Aviation Flight Standards Service (AFS) carries out these activities with a workforce of 4,600 employees located in FAA headquarters, eight regional offices around the United States, and more than 100 field offices around the U.S. and the world. The budget of Aviation Safety Oversight was about \$870 million in 2004.²

The philosophy that guides the FAA oversight system has evolved over the last two decades. Until the 1970s, the system focused mainly on developing rules and regulations, and implementing them through the processes of certification, inspection, and enforcement. The latter includes a wide range of responses, from informal, on-the-spot corrections, to formal administrative actions, to legal penalties such as fines and certificate actions.³ In short, the goal of the system was to “convert safety standards from concept to reality” and “to protect the many conscientious pilots from the irresponsible few.”⁴

The mantra of the present safety oversight system, conspicuously displayed on the AFS seal, is “System Safety.” This is a broader, and somewhat more elusive concept than the pursuit of safety through regulatory compliance. Two published definitions of system safety are “the application of technical and managerial skills to identify, analyze,

¹ Donald Kemp, “The Federal Aviation Administration’s Air Safety Program,” *Government-Industry System Safety Conference*, NASA Safety Office, National Aeronautics and Space Administration, May 1-3, 1968.

² Office of Management and Budget, Budget of the United States Government, Fiscal Year 2005, Appendix, Page 767. <http://www.whitehouse.gov/omb/budget/fy2005/appendix.html>. Website visited July 31, 2004.

³ U.S. General Accounting Office, “Aviation Safety: Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks,” GAO/RCED-98-6, February 1998.

⁴ Robert Burkhardt, *The Federal Aviation Administration*, Frederick A. Praeger: New York, 1967.

assess and control hazards and risks” or, more succinctly, “organized common sense.”^{5, 6} Major elements of the system safety philosophy include the allocation of oversight resources on the basis of risk, the need for airlines to have operating systems in place to identify and mitigate hazards and risks, and the promotion of safety culture in which workers throughout an organization strive to increase safety.⁷

AFS has initiated the System Approach to Safety Oversight (SASO) program, whose aim is to “transform the AFS and the aviation industry to a national standard for safety risk management.”⁸ The SASO program will include improved decision support and training for inspectors, incentives for industry to implement safety management programs, and enhanced systems for acquiring, storing, sharing, and analyzing safety data.

To support AFS’s system safety initiatives, and in particular the SASO program, the FAA Air Traffic Organization Operations Planning Research and Development (ATO-P/R&D) has undertaken a research and development program entitled Risk Management Decision Support (RMDS). The goal of the program is “to improve safety by making AFS oversight more systematic, effective, efficient, and targeted to deal with identified risks by developing risk management and decision support capabilities.”⁹ RMDS research activities currently focus on commercial aviation and they address six requirements including the development of an aviation safety evaluation construct, risk analysis methods, safety performance measures, diagnostic procedures, decision support tools for aviation safety evaluation, and, finally, the development of a methodology for evaluating the safety oversight system.

⁵ U.S. General Accounting Office. “Aviation Safety: FAA’s New Inspection System Offers Promise, but Problems Need to Be Addressed. GAO/RCED-99-183. June 1999.

⁶ Lederer, Jerome. “National Aeronautics and Space Administration.” *Government-Industry System Safety Conference*. NASA Safety Office. National Aeronautics and Space Administration. May 1-3, 1968.

⁷ U.S. General Accounting Office. “Aviation Safety: Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks.” GAO/RCED-98-6. February 1998.

⁸ U.S. Federal Aviation Administration. “AFS-900, System Approach to Safety Oversight—Mission Need Statement.” March 2001.

⁹ Kathleen M. Fazen, “Risk Management Decision Support – RPD 676 Commercial Aviation Requirements Program Review,” Risk Management Decision Support Meeting, Atlantic City, New Jersey. July 13, 2004.

In this report, we present ATO-P/R&D-sponsored research on the last of these areas, a methodology for evaluating the safety oversight system. While the other RMDS tasks are intended to yield concepts, methods, and tools that will be incorporated into the safety oversight system itself, our task is geared to evaluating the health and performance of a generic aviation safety oversight system. In a sense, we want to design a methodology for overseeing oversight. The evaluation system is expected to be independent of the safety oversight system itself so that it can track improvement, or degradation, as the safety oversight system evolves. It is expected to identify situations in which more oversight is needed and in which oversight can be curtailed. Above all, the evaluation system must inform the easily posed but difficult question: what does the safety oversight system contribute to aviation safety?

While there is no shortage of methodological research geared toward assessing the performance and results of government programs, the application of generic methodology is of little value—indeed, can cause considerable harm—without a detailed understanding of the program under study. Thus, a large portion of the effort documented in this report has gone to understanding how the current oversight system works. In addition to reviewing a wide range of published and unpublished government documents, we have interviewed many individuals who are presently or have recently worked in the system, including inspectors, managers, analysts, and key SASO staff. We have also consulted with current and former airline employees concerned with safety, including those who interfaced with the government overseers. Finally, we have conducted a comprehensive review of previous oversight system evaluations produced by the Government Accountability Office (GAO) and the Department of Transportation Office of Inspector General (DOT IG).

The aviation safety oversight system is rarely the subject of headline news, but there is abundant controversy among its stakeholders, both internal and external, about what it does, how well it works, how it should evolve, and how it should be evaluated. The individuals and sources we consulted thus presented a wide range of viewpoints. Our task as researchers was to analyze and synthesize them. Thus, unless otherwise stated, the observations in this report represent our own conclusions drawn from the various inputs described above.

Chapter 2 of this report discusses how the safety oversight system works, its activities, and the information that it produces. This understanding of the system is the basis for a discussion of the evaluation of safety oversight systems. Readers who are already familiar with the FAA's safety oversight system may want to skip Chapter 2. Chapter 3 introduces information and findings about external evaluations of the FAA safety oversight system. External evaluations provide insight about evaluation methods and the criteria by which an evaluator could judge the safety oversight systems. Chapter 4 discusses how safety oversight systems, particularly those with a systems approach, pose challenges for developing an evaluation methodology. Chapter 5 presents our approach to developing an evaluation methodology for the FAA safety oversight system. Chapter 5 also presents a preliminary conceptual design of this evaluation methodology with examples of analyses. Finally, Chapter 6 presents key findings of our research.

2. Understanding of the Safety Oversight System

To successfully and usefully evaluate the Federal Aviation Administration (FAA) safety oversight system, we must begin with a sound understanding of this system. Our understanding is based on extensive reviews of documents from and about the FAA safety oversight system and through numerous interviews with people who carry out the management and activities of the safety oversight system. To gain a more general understanding of safety oversight systems, we also used information about safety oversight systems from other industries, including nuclear power and railroads.

In this report we focus on the two principal FAA safety oversight systems that apply to large, commercial carriers operated under Federal Aviation Regulation Part 121. Though this work concentrates on the FAA's safety oversight activities in the United States with domestic air carriers, we also discuss briefly how the FAA safety oversight system relates to international programs that oversee the safety of international air travel and foreign air carriers.

The Figure 1 illustrates some of the key components of the FAA oversight system for commercial air carriers.

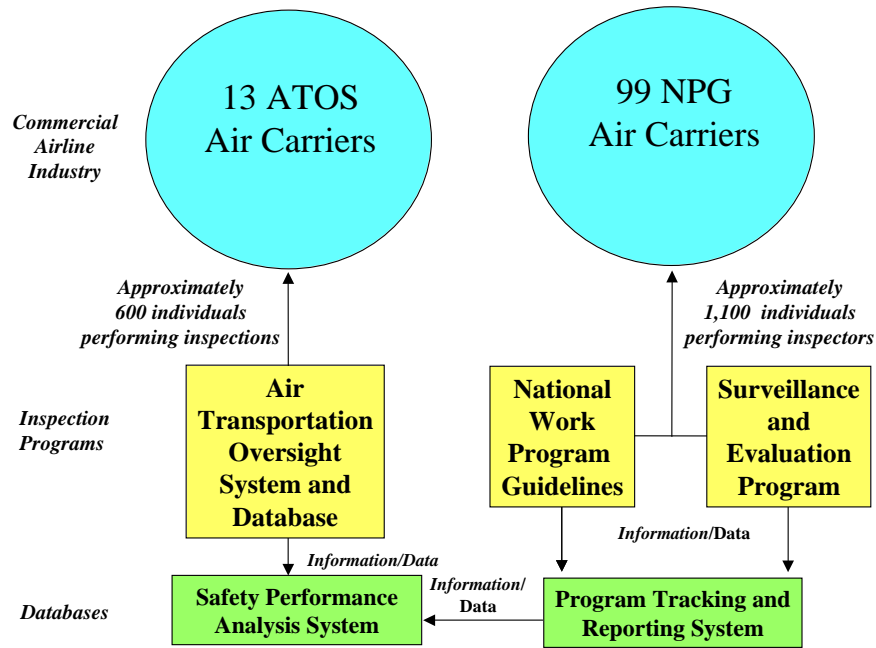
Two principal oversight systems are:

- National Program Guidelines (NPG)
- Air Transportation Oversight System (ATOS)

Of these two oversight systems, NPG has been in existence longer; in the late 1990s important surveillance elements of the NPG were re-engineered, resulting in the creation of the Air Transportation Oversight System (ATOS). In 1998, ATOS was implemented for the 10 largest passenger airlines. NPG remains operational and is applied to non-ATOS air carriers operating under 14 CFR Part 121. This report discusses both safety oversight systems.

Throughout this section, we maintain a systems perspective on safety oversight and consider both the elements of the safety oversight system and the relationships among them.

Figure 1: Federal Aviation Administration Oversight Program for Air Carriers Operating Under 14 CFR Part 121



Note: The number of ATOS carriers will increase as more carriers transition to ATOS.

2.1 Elements of Safety Oversight

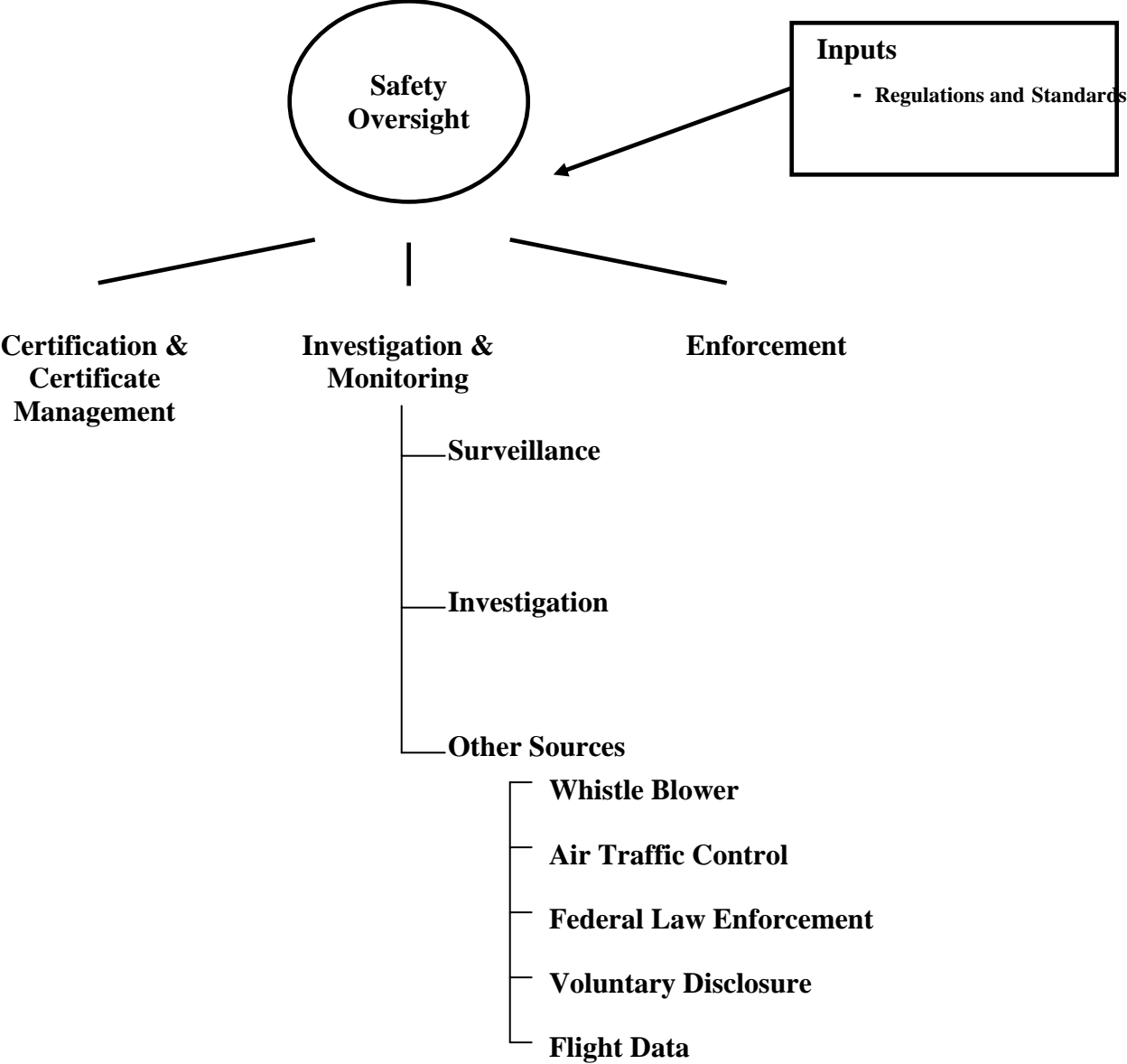
The FAA safety oversight system is a process by which the FAA sets safety standards, licenses pilots, certifies air carriers and their aircraft, and ensures compliance with these standards through surveillance and enforcement. Within the FAA, AFS is the primary steward of the safety oversight system and the certification and regulation of the operators.¹⁰ The FAA updates its safety regulations through formal rulemaking processes that use information collected through the safety oversight system. The safety oversight system also investigates incidents and accidents, as well as industry activities that may affect safety (e.g., suspected unapproved parts). Though many of the processes within the safety oversight system are formal, such as rulemaking, many informal processes, such as the transmission of information between individuals in the system, work to carry out the mission of safety oversight.

¹⁰ U.S. Federal Aviation Administration, *Inspector Handbook Volume I: Chapter 2. The Federal Aviation Administration and Flight Standards: History and Organization, Section 1*, 8400.10 CHG 1, December 20, 1994.

One way to understand the safety oversight system is to construct a general framework describing the relationships among its elements. Figure 2 is a diagram of one such organizing framework. The safety oversight system, including NPG and ATOS, has three main programmatic functions: *certification, enforcement, and monitoring*. These programmatic functions are performed by AFS and by partners in safety oversight, including the National Transportation Safety Board (NTSB), air carriers, federal law enforcement, and other entities.

The following sections describe the programmatic functions of the safety oversight system in more detail.

Figure 2: Organization of the Safety Oversight System's Elements



2.1.1 Air Carrier Certification and Certificate Management

The cumulative knowledge of safe designs and practices that is formalized through the certification process makes a significant contribution to the effectiveness of the safety oversight system. The standards by which a carrier is certified are the baseline for its future operations. These standards apply to aircraft, routes, airports, maintenance and operations plans, training, and other elements of service.

As air carriers evolve their certificates change too.. Through a process called certificate management, air carriers request certificate changes, and FAA approves those changes and modifies the certificate. Certificates change when an air carrier adds new aircraft, new service to additional airports, new routes, special operations, new safety systems, and other elements that are captured in a certificate. Inspectors are responsible for validating that a new aircraft, for example, has the equipment required per the air carrier's certificate.

In 1997, the FAA created a special program to certify new entrant air carriers and perform certificate management tasks, the Certification Standardization and Evaluation Team (CSET) program.¹¹ The CSET process was designed to encourage consistent application and interpretation of regulations for each prospective certificate holder. This was accomplished with teams of dedicated CSET inspectors working with regional offices. In February of 2005, CSET became part of ATOS as the Certification Section. With this new organization, CSET inspectors now work with Certificate Management Offices. The certification process for a new entrant carrier can take approximately 18 months. In order to facilitate and expedite this process, the FAA allows private companies to help carriers through the certification process.¹²

2.1.2 Monitoring

Monitoring is the process of *detecting* safe or unsafe conditions and *providing* that information to the people in the safety oversight system. Monitoring information is used in many ways, including the enforcement of Federal Aviation Regulations (FARs), retargeting inspection activities, and informing new safety standards and regulations. The ability of the

¹¹ U.S. Federal Aviation Administration, "FAA Order Number 1100.159," Feb. 7, 1997. Initiated by AFS-120. Guy S. Gardner, Associate Administrator for Regulation and Certification

¹² U.S. Federal Aviation Administration, "Certification Process Study: An Evaluation of Selected Aircraft Certification, Operations, and Maintenance Processes," April 2, 2002.

safety oversight system to use monitoring information is a key element to its ongoing success. Monitoring is carried out in a diversity of ways by multiple programs within the safety oversight system, including surveillance, investigation, and the support of outside programs that provide additional monitoring information to the safety oversight system.

In its broadest sense, monitoring provides current information about safety to members of the safety oversight system who can correct noncompliance or safety issues. AFS is not alone in its monitoring role as there are several other sources of information that complement the activities of the safety oversight system. These other sources include, for example, the programs and employees of air carriers, security programs operated by other parts of government, the National Transportation Safety Board, the U.S. Department of Defense, and other parts of the FAA such as Air Traffic Control that have access to information about safety. AFS's role includes supporting its partners' monitoring efforts and using information from these sources to carry out its own safety oversight. For example, the FAA safety oversight program has developed self-disclosure and whistle-blower programs as alternate ways in which crucial information about safety can enter the safety oversight system.

The Government Accountability Office (GAO) has found that both the FAA safety oversight system and its partnership programs contribute significantly to the safety of the aviation system. Table 1 summarizes selected findings from GAO reports. Although the GAO's findings relate to a broader set of air carriers and types of violations and enforcement actions, its analysis demonstrates clearly the important role of partnerships in system safety.

The GAO reported that "FAA frequently became aware of violations that led to emergency revocation orders during accident investigations or investigations performed as a result of tips from employees, competitors, or customers..."¹³ The GAO's study of air taxis found that 46 percent of discoveries of violations were made by FAA inspectors conducting routine or special inspections; 37 percent of these discoveries were made and reported by tips from other sources; and that in 17 percent of the cases, accident investigations were the source of the discovery.¹⁴ There is evidence that this pattern applies to large commercial carriers operating under Part 121. In a separate report for Part 121 and Part 135 carriers, the GAO reported that Air

¹³ U.S. General Accounting Office, "Emergency Revocation Orders of Air Carrier Certificates," GAO/RCED-92-10, October 17, 1991.

¹⁴ *Ibid.*

Traffic Control, local and state government, self-disclosure, public complaint, and other non-FAA sources were the basis for 39 percent of all safety and security enforcement cases identified in the study.¹⁵

Table 1: Sources of Information about Safety and Security Violations and Enforcement Cases

<i>Part 135 Carriers</i>	
Source of Violation Discovery	Contribution to Violation Discovery
FAA Safety Oversight	46%
Tips from Employees, Others	37%
Accidents	17%
<i>Part 135 and Part 121 Carriers</i>	
Source of Enforcement Cases	Contribution to Enforcement Cases
FAA Safety Oversight	61%
Other Sources:	39%
Air Traffic Control	
Local Government	
State Government	
Self-Disclosure	
Public Complaint	
Other, Non-FAA Sources	

Source: U.S. General Accounting Office. (1991). "Emergency Revocation Orders of Air Carrier Certificates." GAO/RCED-92-10.; U.S. General Accounting Office, "Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks," GAO/RCED-98-6, February 27, 1998.

2.1.2.1 Surveillance

Surveillance is one of the most important functions of the safety oversight system. According to the FAA's *Inspector Handbook* for operations inspectors, "Surveillance programs provide the FAA with a method for continual evaluation of operator compliance with the FARs and safe operating practices."¹⁶ Over 3,000 aviation safety inspectors perform surveillance of

¹⁵ U.S. General Accounting Office, "Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks," GAO/RCED-98-6, February 27, 1998.

¹⁶ U.S. Federal Aviation Administration, *Inspector Handbook Volume 6: Surveillance, Chapter 1. General Policies and Procedures.* 8400.10 CHG 3, July 31, 1990.

commercial air carriers and general aviation. The primary elements of the surveillance system include planning and executing inspections, determining inspection requirements, and evaluating inspection results.

Inspections are the main component of the surveillance system and are categorized by the inspectors' specializations: maintenance, operations, and avionics. Inspection records and findings are collected in databases—the Program Tracking and Reporting Subsystem (PTRS) for NPG inspections and the Air Transportation Oversight System Database (ATOS Database) for ATOS inspections. Each year, the FAA performs hundreds of thousands of routine NPG and ATOS inspections, as well as special inspections. According to the FAA's *Inspector Handbook* for operations inspectors working in the NPG system, an inspection must have the following five characteristics:

- (1) A title and a Program Tracking and Reporting Subsystem (PTRS) code;
- (2) A beginning and an end;
- (3) Defined inspection procedures;
- (4) Specific objectives; and
- (5) A finding: either positive, negative, or both.¹⁷

2.1.2.2 Investigation

Investigation is similar to surveillance because it uses information produced through observation and discovery to prevent future accidents, incidents, and unsafe conditions. Investigation is distinct from ongoing monitoring because it is the process of discovery that happens *after* an accident or incident, or in response to a violation or possible criminal action.¹⁸ Surveillance could be considered an *interactive* approach to preventing unsafe conditions and investigation would be a *reactive* approach to improving safety.

FAA employees, including aviation safety inspectors, and NTSB investigators cooperate during an investigation, contributing to accident prevention with the knowledge they produce through the investigation. The FAA's role in accident investigations is described in FAA Order

¹⁷ *Ibid.*

¹⁸ See U.S. Federal Aviation Administration, *Inspector Handbook Volume II: Certification, Chapter 211. Conduct and Accident Investigation*, 8300.10 CHG 12, December 14, 1999. See also, U.S. Federal Aviation Administration, *Inspector Handbook Volume 7: Investigations, Chapter 1. Accident and Incident Investigation and Reporting*, 8400.10 CHG 10, December 20, 1994.

8020.11B. The Order states that the FAA’s role in accident investigations is to prevent future accidents by ensuring that all facts are gathered during an investigation, and to determine whether factors that are critical to the FAA’s mission to ensure safety in the public air transport system – the FARs, the certified aircraft, airmen, facilities, and other essential elements of FAA’s operations and policies – were involved in the accident and in what way. The FAA is also responsible for determining whether violations of FARs contributed to the accident. The Flight Standards offices in the regions and districts play a role in the investigation as well. Flight Standards employees, including aviation safety inspectors, participate in investigation when the accident or incident occurred in their geographic region.¹⁹

FAA inspectors also participate in the investigation of criminal activity, and other types of investigations. To facilitate these, FAA inspectors work with the Federal Bureau of Investigation (FBI), as well as the Inspector General for the Department of Transportation.

2.1.2.3 Information from Other Sources

Another important function of the safety oversight system is to support partners’ surveillance programs, including, but not limited to whistle-blower programs, flight data collection, and voluntary self disclosure programs. The GAO has identified these programs as being major contributors of information to the safety oversight system.²⁰ In general, information protected under voluntary self-disclosure agreements is not available for all FAA inspectors. Instead, sensitive data are kept at the local office or reviewed by committees who recommend further action.

¹⁹ U.S. Federal Aviation Administration, “Aircraft Accident and Incident Notification, Investigation and Reporting,” FAA Order 8020.11B, August 16, 2000.

²⁰ See, U.S. General Accounting Office, “Emergency Revocation Orders of Air Carrier Certificates,” GAO/RCED-92-10, October 17, 1991. See also, U.S. General Accounting Office, “Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks,” GAO/RCED-98-6, February 27, 1998.

Whistle-Blower Program

The FAA whistle-blower program for air carriers allows employees to report violations without retaliation by air carriers. Information from whistle-blowers is one way in which the surveillance system can detect non-compliance or even unsafe conditions.²¹

If an air carrier employee offers a protected complaint, FAA inspectors research the complaint, keep records of the investigation in PTRS, and write a report. This information is then shared with the inspector's regional division manager.²²

Aviation Safety Reporting Program (ASRP)

The Aviation Safety Reporting Program (ASRP) was created in 1975.²³ Reports "of unsafe occurrences and hazardous situations are voluntarily submitted by pilots, air traffic controllers, and others."²⁴ ASRP is similar to the whistleblower protection program because it encourages people to volunteer safety information that they may otherwise not share because of a fear of retaliation or the threat of enforcement action. Also like the whistleblower program, ASRP protects the identity of the person who submits information.

Analysts at the National Aeronautics and Space Administration (NASA) screen the data collected through the program. The data are then de-identified and collected in a database, the Aviation Safety Reporting System (ASRS), which NASA administers. This database is available to the public.²⁵ ASRS information is used to:

1. "Identify hazards and safety discrepancies in the National Airspace System."

²¹ U.S. Federal Aviation Administration, *Inspector Handbook Volume 1: General Concepts, Direction, Guidance, and Definitions, Chapter 6. Whistleblower Protection Program (Air Carrier)*, 8400.10 CHG 19, April 2, 2003.

²² *Ibid.*

²³ U.S. Government Accountability Office, "FAA's Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs' Performance," Statement by Gerald L. Dillingham, GAO-06-266T, November 17, 2005.

²⁴ National Aviation Safety Data Analysis Center, www.nasdac.faa.gov, website visited December 19, 2005. See also <http://asrs.arc.nasa.gov/main.htm>, website visited March 18, 2006.

²⁵ U.S. Government Accountability Office, "FAA's Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs' Performance," Statement by Gerald L. Dillingham, GAO-06-266T, November 17, 2005.

2. “Formulate policy and to strengthen the foundation of aviation human factors safety research.”²⁶

Voluntary Disclosure Reporting Program (VDRP)

The Voluntary Disclosure Reporting Program (VDRP) began in 1990. This program applies to air carriers, repair stations, and aircraft and parts manufacturers. Through this program, employees and operators can report an eligible violation to the FAA within 24 hours without enforcement action. The individual or air carrier must also report how the violation was discovered and what corrective actions were taken. FAA inspectors participate in the program by approving the disclosure and reports.²⁷

Flight Operations Quality Assurance (FOQA)

The Flight Operations Quality Assurance (FOQA) program began in 1995.²⁸ The objective of the FOQA program is to identify and correct safety deficiencies in flight operations using information about trends and safety risks. The FAA can use this information to identify trends in the aviation system.²⁹ Currently, 13 air carriers participate in the FOQA program by equipping aircraft with data collection devices that monitor the aircraft engines, flight paths, and

²⁶ National Aviation Safety Data Analysis Center, www.nasdac.faa.gov, website visited December 19, 2005

²⁷ U.S. Government Accountability Office, “FAA’s Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs’ Performance,” Statement by Gerald L. Dillingham, GAO-06-266T, November 17, 2005.

²⁸ *Ibid.*

²⁹ U.S. Federal Aviation Administration, *Inspector Handbook Volume 1: General Concepts, Direction, guidance, and Definitions, Chapter 5. Voluntary Safety Programs, Section 2. Flight Operational Quality Assurance (FOQA)*, 8400.10 CHG 30, October 14, 2004.

other variables.^{30,31} The air carriers own the FOQA data and use the data to identify possible safety trends or problems.³²

Aviation Safety Action Program (ASAP)

The Aviation Safety Action Program (ASAP) began in 1997.³³ This program for air carriers and air carrier employees facilitates the voluntary disclosure of safety information. The program also applies to repair stations.³⁴ The objective of ASAP is to prevent future accidents and incidents, educate the members of the aviation system, and analyze safety data.³⁵

The program applies to air carriers that have an agreement with the FAA under which employees can voluntarily disclose eligible safety violations, incidents, or other issues without enforcement action. According to the GAO, 54 air carriers participate in the program and the FAA has received more than 80,000 reports through the program.³⁶

Continuing Analysis and Surveillance System (CASS)

The Continuing Analysis and Surveillance System (CASS) is required for each certificate holder. The CASS is a quality assurance program that air carriers use to monitor and evaluate their maintenance programs. The program is designed to identify and correct any deficiencies in an air carrier's maintenance program. FAA inspectors follow the development of an air carrier's

³⁰ U.S. Government Accountability Office, "FAA's Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs' Performance," Statement by Gerald L. Dillingham, GAO-06-266T, November 17, 2005.

³¹ U.S. Federal Aviation Administration, *Inspector Handbook Volume 1: General Concepts, Direction, guidance, and Definitions, Chapter 5. Voluntary Safety Programs, Section 2. Flight Operational Quality Assurance (FOQA)*, 8400.10 CHG 30, October 14, 2004.

³² U.S. Government Accountability Office, "FAA's Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs' Performance," Statement by Gerald L. Dillingham, GAO-06-266T, November 17, 2005.

³³ *Ibid.*

³⁴ *Ibid.*

³⁵ U.S. Federal Aviation Administration, *Inspector Handbook Volume 1: General Concepts, Direction, Guidance, and Definitions, Chapter 5. Voluntary Safety Programs, Section 1. Aviation Safety Action Program (ASAP)*, 8400.10 CHG 17, February 11, 2003.

³⁶ U.S. Government Accountability Office, "FAA's Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs' Performance," Statement by Gerald L. Dillingham, GAO-06-266T, November 17, 2005.

CASS, attend CASS meetings and presentations at the air carriers, and inspect that the program is in compliance with the regulation.^{37,38}

International Aviation Safety Assessment

The International Aviation Safety Assessment (IASA) is the process by which the FAA ensures that foreign air carriers operating in the United States are regulated by a Civil Aviation Authority (CAA) that is in compliance with the International Civil Aviation Organization (ICAO). The FAA initiated IASA in 1992. The FAA currently has two ratings for CAA's: Category I and Category II. Category I CAA's are those that are in compliance with ICAO standards. Category II CAA's are not in compliance with ICAO standards.³⁹ The minimum international standards for CAA's address, for example, the adequacy of regulatory legislation, advisory documentation, technical data, and staff experience. ICAO standards are set forth in 18 Annexes.

The ICAO standards have traditionally taken a compliance-based approach focusing on compliance with standards, minimum requirements, and the role of the state in policing these standards and requirements. More recently, however, new standards that require aviation organizations to pro-actively manage safety have been added. Specifically, the ICAO standards now include provisions that aircraft operators, air traffic service providers, and airports implement *safety management systems* (SMS). Specifically, states must require that these entities have systems that identify safety hazards, continuously monitor and regularly assess levels of safety, and ensure that necessary remedial actions are taken. A new manual providing guidance on SMS implementation was issued 2005. Like many of the other programs discussed above, the ICAO's SMS requirements are motivated by the idea that they are "the most effective way of responding to the need for effective supervision with a relatively small workforce."⁴⁰ The FY

³⁷ U.S. Federal Aviation Administration, *Inspector Handbook Volume 3: Aircraft and Equipment, Chapter 37. Monitor Continuing Analysis and Surveillance Program/Revision*, 8300.10 CHG 9, August 13, 1993.

³⁸ U.S. Federal Aviation Administration, "Certification Process Study: An Evaluation of Selected Aircraft Certification, Operations, and Maintenance Processes," April 2, 2002.

³⁹ U.S. Federal Register, "Changes to the International Aviation Safety Assessment (IASA) Program," May 25, 2000, Volume 65, Number 102.

⁴⁰ Chérif, Taïb, "Working Together to Achieve Greater Safety," Luncheon Address, Federal Aviation Administration 2nd International Aviation Safety Forum, Oct. 20, 2005.

2006 FAA Aviation Safety Business Plan targets delivery of an SMS standard, for use by Joint Planning and Development Office agencies and AVS, by June 2006.⁴¹

2.1.3 Enforcement

An effective enforcement system complements surveillance by providing a framework for corrective action. Enforcement actions, warnings, civil penalties, and criminal penalties, are methods that reinforce compliance, accident prevention, and the implementation of corrective actions. General knowledge of FAA's enforcement activities helps maintain public confidence in its safety oversight system and air carrier performance. Although many enforcement actions are settled without penalty, and others are reduced in severity, the enforcement of FARs and application of appropriate measures for noncompliance is an important part of the safety oversight system. Data from the enforcement process have also been valuable for GAO evaluations of the safety oversight system.

2.1.4 Supporting Systems and Databases

In addition to the basic programmatic elements of the safety oversight system – monitoring, enforcement, and certification – there are a number of supportive systems or tools that facilitate the operation of these elements of safety oversight. Databases are the primary tools that support these programs. These databases archive inspection information and facilitate its analysis. The following section describes these important systems.

2.1.4.1 Program Tracking and Reporting System (PTRS)

The Program Tracking and Reporting System (PTRS) was implemented in 1990 as part of the Aviation Safety Analysis System (ASAS), a system that integrates, standardizes, and maintains FAA safety databases.^{42,43} It is still used today to track inspections and air carrier

⁴¹ U.S. Federal Aviation Administration, Aviation Safety Business Plan for FY 2006, p. 4.

⁴² U.S. Office of Technology Assessment, "Safe Skies for Tomorrow: Aviation Safety in a Competitive Environment," OTA-SET-381, July 1988.

⁴³ U.S. General Accounting Office, "Problems Persist in FAA's Inspection Program," GAO/RCED-92-14, November 20, 1991. See also, FAA Inspector Handbook.

compliance with FARs. Its purpose is to keep track of inspection data, enabling managers and inspectors to use those data to plan future inspections.

Each safety inspection has a unique PTRS code, as well as codes to classify specific instances of non-compliance or safety problems.

2.1.4.2 Safety Performance Analysis System (SPAS)

Development of the Safety Performance Analysis System (SPAS) was completed in the late 1990s, and it is continually maintained. It serves as a source of information to assess airline safety risk. SPAS provides data from PTRS and Vital Information Subsystem (VIS) databases, as well as seventeen other sources of information and databases.⁴⁴ SPAS is also programmed with its own set of analysis tools that enable users to create indicators and metrics of safety in the aviation system.

Inspectors are responsible for monitoring the SPAS data that relate to their inspection duties. Principal Inspectors are responsible for regularly monitoring and evaluating SPAS data for indicators (referred to as flags) that signify a safety concern or possible non-compliance.⁴⁵ In addition to identifying flags, inspectors can investigate issues by accessing the numerous databases available through SPAS, or through stand-alone applications and systems. The information available to inspectors is relevant to the task of creating an evaluation methodology for two reasons. First, these data could be sources of information used in an evaluation process. Second, information has an important role in systems safety because it leads to a shared awareness among members of the system.

2.1.4.3 Enforcement Information System

The Enforcement Information System (EIS) is the FAA enforcement database. An inspector opens enforcement cases by creating a record for the case in EIS. The record can be closed without any further action, or the record can be augmented with recommendations of an enforcement action. The enforcement actions can be administrative or civil.

⁴⁴ U.S. General Accounting Office, "Data Problems Threaten FAA Strides on Safety Analysis System," GAO/AIMD-95-27, February 8, 1995.

⁴⁵ U.S. Federal Aviation Administration, *Inspector Handbook Volume IX: Technical Staff Administration and Responsibilities, Chapter 1. Inspector Responsibilities, Administration, and Conduct*, 8400.10 CHG 10, December 23, 2004.

2.1.5 Regulations and Standards

Air carriers are required to comply with Federal Aviation Regulations (FARs), Airworthiness Directives, Operations Specifications, Manuals, Department of Transportation hazardous materials regulations, NTSB regulations, and Transportation Security Administration (TSA) regulations. Advisory Circulars are additional standards that help ensure the safety of the aviation system, but they are not mandatory.

The safety oversight system cooperates with other government agencies, including the TSA, the Organizational Safety and Health Administration (OSHA), the NTSB, and law enforcement offices to ensure compliance and safety.

2.2 National Program Guidelines and Non-ATOS Surveillance

Having presented general structure for an aviation safety oversight system, we now consider in detail two specific systems, both of which target large jet passenger carriers. First we consider the older of the two—the National Program Guidelines (NPG).

2.2.1 History of NPG

In 1985, the FAA developed National Program Guidelines—a specified minimum number of operations, maintenance, and avionics inspections that were required to be carried out in each region—to address the need for more standardized inspection planning.⁴⁶ This need for standardization stemmed from the difficulty of sharing situational awareness across the nine regions, 94 domestic field offices, and functional specializations.⁴⁷ The need to provide national guidance in the aviation safety arena has been a long standing goal, starting with the 1926 enactment of the Air Commerce Act, and continuing to the present day.

NPG grew out of an earlier effort to coordinate surveillance across regions known as the Work Program Management System (WPMS), a database network created in the middle 1980s that linked each of the regions and facilitated information sharing, safety analysis, and program

⁴⁶ U.S. Government Accountability Office, “FAA’s Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs’ Performance,” Testimony by Gerald Dillingham, GAO-06-266T, November 17, 2005.

⁴⁷ Different policy divisions focus on safety in general aviation, human factors, system safety and analysis, technical programs, investigation, and other specific programs that relate to carrying out the mission of the organization. Inspectors specialize in operations, maintenance, or avionics.

planning.⁴⁸ Another goal of WPMS was to help the regions and headquarters keep track of planned inspections and their execution. WPMS also collected inspection data and allowed employees to analyze the data for safety insights and for program planning. Unfortunately, problems with the computer network, lack of hardware, poor data quality, and lost data prevented the network from working as it was designed to. A GAO report describing the system also noted that the users “lost confidence” in the system.⁴⁹

In the middle 1980s, the GAO published reports in which it criticized the surveillance elements of the safety oversight system.⁵⁰ These reports focused on FAA’s ability to monitor the regions’ planning and execution of surveillance programs, as well as the outcomes of the inspections. The standardization of inspections was a concern, and Congress and the GAO were interested in the FAA’s methods for planning inspections. Individuals at the regional and field office levels made inspection plans, but there was no national program or working database that provided guidance. These were the issues that information systems such as WPMS were intended to address by providing data for inspection planning. The creation of NPG came out of these concerns.

Over the years, NPG has continued to evolve. As the FAA has developed the capacity to target inspections based on risk, the FAA augmented the NPG program. In 2002, the FAA added the Surveillance and Evaluation Program (SEP) to the NPG program, which allows inspectors to analyze safety data and develop an additional inspection plan for each carrier based on information about safety risk for that air carrier.⁵¹

⁴⁸ U.S. Office of Technology Assessment, “Safe Skies for Tomorrow: Aviation Safety in a Competitive Environment,” OTA-SET-381, July 1988.

⁴⁹ U.S. General Accounting Office, “FAA’s Safety Inspection Management System Lacks Adequate Oversight,” GAO/RCED-90-36, November 13, 1989.

⁵⁰ See, U.S. General Accounting Office, “Compilation and Analysis of the Federal Aviation Administration’s Inspection of a Sample of Commercial Air Carriers,” GAO/RCED-85-157, August 2, 1985. See also, U.S. General Accounting Office, “FAA’s Airline Inspection Program,” Testimony by Herbert R. McLure, 129875, May 14, 1986; U.S. General Accounting Office, “Needed Improvements in FAA’s Airline Inspection Program Are Underway,” GAO/RCED-87-62, May 19, 1987; U.S. General Accounting Office, “FAA Aviation Safety Inspectors Are Not Receiving Needed Training,” GAO/RCED-89-168, September 14, 1989.

⁵¹ U.S. Government Accountability Office, “FAA’s Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs’ Performance,” Testimony by Gerald Dillingham, GAO-06-266T, November 17, 2005.

2.2.2 How the NPG Surveillance Program Works

Surveillance under NPG comprises a significant part of the safety oversight system. Section 49USC44713 of the U.S. Code sets forth information about inspection and maintenance in the aviation system and the role of the FAA. The salient themes in this section of the code are:

(1) The purpose of inspections is to provide information that can be used to determine whether the aviation system is safe;

FAA inspectors should “advise and cooperate with” air carriers;

Inspectors have the authority to decide whether a part of the system is safe;

The FAA has the authority to create regulations that enables its employees to carry out surveillance and it should communicate with interested parties during this rulemaking process;

The FAA should use a system to target surveillance and its employees need to have access to that system; and

Information about safety events such as accidents and incidents (based on independent investigation) should be integrated into the safety performance and analysis system (SPAS) database.

It is FAA’s mission to implement this code, which is manifested in a set of policies, procedures, training, decision-making, and other activities that comprise the safety oversight system.

The FAA *Inspector Handbook*, Orders 8300.10 and 8400.10 are a major source of information about the operation of the safety oversight system. Interviews with subject matter experts and selected FAA employees complement the information from the *Inspector Handbook* and provide insights about the nuances and dynamics of the system.

2.2.2.1 Inspection Planning

Under NPG, the surveillance program has a four-stage surveillance planning and execution program. The first stage is to develop a surveillance plan based on data, prior inspection results, and other information. These NPG plans are the baseline data that guide the level of surveillance that is planned for and executed in a given year. A committee at FAA headquarters creates the annual NPG plan, which specifies the minimum number of required inspections by inspection type for each region. NPG inspections account for about 20 percent of total inspections, and the remaining 80 percent are left to the regions and field offices to plan.

Regions and field offices also decide how to *allocate* the required NPG inspections across the certificate holders in that region in addition to planning the remaining 80 percent of inspections. Because inspection planning needs to be informed by these historical inspection findings, the FAA continued its effort to create data collection systems and information systems to support the program. Thus, the FAA developed PTRS, a database that provides certificate information, operator characteristics, and other information to the NPG planners..

In addition to previous inspection results, NPG requirement planning also considers the following:

- (1) Program objectives;
- (2) Available resources;
- (3) Information about incidents and accidents;
- (4) Information about prior compliance and enforcement actions; and
- (5) Information from the public such as complaints.⁵²

The *Inspector Handbook* states, "...for a routine surveillance program, there should be a representative number of each type of inspection," unless there is information that suggests some areas should receive more information than others. The handbook also states, "...surveillance data may indicate that certain types of inspections are ineffective or that fewer inspections can effectively accomplish the objective."⁵³ Based on the NPG plan from Headquarters, local Principal Inspectors determine inspection requirements in more detail.

Once inspectors have an inspection plan that outlines the inspections that they are required to conduct in a year, each individual inspector is responsible for making decisions about how to allocate their time across inspection and other activities. Though inspectors decide their own schedules, they are required to finish their required inspection program for the year.

Under the NPG program, inspectors can also plan for inspections using the SEP program. According to the GAO, planning meetings for SEP inspections are held at least two times each year. Inspectors may adjust the required inspections to reflect new information or analysis from

⁵² U.S. Federal Aviation Administration, *Inspector Handbook Volume 6: Surveillance, Chapter 1. General Policies and Procedures*. 8400.10 CHG 3, July 31, 1990.

⁵³ *Ibid.*

the SEP program or other sources. Thus, the required inspections can be “retargeted” to areas of higher risk by using the SEP process.⁵⁴

In practice, many factors affect inspection planning, including data analysis. Interviews with selected experts and personnel reveal that an inspector’s analysis of prior experience with the carrier, information from informal sources such as conversations with air carrier employees, and even gut feelings are all relevant sources of information on which inspection plans are made. Databases such as SPAS that were developed to analyze data and facilitate inspection planning may not substitute for inspectors’ prior experience..

After creating a surveillance plan based on the number and type of inspections to perform in a given year, the second phase is to conduct the inspections in the plan. These inspections are conducted by Certificate Holder District Office (CHDO) personnel and sometimes by inspectors from different regions. The third phase of the surveillance program is to analyze the inspection data and create reports of the findings. The fourth phase of the surveillance program is to take action to improve compliance, implement enforcement actions and corrective actions, or implement additional surveillance.⁵⁵

2.2.2.2 Aviation Safety Inspectors

The surveillance program depends on inspectors to carry out inspection plans and collect inspection data that provide information to the inspection planning, enforcement, and other elements of the safety oversight system. Aviation safety inspectors (ASIs) work within a marbled environment in which public administrative functions interact with private commercial activities. Much energy, and effort, is exerted by ASIs in performing their surveillance function, as prescribed by inspection plans, in a manner that fosters and sustains a cooperative environment.

Literature on inspections and interviews with experts reveals that the competence of the inspectorate is critical to inspection quality and improving compliance and safety performance. Inspectors are hired because of their expertise in a specific field of aviation. Aviation inspectors at FAA typically have aviation backgrounds (e.g., they are mechanics, pilots, or former

⁵⁴ U.S. Government Accountability Office, “System Safety Approach Needs Further Integration into FAA’s Oversight of Airlines,” GAO-05-726, September 2005.

⁵⁵ U.S. Federal Aviation Administration, *Inspector Handbook Volume 6: Surveillance, Chapter 1. General Policies and Procedures*. 8400.10 CHG 3, July 31, 1990.

crewmembers). They use their air carrier experience and training to detect and judge the quality of aviation systems, including hardware, software, policies, procedures, and operating environments.

In the structure of the FAA oversight system, inspectors are the agents that implement oversight policies, apply regulations, provide evidence for enforcement cases, and provide information back to the system about activities in the field. The inspectors work alone or in teams inspecting the quality, safety, and compliance of air carriers. They belong to a labor organization, Professional Airways Systems Specialists (PASS).

Inspectors must balance two important relationships. They must ensure that they fulfill their roles and responsibilities as FAA inspectors and subordinates to FAA management, while concurrently maintaining a cordial and respectful working relationship with the air carriers that they are responsible for overseeing. Inspectors must be capable of obtaining and evaluating information that will reveal compliance with FAA safety regulations, while establishing a climate that fosters voluntary actions to improve safety performance.

Inspectors are responsible for planning, executing, and following up on a wide variety of safety oversight activities, including investigations, surveillance, certification, accident prevention, technical assistance, issuance of authorization, and other functions.⁵⁶ Many of these activities may be planned, but an important part of an inspector's work is managing unexpected demands such as processing enforcement cases and conducting investigations, which may occur without advance indication.

2.2.2.3 Conducting Inspections

FAA's *Inspector Handbook* is only one of many sources of information for inspectors. ASIs conduct inspections according to their ongoing training, orders and other policy documents, and, perhaps most importantly, their experience. In this section, information about conducting inspections comes primarily from two sources, the *Inspector Handbook* and discussions with subject matter experts and selected FAA personnel.

⁵⁶ U.S. Federal Aviation Administration, *Inspector Handbook Volume 9: Technical Staff Administration and Responsibilities, Chapter 1. Inspector Responsibilities, Administration and Conduct, Section 1, Responsibilities of Aviation Safety Inspectors*, 8400.10 CHG 10, December 20, 1994.

The FAA's definition of inspections is the following: "Inspections are specific work activities which have the following characteristics: a specific work activity title and PTRS code; a definite beginning and a definite end; defined procedures; specific objectives; a requirement for a report of finding."⁵⁷ Thus, conducting an inspection is closely related to the *reporting* of inspection findings. Inspectors performing NPG inspections report their findings through PTRS.

PTRS provides inspectors with (1) data sheets, (2) comment codes, (3) job aids, and (4) standard and ad hoc reports.⁵⁸ The data sheet is a hard-copy form in which an ASI can record all of the information that describes an inspection. Instead of using the hard-copy form, the inspector can enter the data directly into a computer. The fields in the form include places to record:

- Inspector name code;
- Record ID;
- Relevant FAR;
- Start date and completion date;
- Data about the job function;
- Overall results;
- Data about the subject of the inspection (e.g., equipment)
- Data about relevant air carrier personnel;
- Comment codes;
- Opinion codes; and
- Comment text.

The ASI uses a PTRS comment code to indicate the subject of the inspection. The FAA organized its comment codes by *primary areas* (e.g., Air Carrier Operations, General Aviation Operations, ATC/Air Space, Airports, Air Agencies, Air Carrier Airworthiness, General Aviation Airworthiness, Aircraft, and Crewmembers), and then by *key word listings* that include personnel, manuals, records/reports, training, facilities/equipment/surface, conformance, operations, maintenance, management, and administration. Within each key word listing are several sub-listings that describe the subject of the code in more detail.

⁵⁷ U.S. Federal Aviation Administration, *Inspector Handbook Volume 6: Surveillance, Chapter 1. General Policies and Procedures*. 8400.10 CHG 3, July 31, 1990.

⁵⁸ *Ibid.*

PTRS job aids are available to help ASIs choose which comment code should be used. The *Inspector Handbook* suggests that writing the comment in plain language before selecting a comment code can help inspectors choose the right code to use.

The inspector also chooses an *opinion code* to represent the outcome of an inspection. The opinion codes include “U” – unacceptable, “P” – potential, “I” – information, and “E” – exceeds. These codes can be interpreted in the following way:

- **Unacceptable** inspections mean that the inspector believed the subject of the inspection was not in compliance with regulations or safe operating practices, or was inadequate or unsatisfactory. This determination may not match the overall outcome of the inspection, which could be satisfactory even if one element was unacceptable. The *Inspector Handbook* states that if an inspector is able to correct the noncompliance during the inspection, then the inspection should record a “U” opinion even though the matter was fixed.
- **Potential** inspections mean that the inspector believed that the subject of the inspection had the potential to be in noncompliance with regulations or safe operating practices. This code is intended to provide inspectors with a way to communicate findings in a “grey area.”
- **Information** inspections mean that the inspector does not have an opinion about the finding, usually because of some lack of additional information. Inspections coded with an “I” may also be instances in which an inspector recorded an observation.
- **Exceeds** inspections mean that the inspector believed that the subject of the investigation exceeded the requirements.

Though PTRS comment codes and opinion codes are extremely important for analysis, the *Inspector Handbook* states that an inspector’s narrative text is “the most important part of the overall work activity report” because these comments “are the only means of accurately recording what the inspector has actually observed.”⁵⁹ The *Handbook* addresses some of the nuances of inspections.

Often an inspector may need to express an opinion that a person, item, or subject area evaluated during a work activity or inspection was unacceptable, but still may find that the overall results of the particular work activity was satisfactory...If a sufficient number of unacceptable opinions are recorded, justification can be established to support a revision to the standards. If an inspector is able to correct a situation or deficiency during the work activity which was unacceptable, the inspector should record an ‘unacceptable’ opinion code. The ‘unacceptable’ opinion code provides information for

⁵⁹ *Ibid.*

*future analysis and trend identification, in this case, however, the inspector's comment should include that corrective action was taken.*⁶⁰

Subject matter experts and selected FAA personnel discussed the procedures for inspections in practice. Though the *Inspector Handbook* provides important information about what ideally happens during an inspection, these sources provided additional information about the practice of inspection. When an inspector conducts an inspection, he or she is required to report every instance of noncompliance. It is difficult to know whether this happens in practice because PTRS records may not contain information about everything an inspector observed while conducting an inspection. It may also be difficult to interpret inspection results if inspectors record opinion codes that do not reflect their true opinions. A survey of inspectors by the GAO found that inspectors did not report all of the violations they found during fiscal year 1996 using PTRS comment codes. Twenty-eight percent of inspectors responded that they reported between 96 and 100 percent of the violations that they find, but 33 percent responded that they reported less than half of the violations that they found.⁶¹ Anecdotally, subject matter experts said that the climate has changed since that time, and that inspectors now report all violations.

Inspectors have an adjudicative function in the safety oversight system. The information that inspectors collect during their formal inspections has implications that can range from whether the FAA will continue to allow the air carrier to operate to whether the FAA will require the air carrier to make an investment that could be costly even if it is made for the purpose of preventing the costs of an accident. In many ways, "Inspection can have a dual character, seeking either to identify non-compliance, or to develop means for assuring compliance, with regulatory measures. In this context, much will depend on the attitude the inspectorate brings to its work."⁶²

Inspector workload is another factor that influences the effectiveness of inspections. Inspectors are responsible for allocating their own time and completing their inspection

⁶⁰ *Ibid.*

⁶¹ U.S. General Accounting Office. "Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks." GAO/RCED-98-6. February 1998.

⁶² Peter L. Strauss, *An Introduction to Administrative Justice in the United States*. Durham: Carolina Academic Press, 1989.

requirements. Through discussions with subject matter experts, we learned that if unexpected investigations or special inspections consume a large part of an inspector's time, then the inspector may rush to complete the required inspections at the end of the year. Rushing the inspections may result in fewer reports of violations that may require enforcement action, or fewer "unacceptable" opinion codes that require follow up with corrective action. Instead of reporting an "unacceptable" opinion, an inspector may instead report that the inspection was inconclusive by entering "I" for "information. One possible explanation is that workload factors influence the outcomes of inspections. Subject matter experts and selected FAA personnel explained that the initial enforcement process requires between 30 and 40 person hours to complete. In GAO's survey of inspectors, 25 percent of the survey respondents said that the additional workload of opening enforcement cases was a "major reason" for not opening enforcement cases in some situations. In the same survey, 24 percent of respondents said that it was a "minor reason" for not doing so. Survey respondents also noted that compliance is more important than enforcement, and that it is more important to gain immediate compliance than open an enforcement case.⁶³

2.3 Air Transportation Oversight System (ATOS)

In response to the crash of ValuJet Flight 592 in 1996, evaluations from outside agencies, and internal evaluations, the FAA developed a new safety oversight system, the Air Transportation Oversight System (ATOS), in conjunction with Sandia National Laboratory. The FAA implemented ATOS for the 10 largest passenger carriers in October 1998, while NPG was still active for the remaining Part 121 air carriers. Because ATOS continues to evolve, the information in this section comes from interviews and discussions with the FAA and industry personnel, Appendix 6 of the *Inspector Handbook*, as well as recent GAO reports.

ATOS has characteristics that are different from the NPG system. ATOS was designed to provide more structured inspections that address the safety systems that commercial air carriers have in place, and the extent to which air carriers follow the policies and procedures of these safety systems. Additionally, ATOS includes new and revised inspection procedures, and operations research analysts and certificate management teams.

⁶³ U.S. General Accounting Office. "Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks." GAO/RCED-98-6. February 1998.

In other ways, ATOS has many of the same components as the NPG system. For example, ATOS is implemented for each certificated air carrier by an array of inspectors, including principal inspectors, ASIs, and regional managers. ATOS is also a method to identify whether air carriers are in compliance with FARs and their certificates.

ATOS was designed with an organizational approach to safety oversight, called a “systems approach to safety oversight.” It emphasizes communications, data reporting, analysis, and evaluation. Appendix 6 of the *Inspector Handbook* describes the policies and procedures of ATOS. The following are definitions from the ATOS document that are helpful for understanding ATOS, and they exemplify the approach of the new oversight system.

Air Carrier System A group of interrelated processes which are a composite of people, procedures, materials, tools, equipment, facilities, and software operating in a specific environment to perform a specific task or achieve a specific purpose, support, or mission requirement from an air carrier.

Analyst The ATOS Operations Research Analyst (ORA) responsible for assisting the CMT in collecting and analyzing air carrier data.

Certificate Management Team (CMT) The team responsible for the surveillance of a specific air carrier. The Certificate Management Team will develop and execute a Comprehensive Surveillance Plan tailored to an air carrier.

Data Evaluation Program Manager The CMT member responsible for reviewing inspection reports and records to ensure they meet quality guidelines.

Safety An inherent attribute of an air carrier’s properly designed systems, sub-systems, and elements.

System Approach The structured, safety-driven means by which the FAA will certificate and monitor elements that are designed to interact predictably within the air carrier’s systems and sub-systems.

System Safety The application of special technical and managerial skills to identify, analyze, assess and control hazards and risks associated with a complete system. System safety is applied throughout a system’s entire lifecycle to achieve an acceptable level of risk within the constraints of operational effectiveness, time, and cost.⁶⁴

The *Inspector Handbook* Appendix that presents the ATOS polices also sets forth the roles and responsibilities of each of the users of the system, including the director of AFS, and

⁶⁴ U.S. Federal Aviation Administration, *Inspector Handbook Appendix 6*, 8400.10 CHG 13, October 19, 2001.

AFS personnel in the FAA headquarters office, regional offices, the certificate holding office, the district offices, and other teams.⁶⁵ Accordingly, ATOS acknowledges the organizational elements that define how safety oversight is to be executed. This is one of the important differences between ATOS and the NPG system; the NPG system did not specify clearly the roles of the people who carry out the program.

Under NPG, the guidelines provide the baseline minimum surveillance activities that should be included in a surveillance plan. The *Inspector Handbook* states that NPG surveillance plans should be made at headquarters, regional and district offices, and at the level of the individual inspector, but it does not specify what the plans should include or how they should be integrated with each other. Once each group creates an NPG plan, the *Inspector Handbook* states that it should be accomplished by carrying out inspections. These inspections should result in “accurate and qualitative” reporting because quality reporting is necessary for analysis and possible enforcement.⁶⁶

Under ATOS, aviation safety inspectors have new training and a different organizational context to support their activities. Instead of autonomous inspection planning at various levels of AFS, ATOS inspections are planned by Certificate Management Teams. Each year, the team participates in a planning meeting. This meeting is part of the Certificate Management Process. A planning meeting coordinator arranges the meeting, and receives assistance from principal inspector, cabin safety inspector, aviation safety inspectors, and an analyst for the carrier.

The *Inspector Handbook* explains the process for planning the meeting in detail, including how the coordinator should assign the meeting tasks, date, time, and location. The *Inspector Handbook* also explains in detail how the coordinator should plan the audio-visual equipment, computer equipment, and Internet access needed for the meeting. This meeting should result in a surveillance plan, but according to the *Inspector Handbook*, the meeting also builds team skills, establishes team norms, communicates the team’s expectations, and provides an opportunity to share information.⁶⁷

⁶⁵ *Ibid.*

⁶⁶ U.S. Federal Aviation Administration, *Inspector Handbook Volume 6: Surveillance, Chapter 1. General Policies and Procedures*. 8400.10 CHG 3, July 31, 1990.

⁶⁷ U.S. Federal Aviation Administration, *Inspector Handbook Appendix 6*, 8400.10 CHG 13, October 19, 2001.

The ATOS Appendix in the *Inspector Handbook* explains the system safety approach: “System safety is a multidisciplinary approach to systematically make a system, product, process, operation, or facility as safe as is practical. System safety covers the entire spectrum of activities from the design of hardware to the culture and attitudes of the people involved.”⁶⁸ The principles of system safety are that safety is inherent in the system and must be designed into the system. The document states explicitly, “...safety cannot be inspected into the system.”⁶⁹ (Emphasis in original.)

Teams of inspectors carry out safety attribute inspections that assess the safety systems that they have in place such as quality assurance programs. Then, individual inspectors carry out element performance inspections to determine if an air carrier follows its written procedures and controls for those safety systems. Inspectors use checklists to guide their inspections, and they use databases collect their inspection data. The data collection process is very structured and iterated. An inspector opens a record in which to collect information and then can choose to save the record as a work in progress or as a final draft of the information. Then, the information goes to a data quality analyst who audits the data and will send the record back to the inspector if information is missing. The operations research analyst, who then provides his or her reports back to the CMT, analyses the inspection data.

There are special reports in the *Inspector Handbook* section on ATOS that address reporting guidelines. ATOS’s systems approach seems to have created a good approach to inspection planning and a more structured set of inspection protocols, but it did not make the data collection and reporting tasks of inspectors easier.

The ATOS data collection process arguably took steps backwards with respect to data collection, though the design was intended to simplify the process for inspectors. Instead of creating a system that could provide robust analyses of textual information, the first iteration of ATOS used a binary yes/no checklist. Not only did this design preclude many meaningful quantitative analyses of inspection data, it also precluded forms of qualitative analysis.⁷⁰

⁶⁸ *Ibid.*

⁶⁹ *Ibid.*

⁷⁰ U.S. General Accounting Office, "FAA's New Inspection System Offers Promise, but Problems Need to Be Addressed," GAO/RCED-99-183, June 28, 1999.

2.3.1 What ATOS Generates

Under ATOS, ASIs assess commercial air carriers against established performance measures. The basis for these performance measures are regulatory requirements and safety characteristics applicable to a carrier's systems, subsystems, and elements.⁷¹ Using processes embedded in ATOS, ASIs identify and prepare a report documenting noncompliance or weaknesses and deficiencies in air carrier systems that jeopardize safety. This information is used by CMTs to identify trends, analyze problems for root causes, develop corrective actions and preventive measures, and plan subsequent inspections.

The safety system process forms the foundation of ATOS. It is based on seven air carrier systems:

- Aircraft Configuration and Control – maintenance of the physical condition of the aircraft and related components
- Manuals System – information and instructions to define and govern air carrier activities
- Flight Operations – aircraft movement
- Personnel Training and Qualification – processes to ensure that personnel are competent for their assigned tasks and responsibilities
- Route Structures – maintenance of facilities on approved routes
- Airmen and Crewmember Flight, Rest, and Duty Time – service limitations for employees
- Technical Administration – certification and operations

Associated with each of these systems are discrete subsystems and elements. The inspection process is structured around these systems and their associated subsystems and corresponding elements.

⁷¹ Seven air carrier systems provide the foundation for the ATOS system-based approach to safety. A defined set of subsystems and elements are associated with each system. Elements are discrete activities or actions that support subsystems and systems. (See 8400.10, Appendix 6.)

2.3.2 ATOS Inspections

The ATOS process is based on two distinct types of inspections yielding complementary results. Safety attribute inspections (SAIs) identify the safety attributes that are expected to be inherent in well-designed air carrier systems. Six safety attributes are defined in ATOS as follows:

- Procedures – there are documented methods for doing a process
- Controls – there are checks and restraints designed into a process to get a desired result
- Process Measurement – the air carrier identifies and assesses its processes to identify and correct problems or potential problems
- Interfaces – the air carrier identifies and manages the interactions between processes
- Responsibility – there is a clearly identifiable, qualified, and knowledgeable person who is accountable for the quality of a process
- Authority – there is a clearly identifiable, qualified, and knowledgeable person with the authority to set up and change a process

Using ATOS-prescribed data collection tools (DCTs), ASIs perform inspections that focus on ensuring that these safety attributes are included in air carrier programs consistent with applicable regulations.

The ATOS process also requires that element performance inspections (EPIs) be performed. These inspections focus on determining whether an air carrier complies with its written procedures and controls, and is operating in a manner that is consistent with established performance measures for each system. Further, EPIs are performed at the element level; there are 122 ATOS elements associated with 20 subsystems that are in turn sorted among the aforementioned seven ATOS systems (see 8400.10, Appendix 6).

Both SAIs and EPIs yield “Yes”, “No”, or “Not Applicable” responses to specific questions. Generally, a “Yes” response is always viewed as positive, as it indicates compliance with applicable specific regulatory requirements (SRRs), related Code of Federal Regulations (CFRs), and/or any FAA policies or guidance appropriate to that element. A “Yes” response also indicates that the observed procedures and system safety principles approved/accepted for

the Certificate Holder are being followed. A qualified “Yes” response suggests that the response should actually be a “No.”

A “No” response means that the certificate holder either is not complying with observed specific regulatory requirements (SRRs), related CFRs, and/or applicable FAA policies or guidance for that element, or that the certificate holder’s procedures are not being followed. A “No” response can also mean that system safety procedures are weak in the area being evaluated and that the certificate holder’s approved/accepted procedures are inadequate. Hence, a response limited to just “No” presents ambiguities. Therefore, the written report prepared by an ASI must qualify the “No” response with additional information to ensure that appropriate information is provided to the CMTs to facilitate identifying and analyzing root causes, and developing corrective and preventive actions.

2.3.3 Database Characteristics

Implementation of ATOS has yielded a wide range of experiences and related comments. Generally, implementation has been slow. Due to the very large amount of information collected and maintained in ATOS, substantial effort is required to utilize the data, as access is cumbersome. A significant time investment is required to achieve the level of competency to manipulate the database with some ease and proficiency. A challenge to its utility is that implementation guidance has been ambiguous, training inadequate, and inspectors found it difficult to apply the broad ATOS “checklist” format to certain important findings.

In theory, ATOS achieves better standardization than previous surveillance initiatives, as only ATOS-trained personnel perform inspections. Further, in some instances ASIs using ATOS are uncovering violations and unsafe conditions that otherwise would not have been identified (e.g., problems with air carriers’ manuals).

The independence of ATOS-based oversight and the data that it generates remains questionable for a variety of reasons, most important of which is bias. Inspectors finding it difficult to follow ATOS guidance do not always adhere to the ATOS process, and accordingly plan their oversight activities based on past experience. Often this experience is limited to a specific inspector, and therefore is not necessarily consistent with the preponderance of safety performance data. Further, ASIs sometimes participate with carrier employees on their respective self-assessment activities. While this “shadowing” or “piggy-backing” technique

enables an ASI to evaluate carrier self-assessment programs, it can also create inspector bias due to the scope of the self-assessment and/or the interpersonal relationships between an FAA inspector and carrier personnel. According to a review by the DOT IG of FAA surveillance of air carriers' Continuing Analysis and Surveillance Systems, the FAA's approach to evaluating air carriers' self-assessment programs (a fundamental element of ATOS) is hindered by ineffectiveness in its training programs and the implementations of improvements in the safety oversight system.⁷²

Timely reporting of safety violations, deficiencies, weaknesses, issues, and concerns is critical to the effectiveness of the FAA safety oversight system. Accordingly, prompt close-out of corrective actions and broad dissemination of lessons learned in order to develop preventive measures are key to continually improving safety performance. Executing an ATOS-based inspection is time consuming, especially if an ASI is not proficient with the ATOS processes and its associated automated tools. Although comprehensiveness is an important attribute of the ATOS database, some argue that achieving this goal may come at the cost of its timeliness, as prompt feedback is necessary to mitigating future safety risks. Further complicating this concern is the slow speed at which ATOS is being fully implemented. This was clearly evident after the Alaska Airlines accident that occurred in January 2000, as the ATOS database contained little information about the carrier's safety performance based on inspections. Controversy exists today as to whether the non-ATOS data currently resident at the FAA, as well as that already collected by ATOS, has been adequately "mined" for useful information. Some argue for applying more and better resources to analyzing the existing FAA database rather than refining ATOS. Regardless, it is important that the type of information collected be assessed periodically and determined whether it is needed or useful, and revise the data collection mechanism and processes accordingly.

2.4 Practitioner and Air Carrier Perspectives on Safety Oversight

FAA's documented safety oversight policies and program documents are an important source of information about the system. But the documented policies and program documents,

⁷² U.S. Department of Transportation, Office of the Inspector General, "Report on Oversight of Aircraft Maintenance, Continuing Analysis and Surveillance Systems," AV-2002-066, December 12, 2001.

despite their detail, do not contain all wisdom about how to perform safety oversight. This section presents additional information about the activities of the safety oversight system. This section also discusses dynamics of safety oversight that are not discussed in FAA documents concerning the system. The sources of information for this section include selected current and former employees of the FAA safety oversight system and air carrier personnel, including subject matter experts.

Certain topics received attention from nearly all of the participants with whom we spoke. These topics include the standardization of safety oversight, the role of trust in surveillance, improving the safety oversight system, and the skills and training of the FAA inspectorate. Each of these topics is discussed in more detail below.

2.4.1 Standardization of Safety Oversight

The desire to make safety oversight more standardized across regions and carriers is not unique to the FAA. All safety regulators doubtlessly strive to make safety oversight more consistent, objective, and predictable. A legitimate concern is that safety oversight is still subjective, even when it becomes more data-driven.

The ability to treat all carriers the same is difficult for a process that relies heavily on an individual's ability to assess safety. Each inspector inspects differently depending on the type of inspection, the history and attitude of the certificate holder, and the inspector's background and experience. Similarly, inspectors have their own methods for investigating safety concerns, and these methods may be based on areas in which they have special knowledge. The extent to which training and other variables affect the variability of inspection is not well known. Better integration of safety oversight programs, and improving the ability to share and use information across functional and geographic elements of safety oversight could help inspectors be confident that their application of the regulations is consistent across air carriers.

ATOS was designed to improve the standardization of safety oversight with checklists that guide inspections, but the information required to answer those questions may not be easy to obtain in a uniform way. For example, questions from SAIs about the content of an air carrier's manuals may be asked in a uniform way on the checklist, but inspectors who answer these questions may do so differently depending on the organization of the manuals and the ability to identify and interpret text that resolves an SAI question. Air carriers' manuals are extensive and

sometimes decentralized. Finding information in the manuals to answer an ATOS question can be difficult if the answer is not clear, or if the manuals can be interpreted in multiple ways. The process of reviewing manuals, judging the content of the manuals against the question on the checklist can be time consuming for both the inspector and the air carrier and it may not lead to a conclusive yes or no answer, which can frustrate inspectors and air carriers alike.⁷³ The evidence required to resolve an SAI question can also leave room for uncertainty. An example of this is whether air carriers' manuals need to name *specific individuals* who are in positions of accountability, or whether the naming of each *accountable position* is sufficient to complete the description of a process. These examples of subjectivity in interpreting ATOS SAI questions, and interpreting the content to resolve those questions suggests that ATOS has not overcome the long-standing dilemma in safety oversight – standardization.

2.4.2 Role of Trust in Surveillance

Trust is a key element in the relationship between FAA and air carriers. Safety oversight occurs at several levels, and one is between FAA and the certificate holder and the inspector at the front line. This relationship is one in which trust is necessary.

The “ground truth” of voluntary self-disclosure also offers insights about how trust in the safety oversight system operates in practice. Voluntary self-disclosure of violations, or even changes in service, can happen formally through an administrative process, or it can happen informally through conversation. Both kinds of communication may be indicators of trust between industry and the FAA front-line. . It is possible that documenting unsafe conditions and violations threatens the trust between the air carrier and the inspector. Yet, the oversight system has an interest in collecting and documenting such information. Thus, to have access to information, the FAA and air carriers implement self-disclosure arrangements.

According to Subject Matter Experts, interpreting the self-disclosure conditions is not always straightforward. Sometimes, self-disclosure can be misinterpreted and lead to enforcement actions. The anonymity of volunteering information can also be problematic for air carriers, such as during sensitive contract negotiation periods. The role of trust in surveillance can cut both ways.

⁷³ FJLG comments.

Another example of this arises in situations in which either the FAA or the air carrier would like information to be shared. The FAA has an interest in obtaining access to air carriers' information. Similarly, the air carriers sometimes do not have access to FAA information that they believe would improve an aspect of their participation in oversight.

The relationship between the carrier and the safety oversight system sometimes goes beyond the collection of information and includes discussions of strategies to improve some aspect of service. For example, if there is a condition that is not a rule violation, but the FAA principal inspector believes needs to be addressed, then the FAA must use a different strategy to ensure that the condition receives treatment. In these situations, there are no FAA regulations that can be leveraged to control that treatment, and instead the treatment must be negotiated between the air carrier and the FAA. Subject matter experts pointed to the FAA Customer Service Initiative as a program developed to resolve disagreements between air carriers and the FAA. According to the FAA website, the Customer Service Initiative, issues that arise between an air carrier and the FAA should be documented (including any relevant rulemaking or advisories, dissenting opinions, and prior experience with the issue) and raised to the next level for review. Essentially, the Customer Service Initiative allows air carriers to have assessments made by individual inspectors reviewed formally.⁷⁴

The example in the previous section of disagreements over the interpretation of an air carrier's manuals illustrates how ambiguity in ATOS may be related to any deterioration of the relationship between the FAA and air carriers.

2.4.3 Improving Safety Oversight

The implementation of ATOS clearly created additional work for the air carriers, and comments from Subject Matter Experts suggest that although some improvements were made as a result of the additional work, the increase in oversight on the air carrier side was not well understood or anticipated by the FAA.

FAA personnel maintain that the safety oversight system could do a better job using the wealth of resident experience, knowledge, and data. In many ways, employees of the safety

⁷⁴ U.S. FAA, "Customer Service Initiative," http://www.faa.gov/about/office_org/headquarters_offices/avs/offices/afs/cust_service/, Website visited April 17, 2006. Also FJLG comments.

oversight system hold this knowledge. Thus, it may not always be found in databases. The experience of FAA personnel plays a significant role in inspection planning, targeting, and detection, but this role is considered “informal” when safety oversight data are not used.

Inspections focus on document and record reviews as well as observation of work, although the latter area deserves greater attention than what it currently receives. The observation of work depends on the ability to send inspectors to various locations; this becomes more difficult with complex inspection plans that require multiple inspectors to carry out the work.

2.4.4 Training

The capacity to make constant improvements to the safety oversight system depends on the management of the safety oversight system, and the ability of each employee in the system to carry out change. New inspectors are trained already in special skills, and their skills develop over time with experience and additional training.

Selected junior and senior inspectors interviewed said that safety oversight employees need soft-skills training in addition to the traditional skill set. Soft skills, such as negotiation techniques, are important for inspectors when they work with air carriers on the implementation of corrective actions. For example, ATOS has become easier to implement as the air carriers accept the process, and skills that help inspectors gain air carrier acceptance can be as useful as technical skills. Negotiation is particularly important in the ATOS environment where the FAA and air carriers encounter more opportunities to address safety beyond the limits of the regulations.

Inspectors also expressed a need to receive more training in auditing methods. Auditing methods help inspectors evaluate the self-evaluation programs and safety management systems at the air carriers.

The use of databases is increasingly important as the safety oversight system changes to target inspections to areas of relatively higher risk. Selected inspectors reflected on their use of databases and said that SPAS is easy to use. Some inspectors said that SPAS is the best tool the inspectors have. One inspector says that she prefers SPAS because she has to “drill down” much farther when she’s looking at ATOS data, which some inspectors described as “cumbersome.” Other inspectors agreed, but added that it is too time consuming to do much analysis because

there is already enough work to keep them busy. The databases available to inspectors in the safety oversight system provide useful information, but the utility of the databases might increase if separate sources of information were more integrated. Subject matter experts suggested that integrating databases at the level of the air carrier could be an effective way to “present a clear picture of the airline.”⁷⁵

Subject matter experts with experience working for air carriers emphasized the value of selection and training of inspectors. According to these experts, selecting inspectors with air carrier experience, especially experience in multiple aspects of the air carrier’s business, is valuable in the safety oversight system. Training is another way in which FAA could develop these skills in the inspector workforce. Developing checklists and more standardized approaches to safety oversight are not sufficient if individual inspectors have a general lack of knowledge about how airlines work. Selected FAA personnel and subject matter experts discussed the difficulty of keeping up with current practices and technologies. Both managers and inspectors said that knowledge becomes obsolete once one leaves the airline. Subject matter experts noted that airlines and FAA do not attend training courses together, and that it may be a year before FAA employees can receive the same training offered to air carrier employees.⁷⁶ When prior training becomes obsolete, there is a higher chance that inspectors will limit their inspections to areas that they know well. According to Subject Matter Experts, inspections should have more coverage than one inspector’s comfort zone.

Discussions around the theme of training also reveal that air carriers may be interested in ATOS training offered to FAA employees. One subject matter expert said that an air carrier that participated in FAA’s own ATOS training alongside FAA employees, and the information learned through that experience enabled the carrier to understand the ATOS program, its processes and policies, and better adjust to the program.⁷⁷

⁷⁵ FJLG comments.

⁷⁶ FJLG comments.

⁷⁷ FJLG comments.

3. External Evaluation of the Aviation Safety Oversight System

Evaluations of the FAA safety oversight system made by external reviewers such as the GAO, the DOT IG, and the Office of Management and Budget (OMB) are a rich source of information about the methods used to evaluate the performance of safety oversight systems. In many cases, the methods and indicators used in these evaluations are applicable to developing a methodology for internal evaluation. This is because they are focused directly on the safety oversight system, pertain to standards used more generally throughout government, and use information that is already available. Though the methods employed by these agencies and the resulting findings are often criticized by those who are subjected to the reviews, we use this type of evaluation as a source because it is a clear example of how the legislative branch of government looks at safety oversight programs. These evaluations can have a powerful effect on the safety oversight system and result in change. If the methods and findings of these evaluations do not point in the best direction for change, then that is additional incentive to have an internal assessment of safety oversight. With multiple sources of information, the direction of change may be more informed.

This chapter discusses the role of these external evaluators. It includes a discussion of the types of evaluations conducted, the standards used to measure the effectiveness of the FAA safety oversight system, and the indicators and measures used to identify program effectiveness.

3.1 Background

Congress authorizes the creation of safety oversight systems, but does not implement such administrative programs, instead granting safety regulators broad authority to create and enforce safety regulations and review the performance of an industry. Because Congress and other parts of government are invested in the outcomes of safety oversight programs, though they may not observe directly their outputs or processes, these entities must evaluate these programs to know how and what they are doing and whether their activities are effective. Evaluations of administrative programs are sometimes explicit, as in the case of an investigation by the GAO, but they are sometimes implicit, such as in the budget process when agencies justify their programs to Congress.

In addition to informing Congress and justifying expenditure, program evaluations can serve a variety of other purposes. Program evaluations are one way in which overseers prevent fraud and corruption. One could compare the claims of program management against an evaluation of the outputs and outcomes of the program to tell whether there are discrepancies. Similarly, when accountability for program outcomes is diffused across many layers of managers and other employees, an evaluation can bring outcomes to light and link them with accountability. According to Strauss, evaluations ensure “the control as well as the effectiveness of government.”⁷⁸

Some evaluations try to help an agency accomplish the most that it can with constrained resources. These evaluations focus on the management of the system with the motivation to get more productivity or effectiveness from constrained resources, adjust program goals, or extend programs that are effective.

Other evaluations are concerned with the broader goals of government and whether the safety oversight system is carrying out its mission as set forth in the legislation that created it, or whether the mandate or methods of the oversight system should change. For example, Mary Schiavo and Ralph Nader have both published books in which they questioned the principles underlying the FAA oversight system.^{79,80}

Various government watchdogs including the GAO, associated with Congress, the OMB, associated with the President, and the Office of the Inspector General (DOT IG), associated with the Department of Transportation, evaluate safety oversight systems. In evaluations made by the GAO, the OMB, and the DOT IG, safety oversight systems are evaluated against their legislative mandates, and against the standards set forth in the Administrative Procedure Act and the Government Performance and Results Act. For example, the Government Performance and Results Act of 1993 required quantitative assessments of programs such as safety oversight systems. There is a strong current in government to quantify the performance of people, programs, and systems. Measurements can assess the quality of a program and ensure transparency.

⁷⁸ Peter L. Strauss, *An Introduction to Administrative Justice in the United States*. Durham: Carolina Academic Press, 1989.

⁷⁹ Mary, Schiavo, *Flying Blind, Flying Safe*, New York: Avon Books, 1997.

⁸⁰ Ralph Nader and Wesley J. Smith, *Collision Course: The Truth About Aviation Safety*, Blue Ridge Summit: TAB Books, 1994.

Based on these evaluation objectives, the GAO and other external evaluators employ various evaluation methods. Evaluation methodology is discussed in the next section.

3.2 Evaluation Methodology

A basic evaluation methodology could measure and assess the performance of a program by comparing the actual program activities, outputs, and outcomes, to the planned program activities, intended output, and desired outcomes. This is consistent with the definition of evaluation research according to Patton, and it describes a large proportion of the evaluations of safety oversight systems made by external evaluators. According to Patton:

*Evaluation research is the systematic collection of information about the activities and outcomes of actual programs in order for interested persons to make judgments about specific aspects of what the program is doing and affecting.*⁸¹

Is this basic methodology sufficient to assess safety oversight systems? It may be straightforward to account for the activities of the safety oversight system (e.g., inspections, certification), and to account for outputs of the safety oversight (e.g., inspection results, certificates produced), but in a complex system in which the regulator affects safety indirectly by influencing the behavior of industry, and where industry has strong incentives to produce a highly reliable and safe service irrespective of government regulation, it is difficult to isolate exactly what safety oversight contributes to the safety record. While it would be of great interest to determine how much additional safety was produced by the safety oversight system this is especially difficult to answer when the safety record is very good, as it is for aviation, rail, and nuclear power. When there are few bad outcomes from which to make inferences evaluators need to select alternate outcomes that are measurable and easily observable. The challenge, then, is to select indicators that are plausibly (if not definitively) related to safety outcomes but can be more easily tracked and trended.

To see how external evaluators approach this challenge, we reviewed 56 GAO evaluations of FAA's safety oversight of large commercial carriers from 1971 through 2005 to learn what metrics, sources of information, and methods the GAO uses to make its assessments of the effectiveness of safety oversight. A table listing the documents reviewed is included in an

⁸¹ Michael Quinn Patton, *Utilization-Focused Evaluation*, Beverly Hills: Sage Publications, 1978.

appendix to this report. Our review did not include GAO evaluations of FAA oversight or control of national airspace, aviation security, or other topics not directly related to the FAA safety oversight system.

For each report, we identified statements in which the GAO made a judgment about the effectiveness, efficiency, or some other attribute of the FAA safety oversight system. Then, for each statement we identified the evidence that the GAO used to support its findings and the method by which GAO collected the evidence.

The following section presents and discusses the findings from our analyses of GAO evaluations of the FAA safety oversight system.

3.3 Frequency of Evaluation

A search of the GAO's database of reports and testimony from 1970 to 2005, reveals that the GAO published over one hundred evaluations of the FAA safety oversight system, aviation safety, program management, and regulation.⁸² The frequency of these evaluations increased dramatically in the middle-1980s, as is shown in Figures 3 and 4.

3.4 Focus of Evaluation

The titles of the GAO reports demonstrate that the increased rate of publication was accompanied by sharper focus on specific operations of safety oversight elements (e.g., inspections, data collection). For example, a report from 1973 discussed FAA regulation of air taxi operators and a report from 1978 discussed FAA's role in requirements governing alcohol use among civilian pilots. But in 1986, the GAO wrote one of its first evaluations detailing its findings on the effectiveness the FAA inspection program. The increasing interest in oversight operations coincided with the increase in the GAO's rate of publication of evaluations of the safety oversight system.

In several cases, a single issue such as aging aircraft, the inspection program, deicing, or repair stations received attention in multiple reports and testimony in a single year and sometimes across multiple years. In this way, the GAO tracks evaluation topics over time and changes in specific performance areas over time.

⁸² Over 100 reports discussed aviation safety, regulation, program management, or safety oversight. Other GAO reports address FAA finance, air traffic control, or other issues that were not in the scope of this report.

Figure 3: Number of GAO Evaluations of the FAA Safety Oversight System Published per Year, 1970-2005

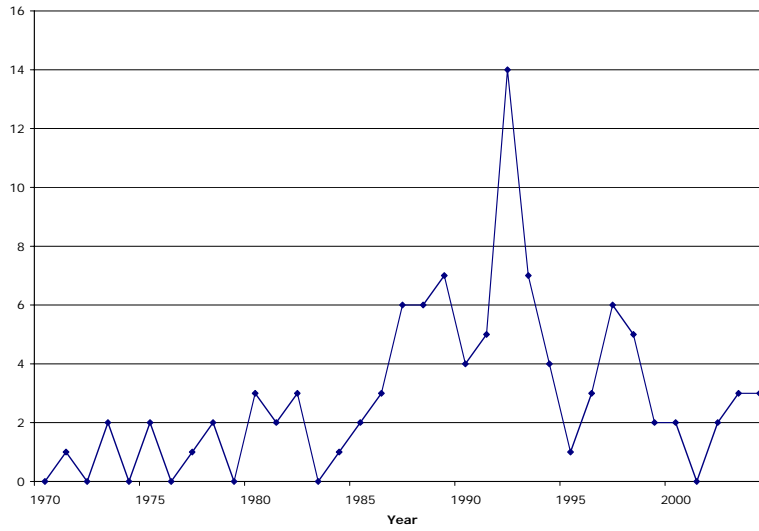
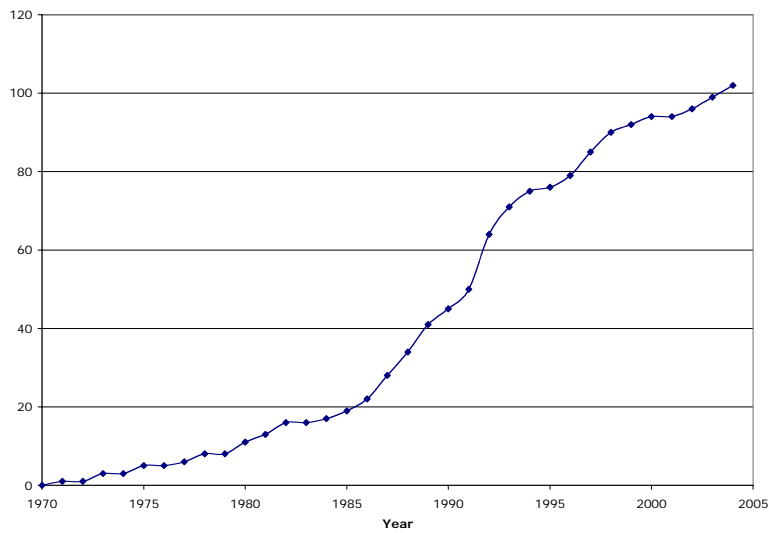


Figure 4: Cumulative Number of GAO Evaluations of the FAA Safety Oversight System, 1970-2005



3.5 Indicators of Effectiveness

We summarized the GAO's oversight evaluation methods using data collected from 56 published evaluations by the GAO from 1971 to 2005. We chose these 56 reports because each one addressed at least one aspect of the effectiveness of the safety oversight system. Table 3 presents a summary of the seven primary evaluation methods that the GAO used to evaluate the FAA safety oversight system. The GAO's methods fall into two categories: reviews and analyses of existing data and information (e.g., NTSB reports, inspection databases, prior GAO evaluations) and the collection of original data using interviews and surveys.

In each evaluation, the GAO reviews and analyzes existing documents and data to form the background of the report. The background review brings to light many of the things that are known already about a particular issue in safety oversight. The GAO references its own prior evaluations (e.g., repair stations are an evaluation topic now, but they have been a topic of evaluation since the middle 1980s), the evaluations created by the NTSB (e.g., from accident investigations), or the evaluations of the Office of the Inspector General to establish a basis for judgment. The GAO's reviews of existing documents are not limited to prior evaluations of the safety oversight system. The document review also includes safety oversight policy documents, internal documents about training programs or other specific issues, or information about the safety record of the air carriers.

The GAO also uses documents as a source of information to address specific questions about the performance of the safety oversight system. It uses documents to show that the safety oversight system has specific programs to carry out its mission, and that the program has a record of success (or shortcomings).

The GAO has used inspection, enforcement, and safety data extensively in its evaluations. At times, the GAO evaluates the quality of the data, but in many cases the GAO uses the data to analyze the performance of the safety oversight system. The GAO has used Program Tracking and Reporting System (PTRS), Work Program Management System (WPMS), results of National Aviation Safety Inspection Program (NASIP) inspections, data from the National Air Transportation Inspection (NATI) program study, inspector training records, Accident Incident Data System (AIDS) data, Vital Information Subsystem (VIS), and Enforcement Information System (EIS) to get insight into the performance of the safety oversight system. The GAO also uses SPAS to access relevant data.

Interviews and surveys are another significant method the GAO uses to collect information and assess the safety oversight system. The GAO has interviewed or had discussions with FAA managers and members of FAA’s workforce, members of industry, academia, government agencies, and other experts. The GAO has also conducted structured surveys of FAA inspectors and managers. The GAO uses original analyses of these surveys and interviews to support observations about the effectiveness of the safety oversight system.

Table 2: Summary of Methods Used by GAO to Evaluate the FAA Safety Oversight System

Method	Description	Example of GAO's Use of Method
1. Program document review	Review of policy and program documents including internal memos, legislative acts, and program notices	Discover the nature of the program; purpose of program; extent of program guidance
2. Interviews with workforce, managers, air carriers, outside experts	Semi-structured interviews and discussions with various subjects; usually selected across regions and field offices as well as from headquarters	Discover the nature of the program; extent of guidance; the strengths and weaknesses of the program
3. Studies of accident investigation reports, NTSB comments	Review of NTSB and other accident investigation reports and NTSB comments	Role of safety oversight in preventing or contributing to accidents
4. Review of prior evaluations by GAO, IG, internal reviews by FAA	Review of GAO, Inspector General, or other evaluation documents or documents from special inspection programs	Do problems identified in the past still remain? To what extent have prior problems been corrected?
5. Original analysis of data including inspection records, training records, certification records, enforcement records, and data on safety outcomes	Reviews include special inspections such as NATI and routine surveillance, follow-up inspections, for all types of inspections	Comparison of inspection rates across carriers, regions, field offices
6. Structured surveys of workforce and managers	GAO conducts surveys of a sample of inspectors and managers using a formal survey instrument; often by mail	Used to gather information on inspectors' role in the enforcement process
7. Review of prior research	Review of academic or other research on safety, organizational factors	Justification for use of specific risk precursors such as financial health of air carriers

The methods used by the GAO have changed over the decades. In the 1980s, the GAO reports presented information collected through discussions with FAA managers and analyses of FAA and NTSB databases. Since about the middle 1990s, GAO reports have relied on large surveys of inspectors and other members of the FAA’s workforce, as well as discussions with FAA managers. Contemporary GAO reports are also more likely to make use of analyses of FAA databases and more explicit measures of performance. A third characteristic of the GAO’s

contemporary methodology is the explicit recommendation to use management controls and evaluations to improve the effectiveness of the safety oversight system.

The GAO's methods of collecting and analyzing information about the safety oversight system to evaluate its performance are tied ultimately to the GAO's indicators of an effective program. The GAO's indicators of effectiveness are based on the methods available to support its analysis and its findings. Table 4 presents a description of indicators of the effectiveness of safety oversight, or sometimes the deficiencies in safety oversight, that the GAO used between 1971 and 2005. Again, these indicators were gleaned from 57 evaluations published during this period. These 17 indicators represent what we considered classes of indicators used by the GAO, and each of these classes of indicators can be operationalized or represented by a variety of metrics.

The GAO's logic and methods can be read through these indicators. One of the GAO's basic questions is: Does the FAA have programs in place that are specifically designed and dedicated to providing safety oversight? The GAO typically reviews program documents to learn about the existence of programs and their scope. The GAO also considers specific program elements such as training programs, management oversight, and program guidance because the effectiveness of the parts of the program affects the effectiveness of the whole program. An extension of this idea is that effective programs should have the internal capacity and awareness to adjust to meet program goals.

Many of the GAO's measures of effectiveness focus specifically on the surveillance elements of safety oversight. Effective surveillance programs should have a record of discovery and that record should be archived and usable. Effective surveillance programs also take the opportunity to make corrective actions when that opportunity exists. When the FAA surveillance program is effective, FAA managers and its workforce should have knowledge of the extent of the surveillance activities and the extent of the air carrier operations that are the subject of surveillance. Knowledge of the extent of what needs to be monitored leads directly to developing a surveillance plan that has sufficient coverage. Coverage should also be influenced by knowledge of specific risk factors that can target coverage to the areas of highest risk. Given these factors, an effective surveillance program should collect the evidence necessary to intervene with enforcement actions if they become necessary. In addition, enforcement actions are effective when they deter hazardous activities or behaviors.

These indicators of effectiveness are not applied in each evaluation, nor does each one appear in every GAO evaluation report, but collectively they describe the logic of the GAO's evaluation methodology. Many of these indicators are extended with specific measures of effectiveness, such as the number of empty seats in a training course. The exact measures that the GAO uses depends on the methods and information available.

As the conclusion of each evaluation, the GAO suggests changes for improvement to the agency and then the agency provides comments either accepting or disagreeing with the GAO's statements. In more recent reports, the GAO's evaluation methods are described in detail in an appendix to the report.

Table 3: Indicators of Effectiveness Used by GAO to Evaluate the Performance of the FAA Safety Oversight System

Indicator	Description	Example of GAO's Use of Indicator
1. Existence	Existence of inspection or other programs; programs should have specific purposes	Does a program exist that is designed to serve its specific, safety-related purpose?
2. Record of discovery	Effective programs have a recorded history of identifying safety problems	Was correspondence with repair stations about corrective actions on file?
3. Opportunity	Effective programs identify and take opportunities to intervene and implement corrective actions or strategies	Did the entity have the opportunity to create safe standards and did that entity take that opportunity?
4. Extent	An effective surveillance program knows the extent of (a) a safety problem within a population of carriers, (b) the extent of surveillance coverage, (c) the extent of technological advance of industry and their systems	Carriers' use of repair stations
5. Enforcement	Effective surveillance programs have enforcement that deters bad behavior	Do carriers absorb fines or punishment as a cost of doing business or a substantial cost that it should avoid?
6. Evidence	Effective surveillance programs develop the necessary facts to substantiate a case	Was an enforcement case closed without action because there was not enough evidence to support it?
7. Coverage	Effective surveillance programs have the right surveillance coverage	Does each air carrier receive inspections of each type, per minimum requirements?
8. Program elements	Effective surveillance programs have effective program elements including: standards, training, guidance, priorities, and evaluation elements	Utilization of training programs; program guidance should have the right level of detail
9. Selection criteria / risk factors	Effective surveillance programs have selection criteria, including but not limited to risk factors, to target inspections	Rapid growth; financial health
10. Internal capacity to adjust, meet program goals	Effective surveillance programs have the internal capacity to adjust and meet program goals	Hiring new surveillance employees; training in new skills
11. Documentation and data	Effective surveillance programs elaborate findings and archive information	Detailed surveillance records and archives of these records; need to be accessible, useful
12. Corrective actions / strategy	Effective surveillance programs identify corrective actions and have strategies to help carry them out	Aging aircraft
13. Continuing neglect	Presence of deficiencies that are a result of continuing neglect	
14. Identification without resolution	Identification of specific safety deficiencies without resolution or correction	
15. Variation	Is there variation in a program that makes it more or less effective depending on who carries out the program	Differences among regions; inspector discretion (should discretion and variation be different)?
16. Carriers' level of compliance	What is the level of compliance of each carrier?	A carrier that was in violation had satisfactory inspection results 76 percent of the time
17. Under identification	Surveillance does not detect violations or safety deficiencies	Routine vs. special or team inspections

4. FAA Safety Oversight Evaluation Challenge

4.1 Background and motivation

The philosophy that guides the FAA oversight system has evolved over the last two decades, from setting standards and ensuring compliance with those standards to a more elusive concept, “system safety.” Major elements of the system safety philosophy include the allocation of oversight resources on the basis of risk, the need for airlines to have operating systems in place to identify and mitigate hazards and risks, and the promotion of safety culture in which workers throughout an organization strive to increase safety.⁸³

At the FAA, this approach to safety oversight is most clearly manifested in the Air Transportation Oversight System (ATOS) that AFS uses to oversee the largest passenger airlines. ATOS was conceived during a 90-day review of the safety oversight system requested by Administrator Hinson in the aftermath of the ValuJet crash in the Everglades in 1996. Under ATOS, teams of inspectors are assigned to each airline. Each team develops a comprehensive surveillance plan that “includes a series of inspection tasks to determine whether the airline has systems in place to ensure safety and a second series of inspections to verify that the airline is actually using those systems.”⁸⁴ The plan is developed in a structured fashion that takes into account inspectors’ assessments of the internal safety management systems, safety performance history, operational stability, and operating environment.⁸⁵ Inspection tasks are also heavily structured; inspectors complete detailed checklists indicating whether the airline is satisfying requirements for having systems in place for managing safety, known as Safety Attribute Inspections, and for actually using those systems, termed Element Performance Inspections.

The FAA implemented ATOS for the 10 largest passenger air carriers in October 1998. While external observers were generally supportive of ATOS as a concept, there has been widespread criticism of its implementation. A 2002 report by the Department of Transportation Inspector General identified three areas of weakness: incomplete development of key ATOS processes, in particular inspection data analysis; inspector training and deployment; and lack of

⁸³ U.S. General Accounting Office. “Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks.” GAO/RCED-98-6. February 1998.

⁸⁴ *Ibid.*

⁸⁵ *Ibid.*

national authority leading to inconsistent procedures across regions.⁸⁶ Many inspectors have been dissatisfied with the program because the time required to implement it resulted in curtailed surveillance activity. The NTSB report on the crash of Alaska Airlines flight 261 cited deficient oversight, which inspectors attributed to the ATOS transition, as a contributing factor.⁸⁷

Despite the difficulties with ATOS, the FAA remains committed to the system safety approach and has established several programs to extend system safety concepts to its oversight of other segments of the aviation system. The Certification, Standardization, and Evaluation Team (CSET) program, for example, applies such concepts to the process of certifying new entrant air carriers or when an existing carrier requires a reissued certificate, while the System Safety Approach for General Aviation (SAGA) program does the same for general aviation, beginning with agricultural operators.⁸⁸ On a more ambitious scale, AFS has initiated the System Approach to Safety Oversight (SASO) program, whose aim is to “transform the AFS and the aviation industry to a national standard for safety risk management.”⁸⁹ The SASO program will include improved decision support and training for inspectors, incentives for industry to implement safety management programs, and enhanced systems for acquiring, storing, sharing, and analyzing safety data.

In this section, we examine the motivations for and obstacles to safety oversight system evaluation. We then discuss the challenges, some generic and some unique, to evaluating aviation safety oversight activities. We also present research on how safety oversight evaluation is conducted outside of aviation.

4.2 Why Evaluate the Oversight System?

Our goal is to develop a methodology for evaluating the safety oversight system of FAA. An evaluation methodology should provide AFS and FAA management with the ability to

⁸⁶ U.S. Department of Transportation, Office of Inspector General, “Oversight of Aircraft Maintenance, Continuing Analysis and Surveillance Systems,” AV-2002-066, December 12, 2001.

⁸⁷ National Transportation Safety Board, *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*, NTSB/AAR-02/01, December 30, 2002.

⁸⁸ U.S. General Accounting Office. “Aviation Safety: FAA’s New Inspection System Offers Promise, but Problems Need to Be Addressed. GAO/RCED-99-183. June 1999.

⁸⁹ U.S. Federal Aviation Administration. “AFS-900, System Approach to Safety Oversight—Mission Need Statement.” March 2001.

analyze the health and performance of the aviation safety oversight system. It will help managers determine whether the oversight system is functioning properly and whether it has the desired effects on aviation safety. A sound evaluation methodology should be a basis for an evaluation system that supports management in its efforts to attain continuous improvement through refining oversight practices and re-allocating oversight resources. Further, an evaluation system based on a sound methodology will support AFS and FAA participation in assessments of the oversight system conducted by external agencies such as the IG, the GAO, and the OMB. Ideally, an internal safety oversight evaluation process may reduce the need for such assessments by enabling internal management to see and address problems before they become serious enough to attract the attention of external reviewers.

The logic for developing an evaluation system is not much different from the logic for other evaluation activities underway throughout the federal government. Many of these evaluation activities have roots that extend back several decades. Patton traces the beginning of modern-day evaluation research to “the massive federal expenditures on an awesome assortment of programs during the 1960s and 1970s” when “accountability began to mean more than assessing staff sincerity or political headcounts of opponents and proponents.”⁹⁰ The need for such research derived from two factors: that there was not enough money for government to do everything that the vision of the Great Society might call for, and that more than money was required to solve complex human and social problems. These circumstances, combined with Kennedy-era belief in the ability of science to find solutions to social problems, triggered a spate of evaluation studies featuring rigorous experimental design, data collection, and causal modeling.

The federal aviation safety oversight program predated the Great Society by nearly half a century, and developed at a time when the questions that spur modern-day evaluations were not widely asked. For the most part, Flight Standards was what James Q. Wilson termed a “procedural organization” in which “managers can observe what their subordinates are doing but not the outcome (if any) that results from those efforts.”⁹¹ Aside from that, evaluation of the safety oversight system rested on the judgments and ideologies of senior managers, presidential

⁹⁰ Michael Quinn Patton, *Utilization-Focused Evaluation*, Beverly Hills: Sage Publications, 1978.

⁹¹ James Q. Wilson, *Bureaucracy: What Government Agencies Do and Why They Do It*, Basic Books: New York, 1989.

administrations, and Congress. In making these judgmental assessments, decision makers were informed by the statistical safety record, results of accident investigations, feedback from the aviation community, and episodic inquiries and studies spurred by catastrophic accidents, occasional scandals, and the continuing stresses and strains created by rapid growth and technological change in commercial aviation.

With the Clinton administration's National Performance Review (NPR) and contemporaneous passage of the Government Performance and Results Act (Results Act), scientific evaluation received new impetus, extending its reach to virtually every corner of government. In the cause of making government "work better and cost less" these initiatives sought to "refocus managerial attention on outputs rather than inputs."⁹² The Results Act requires all government agencies to prepare annual performance plans focused on achieving measurable progress toward program goals. In order to implement the Results Act, the OMB has developed the Program Assessment Rating Tool (PART) to assess "how program evaluation is used to inform program planning and to corroborate program results."⁹³ Furthermore the OMB has reorganized so that budgeting and program planning decisions are made in an integrated fashion. Thus, the ability to measure outputs has become an important factor in the ability of programs to obtain budgetary inputs.

In addition to the government-wide push toward program evaluation, factors more specific to aviation safety oversight and to AFS motivate and shape the effort to develop an evaluation methodology. First, AFS's safety oversight program has been continually criticized by external evaluators, including the GAO, the National Transportation Safety Board (NTSB), Ralph Nader, and former DOT Inspector General Mary Schiavo.^{94, 95} The latter two have written popular books that give substantial attention to the subject, while the GAO has issued dozens of critical reports on aviation safety oversight beginning in the 1970s. Second, much of the FAA has recently been transformed into a performance-based organization. In 1997, the

⁹² Donald F. Kettl and John J. DiIulio, Jr., *Inside the Reinvention Machine: Appraising Governmental Reform*, The Brookings Institution: Washington, D.C. 1995.

⁹³ U.S. Office of Management and Budget. "Assessing Program Performance for the FY 2005 Budget." <http://www.whitehouse.gov/omb/part/>. Accessed July 31, 2004.

⁹⁴ Ralph Nader and Wesley J. Smith, *Collision Course: The Truth About Aviation Safety*, Blue Ridge Summit: TAB Books, 1994.

⁹⁵ Mary, Schiavo, *Flying Blind, Flying Safe*, New York: Avon Books, 1997.

National Civil Aviation Review Commission (NCARC) recommended “establishment of a Performance Based Organization (PBO) within the FAA for the development, management, and provision of air traffic services” that “would be held accountable by committing to specific measurable goals with targets for improved performance. In exchange, the PBO is granted managerial flexibilities.”⁹⁶ This resulted in the establishment of the FAA Air Traffic Organization (ATO) in 2003. While AFS is not part of ATO, it is no doubt influenced by ATO’s commitment to performance management and measurement.

Finally, the same thinking that has led AFS to adopt the system safety approach also creates the impulse to measure oversight system performance. System safety has its roots in the Total Quality Management philosophy that stresses the minimization of production variation and risk, participative management, continuous improvement, and ongoing performance monitoring. In undertaking an evaluation system, AFS is practicing what it preaches to the industry it regulates.

4.3 Challenges

It is challenging to create indicators that can tell us what the safety oversight system contributes to aviation safety. One way to answer this question is to consider the influence of the safety oversight system on safety outcomes. Safety outcomes could include airplane crashes with fatalities, but outcomes could be extended to include risk precursors and initiating events such as engine failure or insufficient operator training. Or perhaps there are other conceptualizations of safety that may yield meaningful measures that are easier to obtain. In any case, in developing an evaluation methodology we face several significant challenges.

4.3.1 Experimentation Is Infeasible

To answer the question about how much safety the oversight system contributed to safety outcomes, one would like to compare two sets of air transport service, one with a safety oversight program and the other without safety oversight. An experiment similar to this has happened only once. In the early 1920s, before the federal regulation of commercial carriers, the

⁹⁶ National Civil Aviation Review Commission. “Avoiding Aviation Gridlock and Reducing the Accident Rate: A Consensus for Change. <http://www.faa.gov/NCARC/reports/pepele.htm>. Accessed July 31, 2003.

federally operated U.S. Air Mail service had a safety oversight program that included regular medical exams for pilots, careful aircraft inspection, a 180-item checklist used at the end of each trip, and regular engine and aircraft overhauls. The investment in the safety program was significant: the ratio of mechanics to aircraft was nearly four to one and 94 percent of air mail service employees were ground personnel. The safety benefits of this investment were clear: the fatality rate for the U.S. Air Mail service was one per 789,000 miles flown from 1922 to 1925, while the comparable figure for itinerant commercial fliers (for 1924 only) was one per 13,500 miles.⁹⁷

Nowadays, controlled experimentation with the safety oversight system is unacceptable. The consequences of an accident are too great for experimentation. Besides, in a competitive environment, the FAA should regulate equitably. In lieu of a controlled experiment, we must use quasi-experimental methods that estimate the effects of “naturally occurring” variation in oversight activities. For example, an investment analysis of the SASO program prepared by the FAA compared accident rates and inspection hours per aircraft for large jet operators, commuter carriers, and general aviation. Not surprisingly, they are inversely related: large jet operators are inspected more intensively and have lower accident rates. But there are many other factors besides safety oversight that contribute to these differences in safety outcomes.⁹⁸ Another quasi-experiment, reported in the SASO Mission Need Statement, is a comparison by the USN Navel Safety Center, which reports that two aircraft—the F/A-18 and the A-7—designed using system safety principles had accident rates 60 and 80 percent lower than equivalent aircraft—the F-14 and F-4—whose design did not use this approach.⁹⁹

4.3.2 Accident Numbers Are Small

The outcomes that we care about most, the crashes, fatalities, and injuries, have declined significantly since the early years of the industry. Air travel, particularly travel on large scheduled air carriers, is a very safe activity. In 1998, Arnold Barnett and Alexander Wang

⁹⁷ Nicolas Commons, *Bonfires to Beacons: Federal Civil Aviation Policy Under the Air Commerce Act, 1926-1938*, U.S. Department of Transportation, Federal Aviation Administration, 1978.

⁹⁸ Clinton V. Oster Jr., John S. Strong, and C. Kurt Zorn, *Why Airplanes Crash: Aviation Safety in a Changing World*, Oxford University Press: New York, 1992.

⁹⁹ U.S. Federal Aviation Administration. “AFS-900, System Approach to Safety Oversight—Mission Need Statement.” March 2001.

estimated the passenger death risk. They found that the death risk per scheduled jet flight over 1987 to 1996 for all first-world airlines, inside and outside the U.S., was about one in 8 million. Another way to say this is that an individual could take one flight per day for 21,000 years before becoming a fatality statistic.¹⁰⁰ Moreover, more recent work by Barnett reveals that variation in 10-year fatal accident rates among major airlines may plausibly be attributed to chance alone.¹⁰¹ If we cannot reject the hypothesis that all large airlines are equally safe, conclusive statistical results about the effect of safety oversight on large airline safety are virtually impossible. Instead of measuring safety outcomes with accident and fatality rates alone, one could characterize safety using precursors to fatal accidents. Using this method creates a longer causal chain, introducing new uncertainties into the model. To create the first part of the model, one would need to establish how safety oversight affects safety precursors. Then, one would need to identify how those precursors affect safety outcomes such as crashes or fatalities. There is little empirical literature exploring the relationship between precursors and accidents in this context. One recent study by Barnett and Wang discussed the use of precursors (mishaps):“...data analysis fails to support the conjecture that, the greater an airline’s involvement in mishaps, the greater its propensity to suffer the disasters that passengers fear.”¹⁰² They proposed an explanation for their finding, “Suppose that all airlines suffer emergencies at the same rate, but that some are more adept than others at resolving them without dire consequences. Then the more skillful airlines will have relatively few disasters but relatively more nonfatal mishaps, while the less-adept carriers will exhibit the opposite pattern.”¹⁰³ Of course, organizational safety literature suggests that there is a link between precursors and accidents, and that precursors must necessarily be controlled and prevented, which is consistent with the practices of system safety.

¹⁰⁰ Arnold Barnett and Alexander Wang, “Airline Safety: The Recent Record,” NEXTOR Research Report RR-98-7. May 1998.

¹⁰¹ Arnold Barnett, “The Mother Metric? FAA’s New Safety Index,” Moving Metrics: A Performance-Oriented View of the Aviation Infrastructure., Asilomar Conference Center; Pacific Grove, California, January 27-30, 2004.

¹⁰² Arnold Barnett and Alexander Wang, “Airline Safety: The Recent Record,” NEXTOR Research Report RR-98-7. May 1998.

¹⁰³ *Ibid.*

4.3.3 Safety Production Is Not Well Understood

In contrast with the role of the federal government in the U.S. Air Mail service in the example given above, the current FAA safety oversight system does not carry out maintenance, training, or operations directly. The safety oversight system can influence safety only indirectly through certification, rulemaking, rule enforcement, the threat of enforcement, surveillance, and the collection of information through accident investigation. While the regulatory environment can have a major influence on safety, the FAA oversight system achieves its goals only by influencing the activities and production practices of aircraft manufacturers, airlines, labor unions, and other stakeholders and not by flying and fixing aircraft. The effectiveness of the safety oversight system depends on the strength and effect of its influence on these other activities.

Instead of focusing all energy on constructing measures that relate oversight and safety outcomes, most of the de facto performance measures of the safety oversight system set that relationship aside and concentrate on more accessible measures of performance such as the distribution of inspections across airlines over time, whether the FAA carried out its plan for inspections, and whether inspectors have the appropriate training. There are a number of functions that the safety oversight system carries out to produce its broader mission of ensuring safety. Some of these functions, such as surveillance, are more accessible in terms of performance measurement. For example, the surveillance function comprises observable and recordable activities such as performing inspections, researching operator information, and training. Performance measures can be formed using information such as number of inspections, number of inspector hours, number of hours training, number of satisfactory or unsatisfactory inspections, the regulations violated, enforcement measures (e.g., fines), and so forth. If the production of safety were understood better, that is, if we knew more about how firms in the aviation system adapt to the FAA's safety oversight activities, then we could examine the influence of these observable safety oversight activities on the inputs into safety (e.g., use of pilot checklists, use of maintenance checklists, internal evaluation within air carriers), rather than focusing on crashes and fatalities.

4.3.4 Inspection Results Are Difficult to Interpret

Relying on data from surveillance activities also presents a complication for analysis. The results of inspections can be difficult to interpret. Imagine the case in which the surveillance system finds many violations. Discovering the violations could mean that the inspection program is working very well, that the inspections are targeted to areas with the greatest risk, and that the inspectors are skilled and are reporting findings consistently. But the scenario could also point to problems in the oversight system because at some prior time the oversight system allowed underlying problems to persist. Thus, the health of the oversight system could be poor, even if the surveillance program is healthy.

Similarly, imagine another scenario in which the surveillance system finds few violations of FARs. This scenario could mean that the surveillance system is healthy and that the oversight system is healthy. But the situation could also indicate that the surveillance program is not targeting inspections to areas that exhibit weakness.

There is a third explanation of this scenario, one that has been documented by the GAO. A report from 1998 states, "...inspectors do not consistently report violations. Many of those we interviewed and surveyed volunteered that they handle many violations informally and, if compliance can be achieved on the spot, may not enter violations into their tracking system."¹⁰⁴ Relying on inspection outcome data to measure the health of the oversight system reveals only behaviors that have been recorded. If inspectors sometimes find that reporting all behaviors is a hindrance to achieving safer outcomes, then the inspectors may choose to lock-in the safer outcome instead of using the inspection procedure to reach the same result.

4.4 Experience in Other Sectors

To gain perspective, and with the hope of finding an existing system that may have surmounted the challenges identified above, we examined safety oversight systems and their evaluation in other sectors including rail and nuclear power generation. Both sectors have well-developed safety oversight systems, although the specific oversight processes are different for each sector. Although the processes in each sector are different, safety engineers, managers, and

¹⁰⁴ U.S. General Accounting Office. "Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks." GAO/RCED-98-6. February 1998.

policy makers learn from the experiences of all sectors. Each sector relies on the cumulative experience across sectors in human and organizational factors, ergonomics, defense-in-depth, system safety, and risk assessment to carry out their oversight missions. In a similar way, external reviews of safety oversight systems by the GAO and the OMB often pose similar questions about whether regulations and surveillance activities are uniformly (or otherwise justifiably) applied, whether plans for oversight are carried out, and whether oversight system managers are improving safety oversight. As a result of the communication across sectors, and the similar training of safety managers, many safety systems and safety oversight systems exhibit similar designs that differ mainly because of differences in regulatory mandates and organizational characteristics.

Although rail and aviation industries face similar challenges with respect to the interpretation of inspection findings and multiple models of safety production, oversight evaluators in rail do not have the small numbers problem. The safety record shows consistent improvements in safety, as measured by the rates of accidents and fatalities. But the railroad industry has more observations with which to measure the effects of its oversight activities.

Since 1996, the Federal Railroad Administration (FRA) has used a collaborative approach to oversight in which teams comprising regulators and representatives from both railroad management and labor unions assess the safety needs of the railroad and set to work addressing the needs. The FRA named this approach the Safety Assurance and Compliance Program (SACP). The FRA monitors progress on the implementation of the safety plan and considers the completion of the plan and its achievements as indicators of successful safety oversight. This collaborative approach complements, but does not substitute for the traditional rulemaking, surveillance, and enforcement method of oversight, which the FRA uses as well.

The Office of Management and Budget found, as part of its standardized performance review of programs, that the Federal Railroad Association's safety oversight system needed a formal evaluation process, but that it did not have one at the time of the assessment.¹⁰⁵ But even without a formal process for assessing the health and performance of the safety oversight system, the FRA adjusted its oversight system "in gradual shifts" and with a "best practices

¹⁰⁵ U.S. Office of Management and Budget. "Department of Transportation PART Assessments." <http://www.whitehouse.gov/omb/budget/fy2005/pma/transportation.pdf>. Accessed July 31, 2003.

approach” to problem solving.^{106, 107} The FRA’s newest approach to oversight, the collaborative approach, may allow for informal evaluation to occur regularly. The participants in the evaluation include both the regulator and representatives from industry and labor and one of the primary outputs of the evaluation is increased communication among these stakeholders. Safety oversight in nuclear power production relies on the traditional method of oversight. The Nuclear Regulatory Commission (NRC) collects operating data, in the form of performance indicators, from each nuclear reactor. Utilities self-report these indicators to the NRC each quarter and the NRC uses its surveillance program to verify some of the measures. The NRC also measures activities that are not self-reported such as whether a utility has a safety-conscious work environment.

Just as the NRC relies on operations and safety performance indicators to carry out its oversight program it also uses performance indicators to assess the performance of the oversight system. Data for these performance indicators are collected from oversight operations records (number of inspections, date of inspection, et cetera) and through surveys of stakeholders. The NRC uses over 70 performance measures to characterize the performance of the safety oversight system. The performance measures assess the oversight system’s (1) performance indicator program, (2) inspection program, (3) significance determination process (the process through which the NRC determines whether a utility’s operations, as measured through the performance indicators and inspection findings, are safe enough or whether they need more oversight), (4) assessment program, and (5) communication activities and other program issues. A few illustrative self-assessment performance indicators include: the percentage of inspection reports that find the utility is in compliance with their program, the number of significant changes made to the inspection process, the completion of planned oversight activities, the timeliness of oversight activities, the predictability of the significance determination findings, and the accurate communication to the public of oversight findings.¹⁰⁸

¹⁰⁶ U.S. Federal Railroad Administration, “Five-Year Strategic Plan for Railroad Research, Development, and Demonstrations,” March, 2002.

¹⁰⁷ Jolene M. Molitoris, “Statement of Jolene M. Molitoris,” *Testimony before the Senate Committee on Commerce, Science, and Transportation Subcommittee on Surface Transportation and Merchant Marine*, February 25, 1998.

¹⁰⁸ U.S. Nuclear Regulatory Commission, “Reactor Oversight Process Self-Assessment For Calendar Year 2003,” SECY-04-0053, April 6, 2004

Although the NRC's oversight evaluation system is formal and the FRA's is relatively informal, both of these evaluation systems share an important characteristic. The evaluation systems in both industries open the lines of communication between the regulator and the regulated industry. In the context of nuclear power production, utilities and advocacy groups are regular and vocal participants in rulemaking and changes in the safety oversight system. For example, the NRC's self-assessment from calendar year 2003 described stakeholder survey responses on the topic of the performance indicators. The report states, "...responses...indicated that the public and the nuclear industry have varying views on the efficiency and effectiveness of the PI [performance indicator] program. The industry generally believed that the PI program was working well....By contrast, the public has become increasingly concerned that the PIs are being managed by the licensees and have become ineffective as indicators of plant performance."¹⁰⁹ In this example, the utilities believe that the performance indicators are a good model of how the utilities produce safety. According to them, if the performance indicators indicate safe operations, then the operations are probably safe. On the other hand, the representatives of the public interest have a different model of safety production, and they believe that other performance indicators would more accurately describe the relationship between reactor performance and safe outcomes. In this case, the self-assessment process collected information about competing models of safety production.

None of the self-assessment performance indicators in the NRC's evaluation system measure the safety oversight system's direct effect on safety outcomes, but they do address the issues that the NRC has direct control over, which is an important distinction. As illustrated in the example above, the NRC has control over which performance indicators it chooses, but its judgment will be influenced by guidance from stakeholders.

We approached the literature about other industries hoping to learn about the evaluation methodologies used there: the performance measures, the analytical frameworks, and the key components of oversight that reveal the health of the system. We found that other industries do use performance measures, and that the performance measures are not so different from those used by the external evaluators such as the GAO. We also found that none of the indicators link oversight activities to safety outcomes.

¹⁰⁹ *Ibid.*

As we reviewed oversight evaluation in these other industries, we began to learn how these industries use evaluations in practice. From these studies we see that the dominant characteristics of the industry, such as the positions of stakeholders or the preference for informal or formal communication methods, will be characteristics of an evaluation system. From these other industries we also see that the oversight evaluation systems are not independent from the structure of the oversight system: in addition to collaborative approaches, the NRC uses performance indicators to measure the performance of the utilities and its own oversight system; the FRA uses historical data and a collaborative approach to oversight as well as its oversight evaluation.

4.5 Lessons for a Safety Oversight Evaluation Methodology

It is reasonable to investigate the health and performance of the aviation safety oversight system. There is a strong public interest in maintaining and improving aviation safety; the public and its elected representatives are entitled to know the return on a program with a direct public cost approaching \$1 billion, and indirect cost to industry that is considerably higher. In addition, performance measurement is an integral part of performance management.

But evaluation of this system is extremely challenging. Experimentation is impossible and naturally occurring quasi-experiments are difficult to find. Small numbers of accidents, particularly in commercial aviation, while a blessing to all, pose a further hurdle to the evaluator. The analysis of incidents and precursors, while appealing from a statistical standpoint, is difficult to interpret. Inspection results are also difficult to interpret, particularly when inspectors, perhaps for good reason, often don't report violations. Moreover, as AFS re-orient its oversight program toward a system safety approach, measures of compliance, even if reliable, will become less relevant.

Underlying the challenges of creating performance measures of system safety is the question, what is a healthy oversight system and what makes it healthy? If we know the answers to those questions, then we know where to go looking for performance measures. Traditionally, the oversight system is thought to perform well if it has plans for oversight and follows those plans. One can go further and check to see that those plans are sensible: consistent regulation yet targeted to areas of greatest risk. But the system safety literature tells us that there is more to safety production than making and following plans and rules; procedural safety is as outmoded

as is a procedural organization. Inspectorates across sectors may have known this all along, giving rise to an inspector culture that encourages voluntary reporting and smooth relations between inspectors and the certificate holder. What are the signs that tell whether the oversight system is asleep at the wheel, or simply working quietly, subtly, but effectively? Does the NRC's annual report evaluating its oversight system with its 70 indicators really tell us more than the FRA's reporting that the railroads are carrying out their plans to improve safety?

It is certainly possible to create an oversight system evaluation methodology that allows AFS and FAA management more than they now know about how the oversight system is working. A methodology that relates safety oversight activities to safety production must combine a risk analysis framework and an organizational approach. We must strive to create evaluation metrics that capture the link between oversight activity and safety outcomes. But in pursuing this goal we will respect the delicate nature of aviation safety oversight and recognize the well-known fact that evaluation drives behavior. Our approach to assessing oversight system health, like that of the physician concerned with human health, must start with the principle, "first do no harm."

4.6 FAA Safety Oversight Evaluation and Monitoring

Ensuring that the FAA's surveillance and monitoring activities associated with its safety oversight program are effective requires application of effectiveness measures. These measures should be developed and evaluated frequently in order to facilitate implementing changes to improve the safety oversight process, with emphasis on surveillance and monitoring efforts. It is important that the measures used to evaluate safety oversight effectiveness focus on both the goal (e.g., fewer accidents/lower accident rates) as well as the process used to achieve the goal (e.g., number of inspections performed). The effectiveness measures should be comprehensive, and derived from a methodology that is based on systemic as opposed to chronic variables.

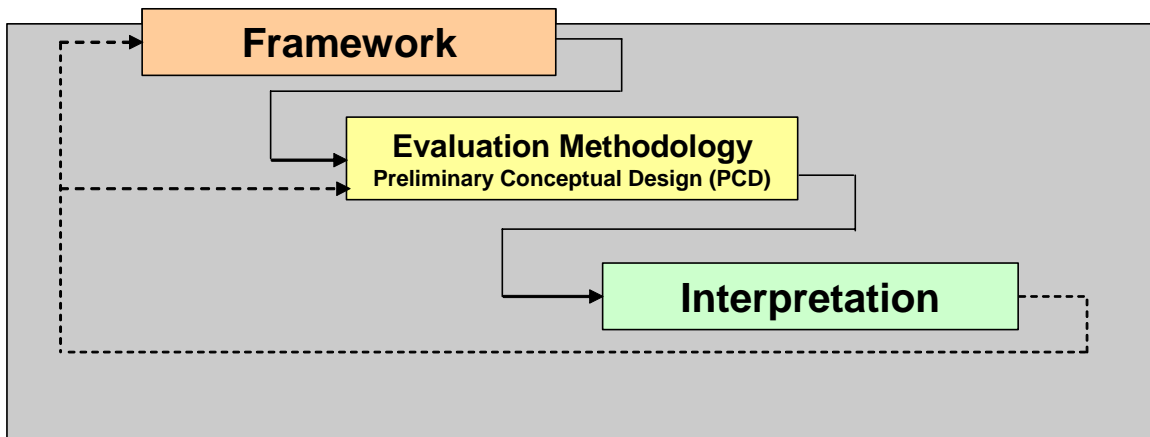
It is likely that no single effectiveness measure will prevail, but rather a group of measures that collectively may be used to evaluate the safety oversight from various perspectives. Contributing to this presumption is that accurate data may not always be available to create robust statistics for all measures. Accordingly, a collection of candidate measures are likely to have greater utility than a few, and may prove to be more accurate.

5. Generic Methodology to Evaluate the Effectiveness and Efficiency of System-Based Safety Oversight

5.1 Preliminary Conceptual Design of an Evaluation Methodology

The FAA needs a comprehensive approach to evaluate the FAA safety oversight system. The Oversight System Evaluation Methodology (OSEM) concept, depicted in Figure 5, comprises a top-level framework, an analysis tool, and a review process for stakeholders.

Figure 5: The OSEM Concept



Decision makers within the FAA need to know if oversight plans and policies are being faithfully executed, if the oversight system is effective in promoting safety and reducing risk, and if the oversight system makes efficient use of the FAA and Industry resources. A framework for an evaluation system needs to address potential visions for future oversight systems, as well as have an understanding of air carriers' internal safety oversight and monitoring systems, and when oversight information has been used in the past or might have been used. All these considerations need to be addressed under the umbrella of a system safety approach. Developing a methodology to evaluate the effectiveness of the current, as well as future, aviation safety oversight systems is a complex task. This overall effort will no doubt be time consuming and involve many stakeholders, information systems, regulations, etc. We suggest an incremental approach for this large overall effort; for the first phase, we began by developing a preliminary

conceptual design (PCD) for an evaluation methodology. One crucial concept needed for success of any oversight evaluation methodology that is not covered in detail in this description is the culture within the FAA. This concept is addressed elsewhere in this report.

5.2 How the Preliminary Conceptual Design Was Developed

The team of UCB and LMI developed the PCD idea and approach after many interactions with the FAA and selected aviation safety subject matter experts (SMEs). The actual PCD was developed in conjunction with inspectors, the FAA and additional aviation subject matter experts, systems engineers, and other analysts with relevant input.

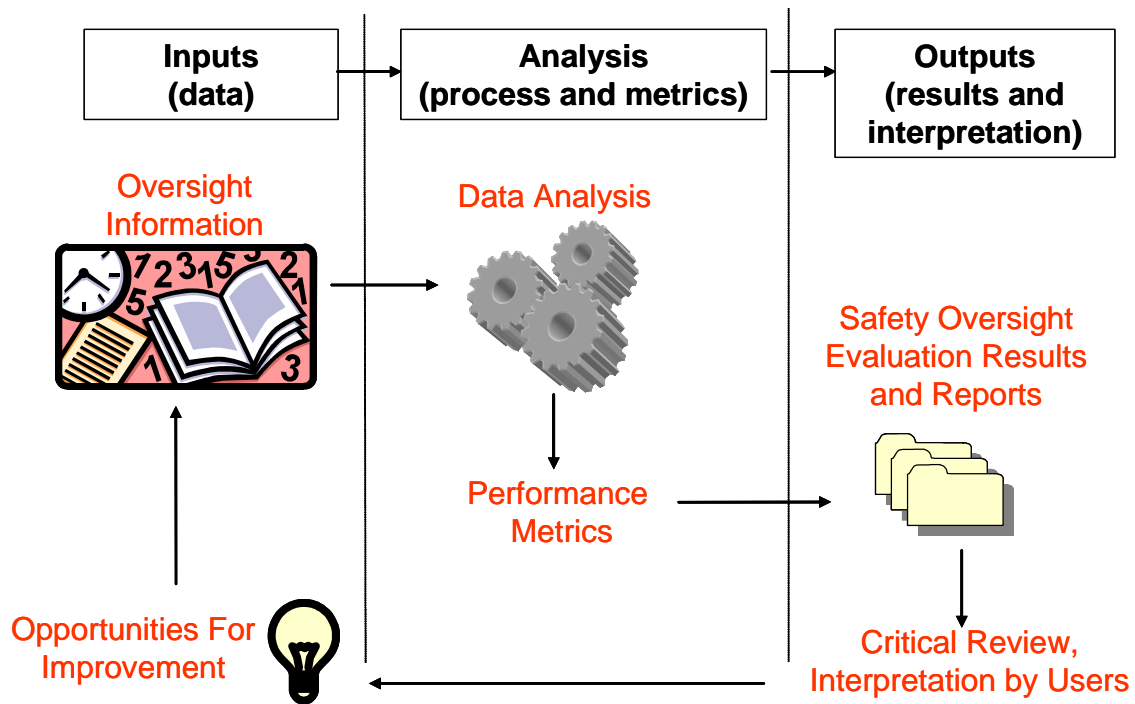
First, we developed a structure for the overall methodology to include a framework, analysis tool and the ability to enable interpretation of multiple data as well as provide an avenue for effective adoption of relevant feedback. Next, we constructed an interview guide around the PCD aimed at eliciting input from various actors within the aviation safety arena, each who provide different and important views into the many facets of this complex function. We validated the interview guide and approach with the same group of initial experts before conducting initial interviews throughout the aviation safety arena with many types of inspectors, inspection office managers, operation research analysts (ORAs), pilots, mechanics, and other aviation safety experts. We captured information about the FAA inspection environment, the pros and cons (from an inspector's viewpoint) of how the current oversight system functions as well as valuable lessons learned from the field. We captured additional specific analyses and data elements useful in providing a comprehensive picture of the overall health of the safety oversight system. These important data and input were then incorporated into the final version of the PCD. It is the hope of this team that this PCD will be further used to elicit additional specific feedback from a broader and higher level of management within the FAA, in order to ensure that the methodology addresses the concerns and issues of aviation safety stakeholders at all levels of government.

5.3 Description of the Oversight System Evaluation Methodology

This methodology enables a future analysis tool, or suite of tools, to be built in concurrence with a framework that guides any oversight evaluation and, with interpretation by users, facilitates the continual improvement of the eventual evaluation system.

The PCD answers the question “how does the OSEM Analysis Tool work?” in the context of the overall OSEM concept. It depicts the various components and process flows, and is preliminary because it is our initial representation to generate feedback leading to requirements analysis. Figure 6 depicts this methodology and the pieces that make it effective, whether they are electronic or human.

Figure 6: Preliminary Conceptual Design



For an evaluation tool for aviation safety oversight to be effective, many people and systems must interact. In particular, the users of the system may be both inspectors and their managers as they assign inspectors to various areas of the aviation system and maintain sufficient training programs and recruitment efforts; the ORAs who assist larger inspection offices in identifying targeted inspection opportunities or effective sampling inspection techniques; as well as headquarters FAA personnel and policy makers who require an overarching or broad view of the entire aviation safety system and how it is working on all levels (including staffing/budget

requirements, overall aviation safety risk assessment efforts, and potential implications outside the FAA - interaction with external stakeholders and decision makers.) Since the SASO program office is chartered to implement a System Safety approach within the FAA, a systems based approach to addressing aviation safety will necessarily involve personnel at all levels throughout the system. Air carriers also would be valuable participants in the OSEM process. The SASO program office continues to explore how the system safety concept will be realized within the FAA, so the evaluation methodology must remain flexible enough to enable new users and actors to be incorporated into the design of such a system over time.

5.4 Features of the Oversight System Evaluation System

5.4.1 Framework

In general terms, the Framework sets forth questions to guide the evaluation of the oversight system. The questions are based on the oversight system goals, its performance objectives, and addresses important issues about any oversight system to include:

- Efficiency - “does the safety oversight system use its resources to maximize effectiveness?”
- Effectiveness - “does the safety oversight system meet its objectives?”, and
- Sustainability - “is the safety oversight system structured to support future oversight activities?”

These broad concepts and issues provide the needed base for a robust and flexible evaluation system.

5.4.2 Analysis Tool

This tool, or suite of tools, provides useful analysis for users and stakeholders, based on the guiding questions in the framework. Both processes for data analysis as well as performance metrics are found in this portion of the methodology. There are multiple inputs that are essential for the vast amounts of analysis required by the many users of such a system. They include, but are not limited to, data regarding surveillance and safety activities (including partnership programs), communication, trends in the aviation environment, and internal and external audit

reports. Some example metrics found in the analysis portion of this methodology may include external reviews, corrective and preventive actions, safety precursors, inspection output (including safety recommendations), inspector workload and readiness. The analysis tool would also perform processes such as monitoring and evaluating incoming data and assess the quality of the data, prepare selected canned analysis reports or charts, perform comparative analyses, ascertain inspector relationship with a carrier, and identify opportunities for improvement that are evident based on data. The outputs of such an analysis tool should provide desired answers for users and stakeholders. Some results that may be important include, but are not limited to, answers to the following questions:

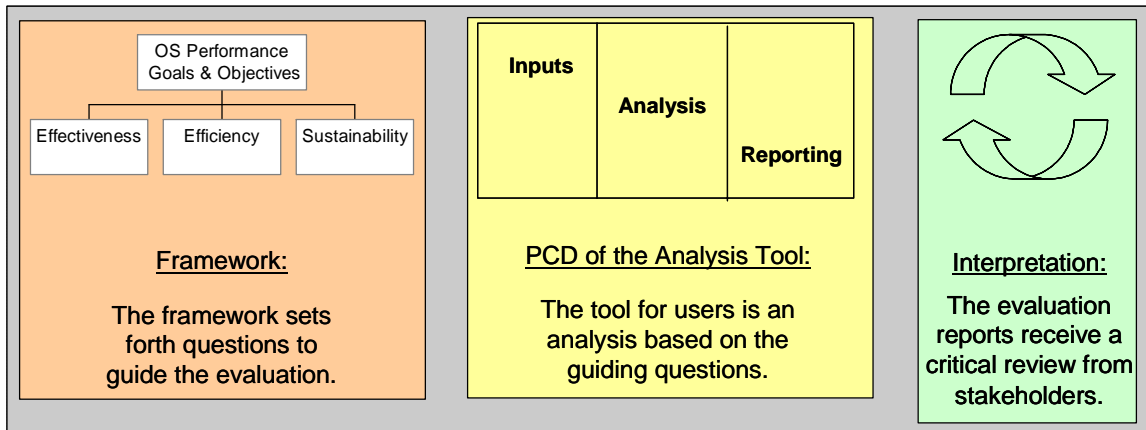
- Is the oversight system effective and efficient?
- Did the existing oversight system identify safety improvement options?
- Did the oversight system reduce unsafe conditions during the past year/quarter?
- What components of the oversight system are no longer needed? What new components are now needed?
- Is the oversight system responsive to changes in service, technology, and the economy?

5.4.3 Interpretation and Feedback

This portion of the methodology employs humans-in-the-loop to critically review evaluation reports and other output from the analysis branch of the methodology. These reviews not only provide key personnel with the information they need to make decisions relating to aviation safety oversight and the assessment of the health of the oversight system, but also to provide useful and crucial feedback to the framework in order to allow the evaluation system to change and mature over time. Undoubtedly, users and stakeholders will have new and changing questions about the current oversight system and how well it operates, and the evaluation methodology we propose enables these issues to be addressed as conditions continually change.

The following figure (Figure 7) shows the relationship between the Framework, Analysis Tool, and the Interpretation/Feedback portions of the methodology.

Figure 7: Components of the Oversight System Evaluation Methodology



Interactions between actors and system, interface between the OSEM and external systems

Top and senior-level management support is necessary to lead cultural acceptance that will promote effective implementation of the OSEM. Roles, authorities, responsibilities, protocols, and procedures must be clearly documented and understood by all players. Systems users, data analysts, and software engineers must collectively communicate and exchange information on lessons learned in order to continuously improve the utility and operability of the OSEM. Interviews during this study yielded responses from all levels of potential users expressing interest in such an evaluation system. Interviewees, representing a wide variety of users from inspectors to FAA headquarters personnel, requested various levels of information and analysis results available to them. Detailed information on the interactions among the various personnel and elements associated with current FAA oversight initiatives, including observation of field inspections and examination of databases, was not available during the study period of performance. Future study efforts will benefit from this information and greatly enhance establishing a methodology for evaluating the effectiveness of FAA oversight.

5.4.4 Primary Data Collection

A wide range of data collection techniques may be used to populate the database necessary to initiate the OSEM; however, timely, accurate, and unbiased data are essential to its utility. Time series and cross-sections data, whether obtained from discrete field observations, or various sampling techniques, must be collected with rigor and be well documented to establish traceability.

5.4.5 Analysis Methods

It is expected that a variety of quantitative and qualitative techniques would be applied to the OSEM database. These would include correlation, regression, trend analyses, and multi-attribute and simple comparative analyses. Collectively, the results of these kinds of analyses will provide input to corrective and preventive actions, identification of precursors to unplanned events, and planning the scope and schedule of inspections. It is important that the discrete data used to support the various analytical findings is identified to the specific results to facilitate comparing subsequent analyses as the database evolves and grows over time. This will help determine whether lessons learned are being adequately addresses and the extent to which oversight is benefiting from continuous improvement initiatives.

As discussed, the framework is the set of questions that judge whether or not the oversight system is attaining its goals and objectives. There must be specific analysis methods and tools that can monitor the condition of the oversight system in each of these areas. While effectiveness, efficiency, and sustainability are all interrelated, some metrics and methods can be seen as more specific to one of the areas. For each framework question, we have identified a few sample criteria that could be used to answer that question. After identifying useful criteria, we identified potential inputs for each criterion.

The table below provides a few examples of criteria that can be used to determine how well the oversight system is performing. Most of these criteria could be further examined for geographic and/or temporal trends. In this way, an office with a positive criterion could share best practices nationally, and negative trends over time could be identified quickly. Stoplight charts for these criteria could be created for each office, allowing each office to better understand which areas of oversight activities need more attention.

Table 4: Sample Criteria to Evaluate the Safety Oversight System

Goal	Evaluation Criteria
Efficiency	<ol style="list-style-type: none"> 1. Inspector productivity 2. Quality of data 3. Organizational efficiency 4. External rating
Effectiveness	<ol style="list-style-type: none"> 5. Responsiveness to corrective actions 6. Safety outcomes 7. Responsiveness to safety recommendations 8. Ability of oversight workforce to assess risk
Sustainability	<ol style="list-style-type: none"> 9. Stability of oversight workforce 10. Ability to plan for technology change 11. Responsiveness to industry changes 12. Preparedness of workforce 13. Ability to complete planned inspections

Table 6 presents a non-exhaustive list of inputs that would be useful in evaluating each of the criteria. We then focus on one specific criterion and see how the input data and output information could be best utilized.

Table 5: Inputs to Evaluation Criteria

Required Inputs	Associated Criteria
Surveillance Activities	
Number of inspections performed	1, 13
Number of inspections planned	13
Completeness of inspection records	2
Number of risk-based inspections performed	8
Number of joint inspections	1
Work observed versus documents reviewed	8
Inspector opinion	13
Safety Activities	
Ratio of acted upon safety recommendations to total safety recommendations	7
Number of corrective actions	5
Number of overdue corrective actions	5
Days to complete corrective actions	5
Human Capital	
Number of inspectors	1
Years of relevant experience of inspectors	12
Ratio of inspectors with current training to all inspectors	12
Number of inspectors requiring re-training certification	12
Number of inspectors requiring new skills certification	12
% employees nearing retirement	9
Number of research and development employees	10
% employees trained in analysis methods	8
Safety Outcomes	
Number of accidents	6
Number of incidents ¹¹⁰	6

¹¹⁰ Number of incidents could include AIDS incidents alone or could include additional items such as turnbacks and operational errors.

External Audits and Reports	
Number of GAO and OIG audit criticisms	4
Number of times FAA oversight is a contributing factor in NTSB reports	6
Air carrier comments on oversight	4
Business Operations	
% budget spent on administrative expenses	3
Days of deliberation with labor	9
Overall funding levels	11
R&D budget	10
External Environmental Changes	
Growth of industry	11

In Table 5 we can see that in order to determine the effectiveness of the oversight system one of the items that we would want to monitor is the responsiveness of air carriers to corrective actions. An effective oversight system should be able to encourage completion of corrective actions in a timely manner. Subject Matter Experts have suggested that air carriers are more responsive to corrective actions when the safety oversight system presents information soundly, adequately, and precisely. Looking at Table 6, we can see that appropriate data inputs to this evaluation criterion would be the number of corrective actions, the time to complete corrective actions, and the number of overdue corrective actions. Appropriate metrics to evaluate this criterion would be the mean time to complete corrective actions and the percentage of overdue corrective actions. Acceptable ranges could be constructed for each of these metrics. Thus, by looking over time and regions, an analyst would be able to determine whether temporal/geographic values for these metrics were unusually high or low. Repeating this process for all 13 metrics would provide a sound basis for evaluating the health of the oversight system.

Not all criteria are as simple to calculate as the example given above, but all of them are based on inputs that are currently, or feasibly could be, collected by the FAA. Furthermore, the evaluation criteria are broad enough to be useful even as the oversight system continues to evolve. The next section presents a more complex analysis of safety outcomes as a measure of effectiveness.

5.5 Example Analyses

5.5.1 Introduction

This section provides some examples of the data analysis methodologies described in the previous sections. The examples are based on the FAA Accident/ Incident Data System (AIDS), a component of SPAS that records aviation accidents and incidents in both commercial and general aviation since 1978. We chose AIDS as the basis for our examples for two reasons. First, it is publicly available. While other SPAS components are considered highly sensitive and access is restricted to FAA facilities, we were able to download the AIDS data base from the World Wide Web. Second, the AIDS data base enables us to focus on outcomes, rather than outputs. While there is clearly a need to monitor oversight outputs, such as completed inspections, problems detected, and corrective actions performed, this cannot directly reveal the impact of the oversight system on aviation safety. By analyzing the AIDS data, we can investigate whether changes in the oversight system, such as the introduction of ATOS, have caused detectable changes in safety outcomes, in particular the occurrence of accidents and incidents.

5.5.2 AIDS Database

The AIDS database derives from FAA Order 8020.11, “Accident and Incident Notification, Investigation, and Reporting”, which requires the operator of any civil aircraft to report aircraft accidents and incidents. The regulation defines an aircraft accident as “an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage” and an incident as “an occurrence other than an accident, associated with the operation of an aircraft, which affects or could affect the safety of operations.” While the latter definition is less than exact, a review of the ten commercial air carrier incidents occurring between December 4-17, 2005 illustrates the variety of occurrences that are classified and reported as incidents. Events included:

- Mechanical issues (3 events—hydraulic system failure, inability to retract nose gear, and an aileron problem)
- Ground incursion with another vehicle(3 events—one with another plane, two with support vehicles)
- Smoke in the cockpit
- Turbulence-caused hard landing
- Door-latch and passenger door warning message
- Aircraft sliding off taxiway¹¹¹

Accident and incident reports made under FAA Order 8020.11 are collected in several data bases. The one used here is the AIDS data base maintained by the FAA AFS-620, which includes records for accidents and incidents for both commercial and general aviation since 1978. The incident data are collected under auspices of the FAA, while NTSB is the source of the accident data.

¹¹¹ Note: all incidents must be reported to FAA, but per 49 CFR Part 830, carriers only need to report the following seven types of non-accident events as incidents to the NTSB:

- Flight control system malfunction or failure;
- Inability of any required flight crewmember to perform normal flight duties as a result of injury or illness;
- Failure of structural components of a turbine engine excluding compressor and turbine blades and vanes;
- In-flight fire; or
- Aircraft collide in flight.
- Damage to property, other than the aircraft, estimated to exceed \$25,000 for repair (including materials and labor) or fair market value in the event of total loss, whichever is less.
- For large multiengine aircraft (more than 12,500 pounds maximum certificated takeoff weight):
 - In-flight failure of electrical systems which requires the sustained use of an emergency bus powered by a back-up source such as a battery, auxiliary power unit, or air-driven generator to retain flight control or essential instruments;
 - In-flight failure of hydraulic systems that results in sustained reliance on the sole remaining hydraulic or mechanical system for movement of flight control surfaces;
 - Sustained loss of the power or thrust produced by two or more engines; and
 - An evacuation of an aircraft in which an emergency egress system is utilized.

The AIDS dataset therefore contains a much wider array of incidents than the NTSB dataset.

5.5.3 Dataset

We compiled a dataset based on the AIDS incident and accident data. Our dataset contains counts of incidents and accidents for Major (over \$1 billion in operating revenues) and National (\$100 million to \$1 billion) passenger carriers over the period from January 1990 to December 2004.¹¹² Each record in our data set corresponds to an air carrier for one calendar month, and contains the following fields:

- Carrier, month, year.
- Number of reported accidents.
- Number of reported incidents.
- Number of flight departures performed.
- Quarterly financial statistics:¹¹³
 - Operating margin ($\frac{\text{operating revenue} - \text{operating expenses}}{\text{operating revenue}}$)
 - Current ratio $\frac{\text{current assets}}{\text{current liability}}$
 - Working capital $\frac{\text{current assets} - \text{current liability}}{\text{total assets}}$

Altogether, the data set contains records for 59 airlines. These include:

- Eight Major carriers that operated as Majors throughout the period of analysis.
- Three airlines that were Majors at the beginning of the period, but exited as a result of bankruptcy or merger between 1990 and 2004.
- Four carriers that began the period as Nationals but ended as Majors.
- 44 National carriers.

Although 59 carriers are represented in the data set, entry and exit was frequent. Thus, data for a specific month generally includes far fewer carriers. Figure 8 shows the number of carriers by month. The total grew from 23—11 Majors and 12 Nationals—in January 1990 to 38—13 Majors and 25 Nationals in 2000. Since then the number has declined—as response to

¹¹² We selected and classified carriers based on the information provided in the Aviation Support Table produced by the Bureau of Transportation Statistics entitled “Carrier Decode”.

¹¹³ We obtained financial data from Form 41—Airline Financial Statistics and Traffic Data.

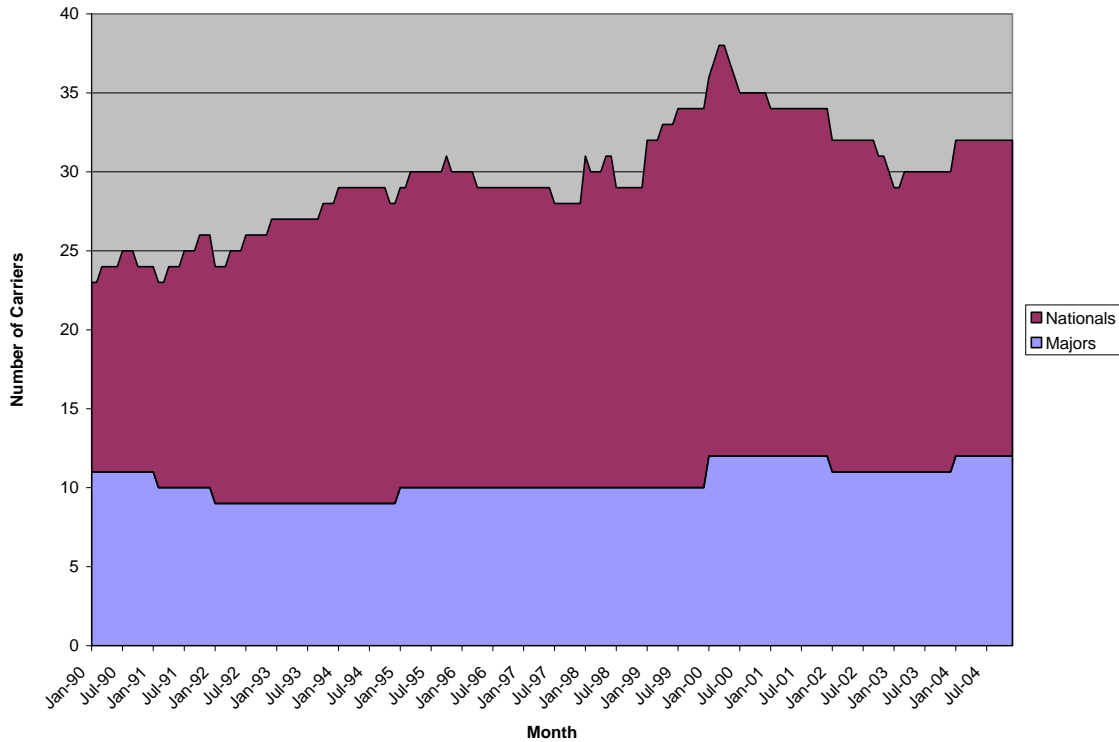
the industry downturn beginning in 2000 and accelerated by the 9/11 terrorist attacks. Major and National passenger carriers are a subset of all carriers that operate under 14 CFR Part 121. We have used Major carriers as a proxy for ATOS carriers. While this is not a perfect substitute, since the initial implementation of ATOS in October 1998, any carrier with “Major” status has been subject to ATOS. As ATOS has grown it has begun to include some National carriers, such as SkyWest Airlines. Unfortunately, since we did not have access to exact dates implementation of ATOS for each carrier, this approximation is the closest we could come to defining two groups of carriers as “ATOS” and “non-ATOS”.

5.5.4 Analysis

This data set permits the exploration of the relationship between safety outcomes and many factors, as well as the assessment of key trends related to safety outcomes. For present purposes, we confine our analysis to a few key issues. First, we will examine trends in incident and accident quantities and rates. Second, we will investigate how incident rate trends have been affected by the changes in the composition of flights between Major and National carriers. Third, we will extend the analysis to consider changes in the specific airlines operating over time. Last, we will perform an analysis designed to determine whether ATOS carriers experienced a change in accident and incident rates after ATOS was implemented.

Figures 9 and 10 show accident and incident trends, respectively, for the years from 1990 to 2004. The annual count is plotted using the scale on the left vertical axis, while rates—the ratio of the count to performed departures—is plotted using the right-hand scale. Figure 9 reveals that Major carrier accident annual counts vary between 4 and 22, while the range for rates is between 0.7 and 4.5 per million departures.

Figure 8: Number of Airlines in Data Set, by Category and Month



Given the inherent fluctuations in these data, trends are not easy to discern. However, Major accident counts and rates declined fairly steadily between 1996 and 2000 before declining even more dramatically in 2001. Results from the year 2003—the worst in the study period in terms of accidents—make it unclear whether the favorable trends in the later 1990s were real and, if so, whether the apparent safety gains have been maintained in the post-9/11 environment. The National carriers have lower counts and, for most years, lower rates than the Majors. There is little evidence of trend in the National data, which fluctuates strongly because of the small numbers of events involved.

Figure 9: Trends in Accidents and Accident Rates, Major and National Carriers, 1990-2004

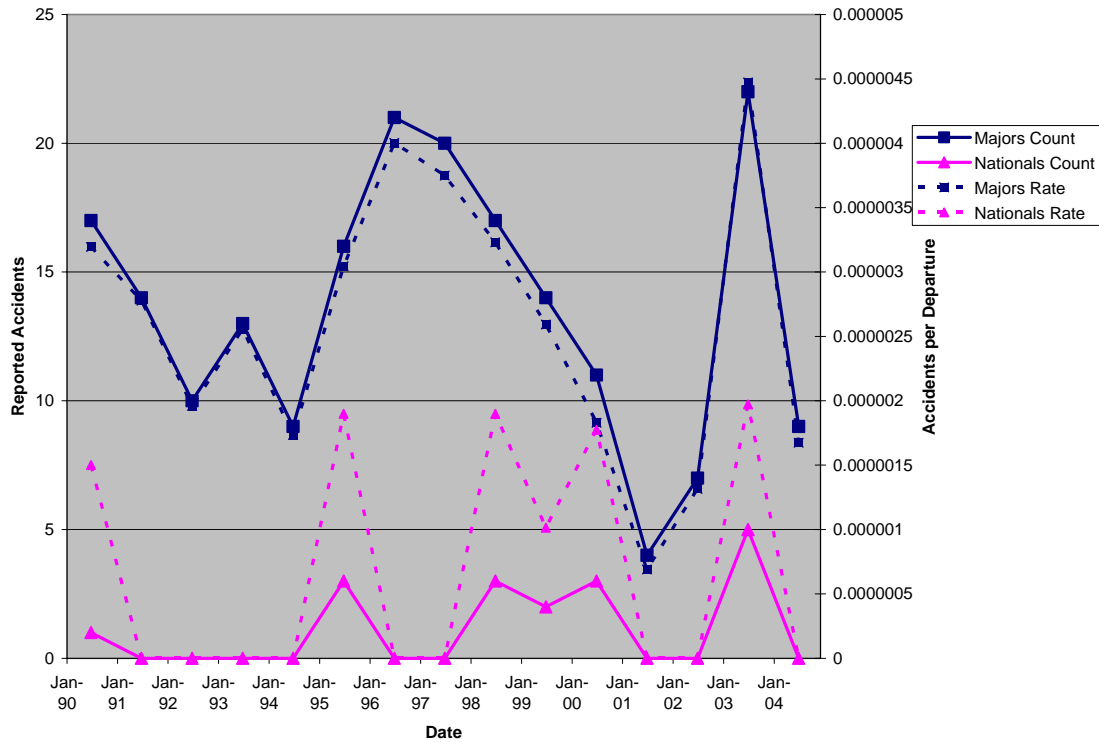
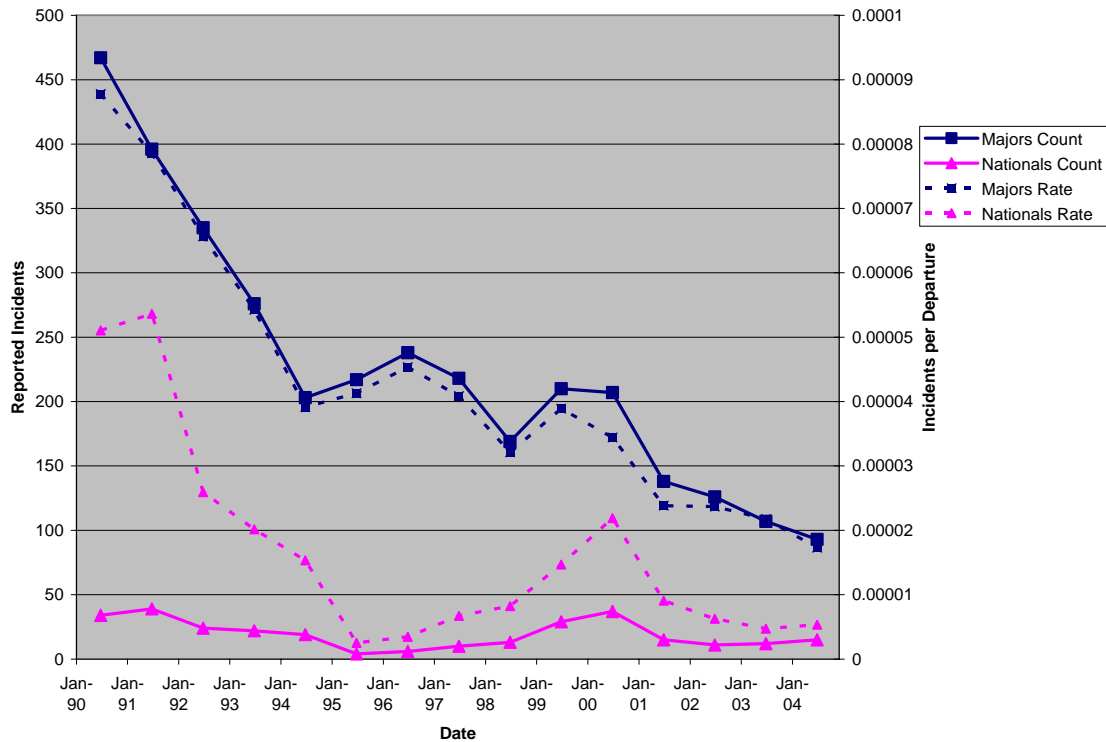


Figure 10 shows substantial declines in both the number of incidents and the incident rate for Major carriers. The count declined consistently during the early 1990s, falling from over 450 in 1990 to 200 in 1994. The decline has continued, albeit at a smaller rate and with greater fluctuation, since that time, going below 100 for the first time in 2004. National air carriers had a similar sharp drop in incidents in the early 1990s, but in this case the rate plummeted far more markedly than the count as a result of the growth in operations. Unlike the Majors, however, the downward trend did not continue after 1995. In fact, counts and rates for incidents both climbed beginning in the mid-1990s, peaking in 2000 before declining again in the early 1990s.

Figure 10: Trends in Incidents and Incident Rates, Major and National Carriers, 1990-2004



The trends documented in Figures 9 and 10 may be attributed to a wide range of factors, some of them related to the oversight system. For example, during the early 1990s, when incidents were decreasing, the FAA had recently implemented the NPG. (See chapter 2.) One area in which it was making substantial progress over this period was in completion rates for required inspections, which the GAO noted approvingly had reached nearly 100 per cent in 1996, as contrasted with the late 1980s when 70 per cent was closer to the norm.¹¹⁴ On the other hand, the early 1990s was also a period of rapid fleet modernization, resulting partly from the recognition of safety problems of aging aircraft, but also from the mandated retirement of Stage II aircraft under the Airport Noise and Capacity Act.¹¹⁵ The apparent safety progress in the early 1990s may be a result of the fleet changes, the oversight changes, or some combination of the two.

¹¹⁴ See Mark Hansen, Carolyn McAndrews, and Emily Berkeley, “History of the Aviation Safety Oversight System,” Nextor Research Report, Institute of Transportation Studies, University of California, Berkeley, 2005.

¹¹⁵ U.S. Federal Aviation Administration, “Historical Chronology, Civil Aviation and the Federal Governments, 1926-1976,” Sep. 19, 1991.

Starting in late 1998, ATOS became the primary oversight system for Major carriers, while Nationals remained under the NPG system. This creates something of a natural experiment, in which the Nationals can be used as a control group and the Majors as a treatment group. If, starting in 1999, we see changes in accident and incident rates for Majors that we do not see for Nationals, ATOS is a possible explanation. To investigate this, we calculated accident and incident rates for both carrier groups over the five years just prior to ATOS implementation (1994-1998) and the five years just after implementation (1999-2003). We also performed statistical tests to determine whether the rate differences observed may be the result of random chance. The results are summarized in Tables 7 and 8.

Table 6: Accidents, Flights, and Accident Rates, Major and National Airlines, by Period

Period	Majors			Nationals		
	Accidents	Flights	Rate	Accidents	Flights	Rate
1994-1998	83	2.63E+07	3.16E-06	6	7.62E+06	7.88E-07
1999-2003	58	2.74E+07	2.11E-06	8	8.65E+06	9.25E-07
1994-2003	141	5.37E+07	2.62E-06	14	1.63E+07	8.61E-07
Z-Score	-2.36			0.30		
Interpretation	Accident rate significantly lower in 1999-2003.			Accident rate not significantly different in 1999-2003.		

Table 7: Incidents, Flights, and Accident Rates, Major and National Airlines, by Period

Period	Majors			Nationals		
	Incidents	Flights	Rate	Incidents	Flights	Rate
1994-1998	1045	2.63E+07	3.98E-05	52	7.62E+06	6.83E-06
1999-2003	788	2.74E+07	2.87E-05	105	8.65E+06	1.21E-05
1994-2003	1833	5.37E+07	3.41E-05	157	1.63E+07	9.65E-06
Z-Score	-6.93			3.44		
Interpretation	Accident rate significantly lower in 1999-2003.			Accident rate significantly higher in 1999-2003.		

Table 7 shows that the accident rate for Majors declined in the post-implementation period, that the decline is statistically significant, and that the accident rate for Nationals did not change significantly between the two periods. Table 8 shows that the incident rate for Majors was also significantly lower in the post-implementation period, whereas the incident rate for Nationals was significantly higher. Both results support the idea that safety outcomes for Majors

improved significantly during the ATOS years and that these improvements were not realized for the National carriers who were not subject to ATOS.

While these results are encouraging, there are several caveats. First, as Figures 9 and 10 show, there have been large changes in accident and incident rates at other times, such as the early 1990s. These fluctuations urge caution in attributing the changes after ATOS implementation to any particular cause. The large incident rate increase for National carriers after ATOS is of particular concern in this regard. Second, incident and accident rates for Nationals are consistently below those for Majors. A satisfactory explanation of this difference is required before we can be confident in using Nationals as a control group. Finally, through discussions with subject matter experts with recent experience in the industry, we learned that the implementation of ATOS required majors to greatly increase expenditures on their safety programs. On the one hand, this provides a plausible mechanism by which ATOS could have substantially influenced safety outcomes. But on the other, it is not possible, on the basis of our analysis, to disentangle the effect of ATOS per se from that of growth in airline safety budgets.

The data base created for this project allows for the investigation of many other questions. Is there a correlation between the financial state of an airline and its accident and incident rates? How has safety performance been affected by the continuing turnover in the industry, particularly the smaller National carriers? And, when such data becomes available, is it possible to use inspection results to predict, in a statistical sense, the likelihood of accidents and incidents. The importance of these and other questions suggest that the use of statistical inference as a means of transforming data into useful information is clearly an important component of the OSEM.

6. Summary of Findings

In this report we have discussed the state of the current FAA safety oversight system, external reviews of the oversight system, and the challenges the FAA faces in carrying out its oversight duties. In addition, we have provided a framework with which the FAA can start to perform self-evaluation of its oversight system. In this report we focused primarily on the FAA's oversight of large, commercial air carriers. This is consistent with the FAA's goal to protect the American public since passengers on large commercial air carriers constitute the overwhelming majority of the flying population.

The current FAA safety oversight system is not one program but rather a network of programs, people, and activities that work together to ensure safety in the skies. Shaped by the FARs, these components come together to provide the certification, enforcement, and monitoring activities that are the basis of the safety oversight system. The FAA has two main oversight programs—ATOS and NPG. Originally implemented on ten of the largest passenger air carriers in October of 1998, ATOS has now grown to 13 passenger air carriers and 2 cargo carriers. Current FAA plans suggest that all carriers will eventually be transitioned into the ATOS program. The NPG program is used to plan oversight of the remaining 99 national air carriers. Thus, the OSEM must be flexible enough to provide meaningful information about this complicated system, even as it changes. The OSEM addresses this issue with its framework of evaluation questions that address the effectiveness, efficiency, and sustainability of the safety oversight system. These three attributes are relevant for the aviation safety oversight system.

While system safety has always been the foundation of ATOS, the FAA NPG program and certification processes have also started to include activities with the system safety philosophy. The OSEM uses criteria that are relevant to both the traditional approaches and system approaches to safety oversight. Because the criteria relate directly to fundamental questions about the effectiveness, efficiency, and sustainability of the safety oversight system, the list of criteria can be expanded to provide a more detailed characterization of these elements as the system safety concept is expanded.

Traditionally, ASIs and PIs working in field offices, or with CMTs, rely primarily on their experience in aviation and investigative skills to monitor air carriers. The FAA's emphasis on system safety now must combine the experience of its inspector workforce with the analytical capacity to monitor the safety and compliance of the entire aviation system. Data collection and

analysis have been issues for the safety oversight system for two decades, and the importance and complexity of these issues only increases. The data required for an OSEM relate to the fundamental activities of safety oversight – detecting and correcting noncompliance and safety deficiencies – and program elements that would be observable in nearly any safety oversight system, primarily monitoring activities, but also including enforcement activities and the implementation of corrective actions. Many of these data are collected by the safety oversight system for some time, and have even been used in evaluation reports by GAO.

Air carriers also participate in the safety oversight system by participating in voluntary disclosure, whistle-blower, self-evaluation, and other programs that help the aviation community develop a better understanding of the root causes of accidents and incidents, and enable the FAA to develop a safety oversight system that targets its resources to areas of relatively higher risk. This is why the OSEM also emphasizes the reporting and communication of evaluation findings, and a process that enables stakeholders such as air carriers to provide feedback to the FAA about the safety oversight system.

Because of the visibility of accidents, the FAA is subject to scrutiny and criticism from the flying public, Congress, and the GAO. The report provides a discussion of how the oversight system has been evaluated by external reviewers in the past. By understanding how outside examiners identify issues and potential weaknesses in the FAA oversight system, the FAA can better understand how to identify issues through internal evaluations. In this way, the FAA can both provide a safer oversight system and preempt critical outside evaluation. The report presents several criteria used by the GAO to evaluate the FAA safety oversight system, as well as methods that the GAO uses to develop its evaluations.

In order to evaluate the safety oversight system, the FAA needs to use a comprehensive approach. As a first step in this process, we have established a generic methodology with which the FAA can evaluate its safety oversight system. Establishing an evaluation methodology is a practical way for the FAA to evaluate how well the objectives and goals of its oversight system are being met. Starting with a framework—a set of questions that can judge whether or not the oversight system is achieving its goals and objective—we have then identified 13 example criteria that will answer these questions. Going further, we suggest 29 data inputs—data that the FAA currently collects, or potentially could collect, that could be used to establish the state of each criterion.

Additionally, ranges of acceptable values for each criterion could be used to create stoplight charts for each office, region, or CMO to identify local strengths and weaknesses in oversight activities. Similarly, using these criteria would be helpful for identifying trends over time. As an example, we provide an in-depth analysis of the relationship between incident rates and oversight activities over time. This analysis focused on the relationship between the FAA safety oversight system and accidents and incidents in the aviation system.

The FAA safety oversight system changes with the goal of constant improvement, and a list of its accomplishments would certainly be quite long. The safety record of the aviation industry is another such indicator. But the FAA has additional opportunities to tune its improvements with information from evaluations, and provide strong evidence of its contributions to the aviation system.

REFERENCES

1. Barnett, Arnold. (2004). "The Mother Metric? FAA's New Safety Index." Moving Metrics: A Performance-Oriented View of the Aviation Infrastructure. Asilomar Conference Center; Pacific Grove, California. January 27-30, 2004.
2. Barnett, Arnold and Alexander Wang. (1998). "Airline Safety: The Recent Record." NEXTOR Research Report RR-98-7.
3. Burkhardt, Robert. (1967). *The Federal Aviation Administration*. Frederick A. Praeger: New York.
4. Chérif, Taïb. (2005). "Working Together to Achieve Greater Safety." Luncheon Address, Federal Aviation Administration 2nd International Aviation Safety Forum. October 20, 2005.
5. Commons, Nicolas. (1978). *Bonfires to Beacons: Federal Civil Aviation Policy Under the Air Commerce Act, 1926-1938*. U.S. Department of Transportation, Federal Aviation Administration.
6. Fazen, Kathleen M. (2004). "Risk Management Decision Support – RPD 676 Commercial Aviation Requirements Program Review." Risk Management Decision Support Meeting, Atlantic City, New Jersey. July 13, 2004.
7. Kemp, Donald. (1968). "The Federal Aviation Administration's Air Safety Program." *Government-Industry System Safety Conference*. NASA Safety Office. National Aeronautics and Space Administration. May 1-3, 1968.
8. Kettl, Donald F. and John J. DiIulio, Jr. (1995). *Inside the Reinvention Machine: Appraising Governmental Reform*. The Brookings Institution: Washington, D.C.
9. Lederer, Jerome. (1968). "National Aeronautics and Space Administration." *Government-Industry System Safety Conference*. NASA Safety Office. National Aeronautics and Space Administration.
10. Molitoris, Jolene M. (1998). "Statement of Jolene M. Molitoris." *Testimony before the Senate Committee on Commerce, Science, and Transportation Subcommittee on Surface Transportation and Merchant Marine*. February 25, 1998.
11. Nader, Ralph and Wesley J. Smith. (1994). *Collision Course: The Truth About Aviation Safety*. Blue Ridge Summit: TAB Books.
12. National Aviation Safety Data Analysis Center. www.nasdac.faa.gov. Website visited December 19, 2005.
13. National Civil Aviation Review Commission. (2003). "Avoiding Aviation Gridlock and Reducing the Accident Rate: A Consensus for Change." <http://www.faa.gov/NCARC/reports/pepele.htm>. Website visited July 31, 2003.
14. National Transportation Safety Board. (2002). *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*. NTSB/AAR-02/01.
15. Niemeier, Dennis. (2004). "SASO Communication Messages." Risk Management Decision Support Meeting. Atlantic City, New Jersey. July 13, 2004.
16. Oster, Clinton V. Jr., John S. Strong, C. Kurt Zorn. (1992). *Why Airplanes Crash: Aviation Safety in a Changing World*. Oxford University Press: New York.
17. Patton, Michael Quinn. (1978). *Utilization-Focused Evaluation*. Beverly Hills: Sage Publications.
18. Schiavo, Mary. (1997). *Flying Blind, Flying Safe*. New York: Avon Books.
19. Strauss, Peter L. (1989). *An Introduction to Administrative Justice in the United States*. Durham: Carolina Academic Press.
20. U.S. Department of Transportation, Office of Inspector General. (2001). "Oversight of Aircraft Maintenance, Continuing Analysis and Surveillance Systems." AV-2002-066.

21. U.S. Federal Aviation Administration. (1990). *Inspector Handbook Volume 6: Chapter 1. Surveillance*. 8400.10 CHG 3.
22. U.S. Federal Aviation Administration, (1991). "Historical Chronology, Civil Aviation and the Federal Governments, 1926-1976," Sep. 19, 1991.
23. U.S. Federal Aviation Administration. (1993). *Inspector Handbook Volume 3: Aircraft and Equipment, Chapter 37. Monitor Continuing Analysis and Surveillance Program/Revision*. 8300.10 CHG 9.
24. U.S. Federal Aviation Administration. (1994). *Inspector Handbook Volume I: Chapter 2. The Federal Aviation Administration and Flight Standards: History and Organization, Section 1*. 8400.10 CHG 10.
25. U.S. Federal Aviation Administration. (1994). *Inspector Handbook Volume 7: Investigations, Chapter 1. Accident and Incident Investigation and Reporting*. 8400.10 CHG 10.
26. U.S. Federal Aviation Administration. (1994). *Inspector Handbook Volume 9: Technical Staff Administration and Responsibilities, Chapter 1. Inspector Responsibilities, Administration and Conduct, Section 1, Responsibilities of Aviation Safety Inspectors*. 8400.10 CHG 10.
27. U.S. Federal Aviation Administration. (1997). *FAA Order Number 1100.159*.
28. U.S. Federal Aviation Administration. (1999). *Inspector Handbook Volume II: Certification, Chapter 211. Conduct and Accident Investigation*. 8300.10 CHG 12.
29. U.S. Federal Aviation Administration. (2001). *Inspector Handbook Appendix 6*. 8400.10 CHG 13.
30. U.S. Federal Aviation Administration. (2001). "AFS-900, System Approach to Safety Oversight—Mission Need Statement."
31. U.S. Federal Aviation Administration. (2002). "Certification Process Study: An Evaluation of Selected Aircraft Certification, Operations, and Maintenance Processes."
32. U.S. Federal Aviation Administration. (2003). *Inspector Handbook Volume I: General Concepts, Direction, Guidance, and Definitions, Chapter 5. Voluntary Safety Programs, Section 1. Aviation Safety Action Program (ASAP)*. 8400.10 CHG 17.
33. U.S. Federal Aviation Administration. (2003). *Inspector Handbook Volume 1: General Concepts, Direction, Guidance, and Definitions, Chapter 6. Whistleblower Protection Program (Air Carrier)*. 8400.10 CHG 19.
34. U.S. Federal Aviation Administration. (2004). *Inspector Handbook Volume 1: General Concepts, Direction, guidance, and Definitions, Chapter 5. Voluntary Safety Programs, Section 2. Flight Operational Quality Assurance (FOQA)*. 8400.10 CHG 30.
35. U.S. Federal Aviation Administration. (2004). *Inspector Handbook Volume IX: Technical Staff Administration and Responsibilities, Chapter 1. Inspector Responsibilities, Administration, and Conduct*. 8400.10 CHG 10.
36. U.S. Federal Aviation Administration (2006), Aviation Safety Business Plan for FY 2006.
37. U.S. Federal Railroad Administration. (2002). "Five-Year Strategic Plan for Railroad Research, Development, and Demonstrations."
38. U.S. Federal Railroad Administration. "Accident/Incident Overview." <http://safetydata.fra.dot.gov/officeofsafety/>. Website visited July 30, 2004.
39. U.S. Federal Register. (2000). "Changes to the International Aviation Safety Assessment (IASA) Program." Volume 65. Number 102.
40. U.S. General Accounting Office. (1971). "Airport Safety Inspection Program Needed to Improve Flight Safety of Civil Aircraft." B-164497(1).
41. U.S. General Accounting Office. (1978). "Commercial Safety Regulations Are Avoided by Some Large Aircraft Operators." CED-79-10.

42. U.S. General Accounting Office. (1978). "Statement of Henry Eeschwege, Director Community and Economic Development Division Before the House Subcommittee on Government Activities and Transportation of the Committee on Government Operations on Uncertified Large Aircraft Operations."
43. U.S. General Accounting Office. (1985). "Compilation and Analysis of the Federal Aviation Administration's Inspection of a Sample of Commercial Air Carriers." GAO/RCED-85-157.
44. U.S. General Accounting Office. (1986). "FAA's Surveillance of Two Contract Military Carriers." GAO/RCED-86-128FS.
45. U.S. General Accounting Office. (1986). "FAA's Airline Inspection Program." Testimony by Herbert R. McLure. 129875.
46. U.S. General Accounting Office. (1986). "Comparison of Airlines with and Without Military Contracts." GAO/RCED-86-185BR.
47. U.S. General Accounting Office. (1987). "Aviation Safety." Testimony of Herbert R. McLure. GAO/T-RCED-87-1.
48. U.S. General Accounting Office. (1987). "Enhancing Policy and Program Effectiveness Through Improved Management." GAO/RCED-87-3.
49. U.S. General Accounting Office. (1987). "Procedures for Registering and Certifying Air Carriers." GAO/RCED-87-115FS.
50. U.S. General Accounting Office. (1987). "Needed Improvements in FAA's Airline Inspection Program Are Underway." GAO/RCED-87-62.
51. U.S. General Accounting Office. (1987). "Is Re-Regulation Needed to Improve Aviation Safety." Testimony by Herbert R. McLure. GAO/T-RCED-88-3.
52. U.S. General Accounting Office. (1988). "Measuring How Safely Individual Airlines Operate." GAO/RCED-88-61.
53. U.S. General Accounting Office. (1988). "Recruitment, Hiring, and initial Training of Safety-Related Personnel." GAO/RCED-88-189.
54. U.S. General Accounting Office. (1989). "FAA Aviation Safety Inspectors Are Not Receiving Needed Training." GAO/RCED-89-168.
55. U.S. General Accounting Office. (1989). "Meeting the Aging Aircraft Challenge: Status and Opportunities," Testimony by Kenneth M. Mead. GAO/T-RCED-90-2.
56. U.S. General Accounting Office. (1989). "FAA's Safety Inspection Management System Lacks Adequate Oversight." GAO/RCED-90-36.
57. U.S. General Accounting Office. (1989). "FAA Needs Comprehensive Plan to Coordinate Government and Industry Actions." GAO/RCED-90-75.
58. U.S. General Accounting Office. (1990). "Staffing, Training, and Funding Issues for FAA's Major Work Forces." Testimony by Kenneth M. Mead. GAO/T-RCED-90-42.
59. U.S. General Accounting Office. (1990). "Serious Shortcomings in FAA's Training Program Must Be Remedied." Testimony by Kenneth M. Mead. GAO/T-RCED-90-91.
60. U.S. General Accounting Office. (1991). "Limited Success Rebuilding Staff and Finalizing Aging Aircraft Plan." GAO/RCED-91-119.
61. U.S. General Accounting Office. (1991). "Additional FAA Oversight Needed of Aging Aircraft Repairs." GAO/RCED-91-91A.
62. U.S. General Accounting Office. (1991). "Emergency Revocation Orders of Air Carrier Certificates." GAO/RCED-92-10.

63. U.S. General Accounting Office. (1991). "Problems Persist in FAA's Inspection Program." GAO/RCED-92-14.
64. U.S. General Accounting Office. (1992). "Air Taxis -- The Most Accident-Prone Airlines -- Need Better Oversight." GAO/RCED-92-60.
65. U.S. General Accounting Office. (1992). "Better Oversight Would Reduce the Risk of Air Taxi Accidents," Testimony by John H. Anderson, Jr. GAO/T-RCED-92-27.
66. U.S. General Accounting Office. (1992). "Commuter Airline Safety Would Be Enhanced With Better FAA Oversight." Testimony by Kenneth M. Mead. GAO/T-RCED-92-40.
67. U.S. General Accounting Office. (1992). "Additional Actions Needed for Three Safety Programs," Testimony by Kenneth M. Mead. GAO/T-RCED-92-90.
68. U.S. General Accounting Office. (1992). "Progress on FAA Safety Indicators Program Slow and Challenges Remain." GAO/IMTEC-92-57.
69. U.S. General Accounting Office. (1992). "Increased Oversight of Foreign Carriers Needed." GAO/RCED-93-42.
70. U.S. General Accounting Office. (1993). "FAA Needs to Follow Through on Plans to Ensure the Safety of Aging Aircraft." GAO/RCED-93-91.
71. U.S. General Accounting Office. (1993). "Important Challenges Affecting Aviation Safety, Capacity, and Efficiency," Testimony by Kenneth M. Mead, GAO/T-RCED-93-33.
72. U.S. General Accounting Office. (1993). "Unresolved Issues Involving U.S.-Registered Aircraft," GAO/RCED-93-135.
73. U.S. General Accounting Office. (1993). "New FAA Approach Needed to Meet Challenges of Advanced Technology." GAO/RCED-93-155.
74. U.S. General Accounting Office. (1994). "FAA and the State Department Can Better Manage Foreign Enforcement Cases," GAO/RCED-94-87.
75. U.S. General Accounting Office. (1994). "FAA Technical Training," Testimony by Kenneth M. Mead, B-258242.
76. U.S. General Accounting Office. (1995). "Data Problems Threaten FAA Strides on Safety Analysis System." GAO/AIMD-95-27.
77. U.S. General Accounting Office. (1996). "Department of Transportation Has Taken Action to Improve Its Certification Process," GAO/RCED-96-8.
78. U.S. General Accounting Office. (1996). "Targeting and Training of FAA's Safety Inspector Workforce," Testimony by Gerald L. Dillingham, GAO/T-RCED-96-26.
79. U.S. General Accounting Office. (1996). "A Comprehensive Strategy Is Needed for Cultural Change at FAA," GAO/RCED-96-159.
80. U.S. General Accounting Office. (1996). "FAA Generally Agrees With but Is Slow in Implementing Safety Recommendations," GAO/RCED-96-193.
81. U.S. General Accounting Office. (1996). "New Airlines Illustrate Long-Standing Problems in FAA's Inspection Program," GAO/RCED-97-2.
82. U.S. General Accounting Office. (1997). "Addressing Management Issues at the Department of Transportation," GAO/T-RCED-AIMD-91-172.
83. U.S. General Accounting Office. (1997). "FAA Oversight of Repair Stations Needs Improvement," GAO/RCED-98-21.
84. U.S. General Accounting Office. (1997). "FAA's Guidance and Oversight of Pilot Crew Resource Management Training Can Be Improved," GAO/RCED-98-7.

85. U.S. General Accounting Office. (1997). "Efforts to Implement FOQA," GAO/RCED-98-10.
86. U.S. General Accounting Office. (1998). "Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks." GAO/RCED-98-6.
87. U.S. General Accounting Office. (1999). "Major Management Challenges and Program Risks, Department of Transportation," GAO/OCG-99-13.
88. U.S. General Accounting Office. (1999). "FAA's New Inspection System Offers Promise, but Problems Need to Be Addressed." GAO/RCED-99-183.
89. U.S. General Accounting Office. (2003). "FAA Needs to Update Curriculum and Certification Requirements for Aviation Mechanics," GAO-03-317.
90. U.S. Government Accountability Office. (2004). "Better Management Controls are Needed to Improve FAA's Safety Enforcement and Compliance Efforts," GAO-04-646.
91. U.S. Government Accountability Office. (2004). "FAA Needs to Strengthen the Management of Its Designee programs," GAO-05-40.
92. U.S. Government Accountability Office. (2005). "System Safety Approach Needs Further Integration into FAA's Oversight of Airlines." GAO-05-726.
93. U.S. Government Accountability Office. (2005). "FAA's Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs' Performance." Statement by Gerald L. Dillingham. GAO-06-266T.
94. U.S. Nuclear Regulatory Commission. (2004). "Reactor Oversight Process Self-Assessment For Calendar Year 2003." SECY-04-0053.
95. U.S. Office of Management and Budget. (2004). "Assessing Program Performance for the FY 2005 Budget." <http://www.whitehouse.gov/omb/part/>. Website visited July 31, 2004.
96. Office of Management and Budget. (2004). "Budget of the United States Government, Fiscal Year 2005, Appendix, Page 767." <http://www.whitehouse.gov/omb/budget/fy2005/appendix.html>. Website visited July 31, 2004.
97. U.S. Office of Technology Assessment. (1988). "Safe Skies for Tomorrow: Aviation Safety in a Competitive Environment." OTA-SET-381.
98. Wilson, James Q. (1989). *Bureaucracy: What Government Agencies Do and Why They Do It*. Basic Books: New York.

Appendix

GAO Documents Reviewed in Chapter 3

1. U.S. General Accounting Office, "Airport Safety Inspection Program Needed to Improve Flight Safety of Civil Aircraft," B-164497(1), January 15, 1971.
2. U.S. General Accounting Office, "Commercial Safety Regulations Are Avoided by Some Large Aircraft Operators," CED-79-10, November 21, 1978.
3. U.S. General Accounting Office, "Statement of Henry Eeschwege, Director Community and Economic Development Division Before the House Subcommittee on Government Activities and Transportation of the Committee on Government Operations on Uncertified Large Aircraft Operations," September 28, 1978.
4. U.S. General Accounting Office, "Compilation and Analysis of the Federal Aviation Administration's Inspection of a Sample of Commercial Air Carriers," GAO/RCED-85-157, August 2, 1985.
5. U.S. General Accounting Office, "Statement of Herbert R. McLure, Associate Director Resources, Community, and Economic Division Before the Subcommittee on Aviation of the Senate Committee on Commerce, Science, and Transportation on Three Safety Issues Relating to Aviation," 128050, October 1, 1985.
6. U.S. General Accounting Office, "FAA's Surveillance of Two Contract Military Carriers," GAO/RCED-86-128FS, March 13, 1986.
7. U.S. General Accounting Office, "FAA's Airline Inspection Program," Testimony by Herbert R. McLure, 129875, May 14, 1986.
8. U.S. General Accounting Office, "Comparison of Airlines With and Without Military Contracts," GAO/RCED-86-185BR, June 6, 1986.
9. U.S. General Accounting Office, "Aviation Safety," Testimony of Herbert R. McLure, GAO/T-RCED-87-1, January 29, 1987.
10. U.S. General Accounting Office, "Enhancing Policy and Program Effectiveness Through Improved Management," GAO/RCED-87-3, April 13, 1987.
11. U.S. General Accounting Office, "Procedures for Registering and Certifying Air Carriers," GAO/RCED-87-115FS, May 5, 1987.
12. U.S. General Accounting Office, "Needed Improvements in FAA's Airline Inspection Program Are Underway," GAO/RCED-87-62, May 19, 1987.
13. U.S. General Accounting Office, "Is Re-Regulation Needed to Improve Aviation Safety," Testimony by Herbert R. McLure, GAO/T-RCED-88-3, October 15, 1987.

14. U.S. General Accounting Office, "Measuring How Safely Individual Airlines Operate," GAO/RCED-88-61, March 18, 1988.
15. U.S. General Accounting Office, "Recruitment, Hiring, and initial Training of Safety-Related Personnel," GAO/RCED-88-189, September 2, 1988.
16. U.S. General Accounting Office, "FAA Aviation Safety Inspectors Are Not Receiving Needed Training," GAO/RCED-89-168, September 14, 1989.
17. U.S. General Accounting Office, "Meeting the Aging Aircraft Challenge: Status and Opportunities," Testimony by Kenneth M. Mead, GAO/T-RCED-90-2, October 10, 1989.
18. U.S. General Accounting Office, "FAA's Safety Inspection Management System Lacks Adequate Oversight," GAO/RCED-90-36, November 13, 1989.
19. U.S. General Accounting Office, "FAA Needs Comprehensive Plan to Coordinate Government and Industry Actions," GAO/RCED-90-75, December 22, 1989.
20. U.S. General Accounting Office, "Staffing, Training, and Funding Issues for FAA's Major Work Forces," Testimony by Kenneth M. Mead, GAO/T-RCED-90-42, March 14, 1990.
21. U.S. General Accounting Office, "Serious Shortcomings in FAA's Training Program Must Be Remedied," Testimony by Kenneth M. Mead, GAO/T-RCED-90-91, June 21, 1990.
22. U.S. General Accounting Office, "Limited Success Rebuilding Staff and Finalizing Aging Aircraft Plan," GAO/RCED-91-119, April 15, 1991.
23. U.S. General Accounting Office, "Additional FAA Oversight Needed of Aging Aircraft Repairs," GAO/RCED-91-91A, May 24, 1991.
24. U.S. General Accounting Office, "Emergency Revocation Orders of Air Carrier Certificates," GAO/RCED-92-10, October 17, 1991.
25. U.S. General Accounting Office, "Problems Persist in FAA's Inspection Program," GAO/RCED-92-14, November 20, 1991.
26. U.S. General Accounting Office, "Air Taxis -- The Most Accident-Prone Airlines -- Need Better Oversight," GAO/RCED-92-60, January 21, 1992.
27. U.S. General Accounting Office, "Better Oversight Would Reduce the Risk of Air Taxi Accidents," Testimony by John H. Anderson, Jr., GAO/T-RCED-92-27, February 25, 1992.
28. U.S. General Accounting Office, "Commuter Airline Safety Would Be Enhanced With Better FAA Oversight," Testimony by Kenneth M. Mead, GAO/T-RCED-92-

40, March 17, 1992.

29. U.S. General Accounting Office, "Additional Actions Needed for Three Safety Programs," Testimony by Kenneth M. Mead, GAO/T-RCED-92-90, August 4, 1992.
30. U.S. General Accounting Office, "Progress on FAA Safety Indicators Program Slow and Challenges Remain," GAO/IMTEC-92-57, August 31, 1992.
31. U.S. General Accounting Office, "Increased Oversight of Foreign Carriers Needed," GAO/RCED-93-42, November 20, 1992.
32. U.S. General Accounting Office, "FAA Needs to Follow Through on Plans to Ensure the Safety of Aging Aircraft," GAO/RCED-93-91, February 26, 1993.
33. U.S. General Accounting Office, "Important Challenges Affecting Aviation Safety, Capacity, and Efficiency," Testimony by Kenneth M. Mead, GAO/T-RCED-93-33, April 26, 1993.
34. U.S. General Accounting Office, "Unresolved Issues Involving U.S.-Registered Aircraft," GAO/RCED-93-135, June 18, 1993.
35. U.S. General Accounting Office, "New FAA Approach Needed to Meet Challenges of Advanced Technology," GAO/RCED-93-155, September 16, 1993.
36. U.S. General Accounting Office, "FAA and the State Department Can Better Manage Foreign Enforcement Cases," GAO/RCED-94-87, March 17, 1994.
37. U.S. General Accounting Office, "FAA Technical Training," Testimony by Kenneth M. Mead, B-258242, September 26, 1994.
38. U.S. General Accounting Office, "Data Problems Threaten FAA Strides on Safety Analysis System," GAO/AIMD-95-27, February 8, 1995.
39. U.S. General Accounting Office, "Department of Transportation Has Taken Action to Improve Its Certification Process," GAO/RCED-96-8, January 11, 1996.
40. U.S. General Accounting Office, "Targeting and Training of FAA's Safety Inspector Workforce," Testimony by Gerald L. Dillingham, GAO/T-RCED-96-26, April 30, 1996.
41. U.S. General Accounting Office, "A Comprehensive Strategy Is Needed for Cultural Change at FAA," GAO/RCED-96-159, August 22, 1996.
42. U.S. General Accounting Office, "FAA Generally Agrees With but Is Slow in Implementing Safety Recommendations," GAO/RCED-96-193, September 23, 1996.

43. U.S. General Accounting Office, "New Airlines Illustrate Long-Standing Problems in FAA's Inspection Program," GAO/RCED-97-2, October 17, 1996.
44. U.S. General Accounting Office, "Addressing Management Issues at the Department of Transportation," GAO/T-RCED-AIMD-91-172, May 21, 1997.
45. U.S. General Accounting Office, "FAA Oversight of Repair Stations Needs Improvement," GAO/RCED-98-21, October 27, 1997.
46. U.S. General Accounting Office, "FAA's Guidance and Oversight of Pilot Crew Resource Management Training Can Be Improved," GAO/RCED-98-7, November 24, 1997.
47. U.S. General Accounting Office, "Efforts to Implement FOQA," GAO/RCED-98-10, December 2, 1997.
48. U.S. General Accounting Office, "Weaknesses in Inspection and Enforcement Limit FAA in Identifying and Responding to Risks," GAO/RCED-98-6, February 27, 1998.
49. U.S. General Accounting Office, "Major Management Challenges and Program Risks, Department of Transportation," GAO/OCG-99-13, January 1999.
50. U.S. General Accounting Office, "FAA's New Inspection System Offers Promise, but Problems Need to Be Addressed," GAO/RCED-99-183, June 28, 1999.
51. U.S. General Accounting Office, "FAA Needs to Update Curriculum and Certification Requirements for Aviation Mechanics," GAO-03-317, March 6, 2003.
52. U.S. Government Accountability Office, "Better Management Controls are Needed to Improve FAA's Safety Enforcement and Compliance Efforts," GAO-04-646, July 6, 2004.
53. U.S. Government Accountability Office, "Better Management Controls are Needed to Improve FAA's Safety Enforcement and Compliance Efforts," GAO-04-646, July 6, 2004.
54. U.S. Government Accountability Office, "FAA Needs to Strengthen the Management of Its Designee programs," GAO-05-40, October 8, 2004.
55. U.S. Government Accountability Office, "System Safety Approach Needs Further Integration into FAA's Oversight of Airlines," GAO-05-726, September 2005.
56. U.S. Government Accountability Office, "FAA's Safety Oversight System is Effective but Could Benefit from Better Evaluation of Its Programs' Performance," Testimony by Gerald Dillingham, GAO-06-266T, November 17, 2005.