NASA/TM–2019-220417

Human Performance Contributions to Safety in Commercial Aviation

Jon B. Holbrook
Langley Research Center, Hampton, Virginia

Michael J. Stewart
San Jose State University, San Jose, California

Brian E. Smith
Ames Research Center, Moffett Field, California

Lawrence J. Prinzel
Langley Research Center, Hampton, Virginia

Bryan L. Matthews and Ilya Avrekh
Stinger Ghaffarian Technologies, Inc., Greenbelt, Maryland

Colleen T. Cardoza and Oliver C. Ammann
San Jose State University, San Jose, California

Viraj Adduru
Universities Space Research Association, Columbia, Maryland

Cynthia H. Null/NESC
Langley Research Center, Hampton, Virginia

November 2019

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at *http://www.sti.nasa.gov*

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM–2019-220417

Human Performance Contributions to Safety in Commercial Aviation

Jon B. Holbrook
Langley Research Center, Hampton, Virginia

Michael J. Stewart
San Jose State University, San Jose, California

Brian E. Smith
Ames Research Center, Moffett Field, California

Lawrence J. Prinzel
Langley Research Center, Hampton, Virginia

Bryan L. Matthews and Ilya Avrekh
Stinger Ghaffarian Technologies, Inc., Greenbelt, Maryland

Colleen T. Cardoza and Oliver C. Ammann
San Jose State University, San Jose, California

Viraj Adduru
Universities Space Research Association, Columbia, Maryland

Cynthia H. Null/NESC
Langley Research Center, Hampton, Virginia

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

November 2019

Acknowledgments

The assessment team would like to thank the NASA Engineering and Safety Center for funding this assessment; Mr. Gary Lohr for assistance in data collection and providing air traffic domain expertise that contributed to development and execution of the assessment; Mr. John Koelling for providing resources that enabled the collection and analysis of data from air traffic controllers; and Ms. Rebekah Hendricks, Ms. Terri Derby, Ms. Linda Burgess, and Ms. Jenny DeVasher for providing invaluable support in planning, coordinating, and editing this report. In addition, the assessment team would like to thank all the peer reviewers for their valuable comments on the report.

Available from:

NASA Center for AeroSpace Information
7115 Standard Drive
Hanover, MD 21076-1320
443-757-5802

Table of Contents

List of Figures

List of Tables

Nomenclature

ARMD	Aeronautics Research Mission Directorate
ASAP	Aviation Safety Action Program
ASDE-X	Airport Surface Detection Equipment Model-X
ASIAS	Aviation Safety Information Analysis and Sharing
ASPM	Aviation System Performance Metrics
ASRS	Aviation Safety Reporting System
ATC	Air Traffic Control
BTS	Bureau of Transportation Statistics
CAST	Commercial Aviation Safety Team
CCFP	Collaborative Convective Forecast Product
CIFP	Coded Flight Instrument Procedures
CIWS	Corridor Integrated Weather System
CWAM	Convective Weather Avoidance Model
D2D	Day-to-Day Safety Survey
DT-MIL	Deep Temporal Multiple Instance Learning
FAA	Federal Aviation Administration
FAR	Federal Aviation Regulations
FOQA	Flight Operational Quality Assurance
ICAO	International Civil Aviation Organization
IRB	Institutional Review Board
KPI	key performance indicators
LOFT	Line Oriented Flight Training
LOSA	Line Operational Safety Audits
NAS	National Airspace System
NESC	NASA Engineering and Safety Center
NM	nautical miles
NOAA	National Oceanic and Atmospheric Administration
NOSS	Normal Operations Safety Survey
NOTAM	Notice to Airmen
NPR	NASA Procedural Requirement
OEM	original equipment manufacturer
PDARS	Performance Data Analysis and Reporting System
PFD	Primary Flight Display
RAG	Resilience Analysis Grid
RNAV STAR	area navigation standard terminal arrival route

SFDPS	SWIM Flight Data Publication Service
STDDS	SWIM Terminal Data Distribution System
SWIM	System Wide Information Management
TAF	Terminal aerodrome forecast
TFMS	Traffic Flow Management
TRACON	Terminal Radar Approach Control
UAM	Urban Air Mobility

Abstract

In the commercial aviation domain, large volumes of data are collected and analyzed on the failures and errors that result in infrequent incidents and accidents, but in the absence of data on behaviors that contribute to routine successful outcomes, safety management and system design decisions are based on a small sample of non-representative safety data. Analysis of aviation accident data suggests that human error is implicated in up to 80% of accidents, which has been used to justify future visions for aviation in which the roles of human operators are greatly diminished or eliminated in the interest of creating a safer aviation system. However, failure to fully consider the human contributions to successful system performance in civil aviation represents a significant and largely unrecognized risk when making policy decisions about human roles and responsibilities. Opportunities exist to leverage the vast amount of data that has already been collected, or could be easily obtained, to increase our understanding of human contributions to "things going right" in commercial aviation. The principal focus of this assessment was to identify current gaps and explore methods for identifying human "success" data generated by the aviation system, from personnel and within the supporting infrastructure.

1.0 Executive Summary

Every day in aviation, pilots, air traffic controllers, and other front-line personnel perform countless correct judgments and actions in a variety of operational environments. These judgments and actions are often the difference between an accident and a non-event. Ironically, data on these behaviors are rarely collected or analyzed.

Data-driven decisions about safety management and design of safety-critical systems are limited by the available data, which influence how decision makers characterize problems and identify solutions. Large volumes of data are collected on the failures and errors that result in infrequent incidents and accidents, but in the absence of data on behaviors that result in routine successful outcomes, safety management and system design decisions are based on a small sample of non-representative safety data. This assessment aimed to find and document "safety successes" made possible by human operators.

With many Aeronautics Research Mission Directorate (ARMD) Programs and Projects focusing on increased automation and autonomy and decreased human involvement, failure to fully consider the human contributions to successful system performance in civil aviation represents a significant risk—a risk that has not been recognized to date. Without understanding how humans contribute to safety, any estimate of predicted safety of autonomous capabilities is incomplete and inherently suspect. Furthermore, understanding the ways in which humans contribute to safety can promote strategic interactions among safety technologies, functions, procedures and the people using them. Without this understanding, the full benefits of an integrated, optimized human/technology or autonomous system will not be realized.

Historically, safety has been consistently defined in terms of the occurrence of accidents or recognized risks (i.e., in terms of things that go wrong). These adverse outcomes are explained by identifying their causes, and safety is restored by eliminating or mitigating these causes. An alternative to this approach is to focus on what goes right and identify how to replicate that process. Focusing on the rare cases of failures attributed to "human error" provides little information about why human performance routinely prevents adverse events. Hollnagel [Ref. 1] has proposed that things go right because people continuously adjust their work to match their operating conditions. These adjustments become increasingly important as systems continue to grow in complexity. Thus, the definition of safety should reflect not only "avoiding things that go wrong" but "ensuring that things go right." The basis for safety management requires developing an understanding of everyday activities. However, few mechanisms to monitor everyday work exist in the aviation domain, which limits opportunities to learn how designs function in reality.

This concept of safety thinking and safety management is reflected in the emerging field of resilience engineering. According to Hollnagel [Ref. 2], a system is *resilient* if it can sustain required operations under expected and unexpected conditions by adjusting its functioning prior to, during, or following changes, disturbances, and opportunities. To explore "positive" behaviors that contribute to resilient performance in commercial aviation, the assessment team examined a range of existing sources of data about pilot and air traffic control (ATC) tower controller performance, including subjective interviews with domain experts and objective aircraft flight data records. These data were used to identify strategies that support resilient performance, methods for exploring and refining those strategies in existing data, and proposed methods for capturing and analyzing new data.

The findings and observations presented in Section 4.0 can be summarized as:

- NASA and industry planning and system design in aviation are based on principles and methods focused on predicting and preventing errors.
- Current safety reporting processes are designed to focus on and capture events that degrade safety, but not positive events that bolster safety.
    - Operators identified cultural barriers to reporting positive behaviors, because adapting to routine disturbances was seen as expected job performance.
    - Existing observer-, operator-, and system-based approaches to data collection and analysis do not systematically include operators' resilient behaviors.
    - Many of the identified behaviors associated with operators' ability to anticipate, monitor, and respond require leveraging experience-based information. No methods currently exist, however, to systematically report or capture this information.
    - Existing operator behavior taxonomies conflate "positive" operator behaviors with "positive" operational outcomes and "negative" operator behaviors with "negative" operational outcomes.

- Defining safety in terms of "things that go right" enabled new methods for exploring existing data.
    - Operators were able to reflect on and provide specific examples of resilient behaviors.
    - Evidence of the use of operator strategies that promote resilient performance were identified in objective data.
- Subjective and objective data sources contributed different information toward building an understanding of operators' resilient performance.
    - Current approaches for safety data collection and analysis are not designed to integrate data from disparate sources.
    - Subjective data sources are necessary to understand the rationale for actions identified in objective data.

These findings resulted in eight NESC recommendations for NASA ARMD's Transformative Aeronautics Concepts Program and Airspace Operations and Safety Program:

- Define safety in terms of the presence of desired behaviors and the absence of undesired behaviors.
- Leverage existing data to identify strategies and behaviors that build resource margins and prevent them from degrading.
- Develop organization-level strategies that promote recognition and reporting of behaviors that support resilient performance.
- Develop expert-observer-based data collection tools to capture strategies and behaviors that support resilient performance.
- Develop methods to collect and analyze operator-reported strategies and behaviors that support resilient performance.
- Develop approaches to expand collection and facilitate analysis of resilient behaviors in adverse event reports.
- Refine data analytics approaches for exploiting Flight Operational Quality Assurance (FOQA) data based on identified resilience strategies.
- Develop a system-level framework for collecting and analyzing resilient performance data that is explicitly designed to integrate information from observer-, operator-, and system-based data sources.

To improve safety, system designers should understand what humans do well and create systems with this understanding in mind. System designers and safety managers should look at what goes right as well as what goes wrong, and learn from what succeeds as well as from what fails. Things do not go well because people follow rules and procedures. Rather, things go well because people exhibit performance variability and make sensible adjustments and adaptations in response to interpretation of what is happening and the demands of the situation. Through understanding "how" and "why" people perform work, in addition to understanding "what," "where," and "when," systems can be designed to ensure the ultra-safe airspace system is not unintentionally made less safe due to loss of resilient properties that are provided by human operators and are not well-understood.

Although the civil aviation system includes humans in many roles, the assessment team focused on commercial air transport operations, specifically on the roles of pilots and ATC tower controllers. The particular focus was on behavior of individual operators rather than that of teams or organizations. Furthermore, analysis was limited to datasets that were rapidly obtainable and already familiar to the team, given the time constraints of the assessment.

2.0 Problem Description, Background, and Scope

2.1 Problem Description

In World War II, Allied aircraft were key to the war effort, yet they were constantly at risk of being shot down over enemy territory. The planes needed more armor, but due to weight restrictions, armor plating could be applied only where necessary. The team tasked with identifying the critical locations for armor plating plotted the pattern of bullet holes on returning aircraft. The initial assumption was that the locations of these clusters were the spots that needed more armor.

However, Abraham Wald, who was brought in to oversee the operation, reasoned that the military did not need to reinforce the spots that had bullet holes. Rather, they needed to reinforce the spots that did not have bullet holes. Planes that had been shot in the bullet-free zones never made it home to be accounted for in the diagrams. For Wald, the revelation was that the bullet-hole data from returning aircraft were not random, but an indication that those bullet-free zones were the most vital [Ref. 3].

Wald’s insight has important implications for current-day civil aviation safety: System designers and safety managers should identify those technical, functional, and procedural domains that do not have bullet holes — success areas that have not been hit and for which no failure maps (i.e., bullet-hole patterns) exist. When segments of a target population — in this case, successful, safe flights — are not systematically analyzed, they are under-represented in the conclusions. In this situation, critical factors causing commercial aviation to display the characteristics of an ultra-safe, socio-technical system are not identified, measured, or analyzed. This report offers methods for discovering and analyzing the missing “success data” in civil aviation.

Successful decisions during routine operations require dealing with phenomena like weather, system problems, gate/runway unavailability, unanticipated autopilot behavior, and many other dynamic, unplanned, or unanticipated events. According to the report of the Performance-based operations Aviation Rulemaking Committee/Commercial Aviation Safety Team (PARC/CAST) Flight Deck Automation Working Group [Ref. 4], 20% of flights experience aircraft malfunctions requiring flight crew intervention. Recovery from these malfunctions occurs seamlessly in most cases. In addition, favorable management and organizational environments (e.g., training) support the actions of operational personnel in dealing with these in-flight malfunctions. These recovery actions by human operators are evidence of “graceful degradation” from nominal technology and procedural functions and organizational processes.

NASA ARMD’s Strategic Implementation Plan proposes that machines take on tasks and responsibilities currently performed by humans [Ref. 5]. One of the future aviation concepts in NASA ARMD’s research portfolio is Urban Air Mobility (UAM). The UAM concept describes a safe and efficient system for vehicles, piloted or not, to move passengers and cargo within a city environment. While private companies are starting to develop the infrastructure to make UAM a reality, NASA's role in UAM is to establish feasibility and help set the requirements to enable the UAM vision [Ref. 6]. NASA has signed a Space Act Agreement with Uber Technologies, Inc. (Uber), to explore UAM concepts and technologies as well as identify and address the challenges facing the UAM market [Ref. 7].

As a part of its vision for the future of on-demand urban air transportation, Uber identified that safety is paramount to establishing market acceptance [Ref. 8]. Uber asserted that safety can be measured on a number of negative dimensions, including injuries, accidents, and fatalities, and that the path to improving safety requires understanding the root causes of historical crashes. Uber further highlighted that pilot error represents a leading cause of fatalities, and concluded that "to fast-forward to the safest possible operational state for vertical takeoff and landing vehicles, network operators will be interested in the path that realizes full autonomy as quickly as possible." This assertion that full autonomy represents the safest possible operation state presupposes that human operators make operations less safe. However, neither Uber's vision for UAM nor NASA's planned activities to enable this vision consider the everyday operational practices of human operators that contribute to system safety. Without understanding how humans contribute to safety, any estimate of predicted safety of autonomous capabilities is incomplete and inherently suspect. Furthermore, this understanding provides the necessary basis for designing machine systems that could perform safety-producing behaviors.

ARMD's Strategic Implementation Plan proposes development of in-time safety monitoring, prediction, and mitigation technologies. Ironically, a critical barrier to measuring safety threats and the impact of mitigation strategies in ultra-safe systems like commercial aviation is the lack of opportunities for measurement. Although it is common practice to relate safety to how many accidents or fatalities occur for a given volume of traffic, very safe systems have very few accidents. Therefore, accident data cannot be readily used to validate safety improvements for at least two reasons. First, the time necessary to observe the effect of a given safety intervention in accident statistics becomes excessively long, with estimates up to 6 years for a system with a fatal accident rate per operation of 10^{-7} [Ref. 9]. Second, attributing improvement to a specific intervention becomes intractable due to the many thousands of changes that a complex sociotechnical system would experience over that time period [Ref. 10]. While ARMD's strategic roadmap for Real-Time System-Wide Safety indicates a need to develop models and metrics to characterize safe operations, the research roadmap and technical challenge descriptions that define the research plans for ARMD's System-Wide Safety Project consistently characterize safety with regard to identifying and avoiding risks. Defining safety in terms of eliminating or minimizing factors that create risk leads at best to an incomplete picture of safety. First, this approach leverages "known" risks identified through analysis of accident or incident data, and therefore does not encompass all risk. Second, this approach does not consider factors that create safety successes. This is particularly relevant for very safe systems, in which safety successes dramatically outnumber safety failures.

Decisions about safety management and design of safety-critical systems should be, and very frequently are, based on data. However, data-driven decisions are limited by what data have been collected and analyzed. In current-day civil aviation, large volumes of data are collected on the failures and errors that result in infrequent incidents and accidents, but data on behaviors that result in routine successful outcomes are rarely collected or analyzed. Thus, data-driven safety management and system design decisions are based on a small sample of non-representative safety data. The data that are available bias how decision makers characterize problems and identify solutions.

The goal of this NESC assessment was to develop methodologies for identifying the beneficial roles humans and organizations play in the aviation system as a complement to human error detection and mitigation. This assessment aimed to identify tools, methods, and data collection mechanisms for finding "success" data generated by the aviation system, from personnel and within the supporting infrastructure. The assessment addressed the following questions:

a) What can be learned about human contributions to safety success from existing data sources?

b) Where are the gaps in data collection and analysis?

c) What are the opportunities for collecting success data in the future?

The answers to these questions could give ARMD and the aviation safety community powerful insights into what is working and why, and provide guidance on where to target safety resources.

This assessment builds upon a growing resilience engineering literature and new approaches to safety. However, the question remains largely open of how to translate the principles, values and concepts described in the literature into specific instruments and tools for specific organizations or work domains [Ref. 11].

2.2 Background

The NAS is a highly complex sociotechnical system that continues to grow and change at an increasing pace. The ever-increasing demand for flights has led to increasingly crowded airspace, an accompanying need for more personnel, and more extensive use of automation. Despite the many changes and advances in commercial aviation, safety thinking, practices, and models have not kept pace. Thus, many of the same simple cause-and-effect relations and linear models used to explain accidents and incidents since at least the 1930s remain in widespread use [Ref. 12].

Historically, safety has been consistently defined in terms of the occurrence of accidents or recognized risks (i.e., safety is typically defined in terms of things that go wrong). These adverse outcomes are, in turn, explained by identifying their causes, and safety is restored by eliminating or mitigating these causes. As new accidents occur, they are explained by new causes, typically relating to technology, human factors, or organizational factors.

The ubiquity of this approach is apparent in regulations requiring detailed reporting of accidents and incidents; organizations and groups dedicated to analyzing these events; databases of incidents and accidents; and numerous models and taxonomies of things that go wrong and their causes. However, for ultra-safe systems like commercial air transport, this effort focuses on a very small proportion of events (i.e., the probability of being in a fatal accident on a commercial jet passenger flight is 2.0×10^{-7} [Ref. 13]). Thus, for every well-scrutinized accident, millions of flights in which things go right receive very little attention.

Hollnagel refers to this common understanding of safety, which focuses on reducing adverse outcomes, as Safety I [Ref. 1]. Safety I describes a bimodal view of system performance, in which things go right because the system is functioning as it should, and things go wrong when something in the system has malfunctioned or failed. The goal of safety management is to ensure that the system stays in the first mode and avoids the second. Thus, safety is defined in terms of what happens when it is absent rather than when it is present. A consequence of this definition is that perfect safety is defined by having no adverse outcomes, and therefore nothing to measure. Thus, for a time period in which no adverse outcomes occur, it becomes impossible to demonstrate the efficacy of efforts to improve safety. Furthermore, predictions of the future safety state of a system are dependent on known risks and are insensitive to uncertain, unanticipated, and unpredicted risks.

Safety I thinking is predicated on the belief that adverse outcomes occur because something has gone wrong, therefore adverse outcomes have causes that can be identified and treated. Because a system can be constructed, it is assumed that this process can be reversed, and systems can be decomposed into meaningful constituents. Furthermore, it is assumed that if this can be accomplished for technological systems, it can be done for tasks and events.

While this approach may be successful for comprehensively defined, highly constrained systems, these assumptions are inappropriate for complex sociotechnical systems (e.g., today's commercial air transport system). With the continuous introduction of increasingly complex and capable technologies comes a commensurate demand for increased operational capacity, intensity, and tempo (i.e., the Law of Stretched Systems [Ref. 14]). This demand results in a system in which functions and services become more tightly coupled, making it increasingly harder to isolate and address individual constituents. Furthermore, increasing system complexity and rate of change means that systems are likely to evolve before they can be fully described. The resulting system has limited predictability and operates in ways that cannot be precisely prescribed in design or operation. Therefore, Safety I thinking is an insufficient basis for safety management.

A further consequence for complex technologies designed using Safety I thinking is that the operator must fill in the gaps for those aspects of performance that were not identified or considered during design. As systems become more complex, the number of unanticipated operating states increases. The resulting regular compensation by the operator for system design flaws comes at a cost of increased workload, performance pressures, and vulnerability to unpredicted outcomes [Ref. 15]. Furthermore, the Safety I focus on failure gives the impression that the human performance variability that results from inadequate design is a major hazard, puts the blame on the operator, and does little to uncover the details of successes and opportunities created by human adaptations [Ref. 16].

Safety I thinking asserts that nominal working conditions can be completely analyzed and prescribed (i.e., "work-as-imagined"). On the other hand, "work-as-done" describes how work actually unfolds over time [Ref. 17]. Thus, an assumption of Safety I thinking is that safety can be achieved by ensuring that work-as-done corresponds to work-as-imagined. In complex environments, work practices (i.e., work-as-done) can differ significantly from the procedures and policies that define work-as-imagined. This is often acknowledged in complex systems by giving operators the latitude to deviate from procedures when, in their judgment, it is in the safety interest of the operation. For example, Federal Aviation Regulations state: "In an in-flight emergency requiring immediate action, the pilot in command may deviate from any rule of this part to the extent required to meet that emergency" [Ref. 18]. However, this also leads to the unavoidable conclusion that safety models predicated on equating work-as-done with work-as-imagined are, at best, inadequate.

An alternative approach to Safety I is to focus on what goes right, and identify how to replicate this process. This approach is identified by Hollnagel [Ref. 1] as Safety II. Focusing on the rare cases of failures attributed to "human error" provides little information about why human performance almost always goes right. Similarly, focusing on the lack of safety provides little information about how to improve safety. Safety II thinking is predicated on the concept that things go right, in part, because people continuously adjust their work practices to match their operating conditions. These adjustments become increasingly important as systems continue to grow in complexity. Hollnagel proposed that the definition of safety should reflect not only "avoiding things that go wrong" but "ensuring that things go right." Thus, the basis for safety management requires developing an understanding of everyday activities.

It can be argued that if something already works, why spend more time on it? However, in complex systems, things often do not work in the ways they are intended or assumed to work. In addition, in dynamic, evolving environments, it cannot be assumed that routines that work today will continue to work in the future. The discrepancy that can exist between work-as-imagined and work-as-done was illustrated in a civil aviation context by Stewart, Matthews, Janakiraman, and Avrekh [Ref. 19]. Area navigation standard terminal arrival route (RNAV STAR) procedures used at major airports are intended to increase predictability and efficiency. These procedures provide vertical, lateral, and speed profiles for aircraft to follow as they descend toward an airport. Analyzing aircraft flight track data for more than 10 million flights into 32 domestic airports revealed that only 12.4% of flights fully complied with the vertical and lateral profiles in the RNAV STARs (i.e., vertical compliance within 300 feet above or below a waypoint altitude restriction, and lateral compliance within one mile left or right of the route centerline). While the Stewart, et al., study provides an example in which published procedures (i.e., work-as-imagined) were misaligned with normal operations (i.e., work-as-done), questions remain with regard to the reasons for the misalignment, and thus how to interpret this finding. For example, these procedural non-adherences may represent necessary adaptations to achieve operational goals in specific contexts (e.g., avoiding encounters with convective weather). One of the foundations of Safety II is that performance is always variable, and the ability to make performance adjustments is an essential human contribution to work. Monitoring everyday work can help to efficiently and proactively identify new strategies that work and conditions under which existing strategies break down. Few mechanisms to monitor everyday work exist in the aviation domain, which limits opportunities to learn how designs function in reality.

It should be noted that Safety I and Safety II represent two complementary views of safety rather than incompatible or conflicting approaches. The primary characteristics of each approach are summarized in Table 2.2-1. Many of the existing Safety I practices can provide valuable insights. When adverse outcomes in complex sociotechnical systems cannot be explained using the principles of decomposition and causality, however, this is an indication that those methods are inadequate. In these circumstances, Safety II approaches to understanding events may add value.

Table 2.2-1. Comparison of Safety I and Safety II.

Traditional Approach (Safety I)	Emerging Approach (Safety II)
Safety: When as few things as possible go wrong	Safety: When as many things as possible go right
Focus on predicting failure probabilities	Focus on preparing for the unpredicted
People are a liability—control, correct	People are an asset—learn, adapt
Variability is a threat—minimize it	Variability is normal—manage it
Focus on incident rates	Focus on learning
Focus on what is not wanted: incidents and accidents	Focus on what is wanted: how safety is created
Procedures are complete and correct	Procedures are always under-specified and must be interpreted and adapted
Systems are well designed, work as designed, and are well maintained	Systems are complex and will degrade; there will always be flaws and glitches

The assessment team used the Resilience Analysis Grid (RAG) framework to provide a structure for instantiating Safety II thinking [Ref. 20]. The RAG identifies four capabilities of resilient performance:

- ***Anticipate***: Knowing what to expect, or being able to anticipate developments further into the future (e.g., potential disruptions, novel demands or constraints, new opportunities, changing operating conditions.
- ***Monitor***: Knowing what to look for, or being able to monitor that which is or could positively or negatively affect the system's performance in the near term.
- ***Respond***: Knowing what to do, or being able to respond to regular and irregular changes, disturbances, and opportunities by activating prepared actions or by adjusting current mode of functioning
- ***Learn***: Knowing what has happened, or being able to learn from experience, in particular to learn the right lessons from the right experience.

Each of these four capabilities can be systematically probed with questions designed to gauge the potential for that ability to be present and functional. The assessment team adopted this framework to aid in identifying operator behaviors that might be examples of resilient performance, defined as the ability of a system to sustain required operations under both expected and unexpected conditions by adjusting its functioning prior to, during, or following changes, disturbances, and opportunities [Ref. 2].

2.3 State of Practice

Several challenges are associated with existing accident and incident reporting and lessons-learned approaches to understanding system safety:

- Accident- and incident-based mitigations are reactive rather than proactive.
- Accidents and incidents can be unique, with different patterns of contributing factors. Therefore, designed mitigation strategies may not generalize to other adverse events.

- In highly safe systems, there are few incidents or accidents, so these events cannot be relied upon for timely safety monitoring or improvement.

Alternatively, risk assessment-based approaches to safety management take a more proactive stance by looking at what abnormal events could happen. However, none of these approaches focus systematically on what does happen every day.

Observational safety surveys are being used in the ATC and flight deck domains as tools for helping organizations notice when performance is drifting toward a less safe state, or becoming more variable than desired. These surveys are direct observations conducted during a normal work period by trained observers who are also domain experts. The Normal Operations Safety Survey (NOSS) and the Day-to-Day Safety Survey (D2D) have been used in the ATC domain, and Line Operational Safety Audits (LOSA) have been used with pilots.

NOSS and LOSA are based on a "threat and error management" model [Ref. 21]. In these approaches, trained observers capture data on recovery from undesirable states that occur from threats and errors that may develop during daily operations. Data from these observations provide lessons learned with regards to threats to safety and unexpected situations. Through the coding of threats and errors, key problem areas (e.g., incomplete handover briefings, distracting non-operational conversation, using incorrect procedure, etc.) are identified. This supports the goal of identifying and setting clear targets for operational safety enhancements. However, many safety-producing actions are taken by operators to anticipate and monitor events before threats manifest.

The D2D survey, used by the National Air Traffic Services United Kingdom and the Irish Aviation Authority, includes direct observations of air traffic controllers and is focused on collecting data about positive and proactive behaviors [Ref. 22]. The intent of the D2D survey is to reveal trends to enhance understanding of what air traffic controllers do to maintain safe operations before a situation evolves into an undesirable state. The D2D survey is currently limited to five areas: visual scanning cycle, active listening, defensive controlling, write-as-you-speak-read-as-you-listen; and strip management. This is a recent approach that is limited in scope and not in widespread use, but over time may provide insights into the actions that operators take to promote safe operations.

Data from adverse event analyses have been used to create detailed taxonomies of human error, but rarely to capture human successes. Human Factors Analysis and Classification System [Ref. 23] and LOSA use databases that track systemic deficiencies without a coded structure for recovery (positive) factors. Rather, positive data are captured in the form of written narratives describing what the crew did well [Ref. 24]. These narratives are unstructured, capture positive behaviors in a happenstance fashion, and can be cumbersome to analyze by currently available techniques. Continued advances in natural language processing may represent an approach to facilitate analysis of narrative data [Ref. 25]. New approaches to analysis of narrative data will become increasingly important as the volume of collected data increases. For example, NASA's Aviation Safety Reporting System (ASRS) intakes an average of 1,964 new reports per week [Ref. 26].

Initial attempts have been made to develop taxonomies to capture positive behavior (Appendix B). The International Civil Aviation Organization (ICAO) and the CAST developed a high-level classification of positive concepts, including definitions [Ref. 27]. The French Voluntary

Reporting System explicitly mentions dedicated fields for positive factors [Ref. 28]. Additionally, ASRS coding forms support the recording of “detection and resolutory actions” that can be linked to some occurrences recorded in the ASRS database. This list contains 34 possible items describing, at various levels (e.g., flight crew, controller, aircraft or other), an action or an event that was involved in the resolution of an adverse situation.

However, the emphasis in these taxonomies largely describes specific behaviors which, in the absence of associated situational factors, may or may not represent safety-producing performance (e.g., flight canceled/delayed, rejected takeoff, proper following of radio procedures). Rejecting a takeoff, for example, is not universally safer than continuing a takeoff, but depends on the specific context in which that decision was made. Thus, these classifications risk conflating operator behaviors with outcomes: Positive behaviors yield positive outcomes, and negative behaviors yield negative outcomes. While it is likely that a strong positive correlation exists between outcomes and preceding behaviors, operators may complete their missions despite risk-taking behaviors, particularly in ultra-safe systems that afford many opportunities to notice and mitigate these risks. Likewise, undesired outcomes can result despite appropriate operator actions. It is often during retrospective analysis that behaviors are identified as positive or negative, depending upon whether the outcome of the event was desired or undesired. Thus it is important to distinguish between behaviors that support resilient performance (i.e., universally desired behaviors) and behaviors that merely precede desired outcomes (i.e., behaviors which may or may not be desired) to support learning. This is critical in ensuring that the “right” lessons are learned from the “right” experiences.

Outcomes can result from many possible conditions or combinations of factors, some of which may be transient. Therefore, causes identified during post hoc analyses of events are often reconstructed rather than found, or may be traced back only until resources for analysis are exhausted. Tools such as root cause analysis and error chain analysis, which seek to break systems down into components and identify likely associated threats or failures, are in widespread use for mishap investigation [Ref. 29]. A previous NESC assessment concluded that reliance on linear causal models (e.g., root cause analysis, fault tree analysis, error chains) contributed to missed identification of cues associated with mishap events [Ref. 30]. The simplifying assumptions of these models can work well for technological systems that can be decomposed into constituent parts, but they fail to accurately describe how success and failure occur in complex sociotechnical systems. Application of Safety II principles could enhance existing practices in mishap analysis in complex sociotechnical systems.

Despite the large, and growing, number and volume of data sources that describe performance in commercial aviation, efforts to integrate these data are relatively new. One effort is the Federal Aviation Administration’s (FAA) Aviation Safety Information Analysis and Sharing (ASIAS) System. This system contains multiple databases of safety data, and is intended to promote the “open exchange of safety information in order to continuously improve aviation safety” [Ref. 31]. However, much of the data shared through ASIAS has been de-identified by removing time-of-flight and airline information. While this can protect operators and organizations from punitive or legal action, de-identification prohibits reliable integration with other data sources, which depend upon that information to sync disparate databases.

The global aviation enterprise is a complex, adaptive system full of variability. Daily operational flexibility by human and organizations within the governing technologies, functions, and procedures keep the system extraordinarily safe and efficient. Additional approaches are needed to systematically detect and analyze the resilient behaviors and strategies employed within this complex sociotechnical system. Understanding how operators contribute toward creating and maintaining safety will better prepare the research and operational communities to address the following questions:

- What factors contribute to the current safety record in commercial aviation?
- How robust is system safety to the intended and unintended consequences of changes in technology, functions, and procedures?
- How should system safety be evaluated during design of new technologies, functions, and procedures?
- How can operational system safety metrics be developed that are sensitive to early signals of departure from expected operations which could either negatively or positively affect system performance?

2.4 Scope

The civil aviation system includes humans in many roles, including but not limited to:

- Pilots
- Cabin crew
- ATC tower controllers
- Terminal radar approach controllers
- En route controllers
- Traffic management units
- ATC system command center personnel
- Airline dispatchers
- Aircraft maintainers
- Ramp personnel

The assessment team focused on commercial air transport operations, and specifically on the roles of pilots and air traffic control tower controllers. The assessment team focused primarily on the behavior of individual operators rather than teams or organizations. Furthermore, the assessment team limited analysis to datasets that were rapidly obtainable and/or already familiar to the team members. The assessment was scoped in these ways to accommodate the time constraints of the assessment, while providing a representative and extensible approach to assessing human contributions to safety in civil aviation.

3.0 Data Analysis

The assessment team conducted several tasks as part of data analysis, including identification of:

- Candidate data and data sources that could be included in analysis of resilient operator performance in routine operations.
- A theoretical framework for structuring and organizing identified resilient actions.
- A candidate list of human actions that reflect resilient performance in routine operations.
- Techniques for exploring the identified behaviors in objective system-based datasets.

To begin to understand and characterize resilient behaviors, the assessment team leveraged multiple data sources in complementary ways. First, potential data sources were identified and down-selected for inclusion in the NESC analysis. Second, an existing theoretical framework for resilience strategies was identified to guide data acquisition strategies and to structure and organize identified behaviors. Third, subjective data sources were used to identify insights into specific resilient operator behaviors. Fourth, the assessment team adapted existing data analytics approaches to explore how these resilient strategies might manifest in system-based data.

The results of these analyses were used to generate findings, observations, and NESC recommendations, identified in Section 4, with regard to feasibility of and guidance for data collection and analysis of resilient operator behavior in routine situations.

3.1 Identification of Candidate Data and Data Sources

Systematic collection of resilient behaviors in everyday operations requires leveraging existing potential data sources and, where there are gaps, identifying new opportunities to collect information. Data sources can be broadly grouped into three classes:

- ***Operator-based data*** includes information from self-reports of events (e.g., ASRS) and interviews with operators about their lived experiences.
- ***Observer-based data*** includes information from direct inspection of operators by trained observers (e.g., LOSA, NOSS).
- ***System-based data*** includes information recorded or displayed by a system about its state or environment (e.g., flight data recorder data).

The assessment team identified several potential sources of aviation data that could apply to operator resilient performance. These data sources were classified according to their availability, temporal resolution, source for obtaining data, data type (e.g., continuous, categorical, textual, auditory), data format, and generating source (i.e., system, observer, operator). The identified data sources, listed in Appendix C, provide a representative although not complete sample.

Each of these classes of data affords insights into different aspects of routine performance. The current analysis focused on insights that could be derived from operator-based subjective operator interviews and system-based objective flight recorder data. Subject interviews can provide valuable insights into operator state, intentions, goals, pressures, and knowledge, whereas flight recorder data can provide occurrence rates and objective validation of subjectively-described events. Taken together, subjective and objective data can create a more complete and interpretable description of work-as-done.

3.2 Identification of Factors that Reflect Resilient Actions

Systematic analysis of resilient behaviors in routine operations requires a framework for organizing and making sense of the collected data. The assessment team adapted a framework developed by Rankin et al. [Ref. 16] to describe strategies that support resilient performance. This framework was developed as a tool to structure, organize, and analyze operator behaviors in daily work situations in complex systems. The framework focuses on identification of strategies; the enablers and conditions under which the strategy was applied; the properties of resilient performance supported by the strategy; and the source of the strategy (see Figure 3.2-1).

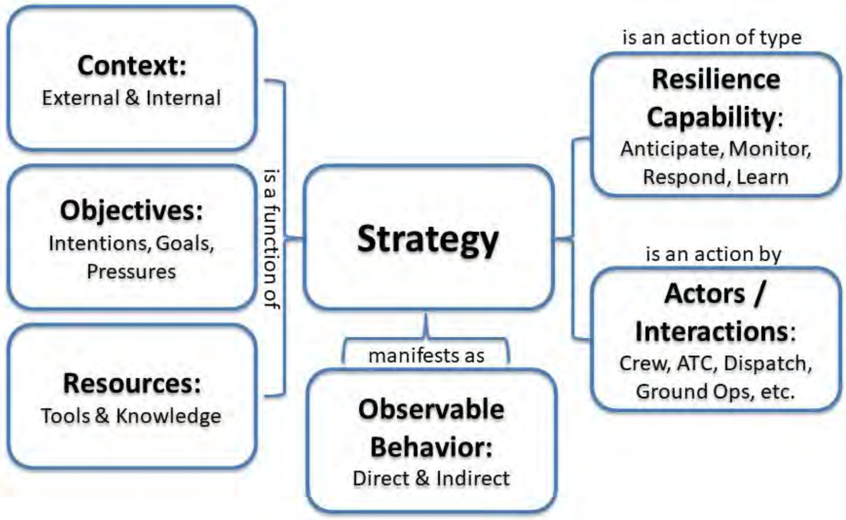


Figure 3.2-1. Strategy framework.

The primary elements of the framework include:

- *Strategy* – The adaptations and countermeasures used to cope with variations in the dynamic environment.
- *Objective* – Intentions and goals, including any competing/conflicting pressures or goals.
- *Context* – Forces and situational conditions under which the strategy is conducted. Includes external (e.g., weather) and internal (e.g., organizational pressures) forces that act on a system to produce the “trade space” in which work is performed.
- *Resources* – The necessary conditions for successfully implementing the strategy. Includes tools (e.g., hardware, software, automation) and knowledge (e.g., training, experience, creativity). Helps identify what enables or hinders implementation of the strategy.
- *Resilience Capability* – Refers to the four capabilities of resilient performance (i.e., anticipating, monitoring, responding, and learning [Ref. 20]. A strategy may pertain to one or more of these capabilities.
- *Actors/Interactions* – Refers to where within the system the strategy is designed and executed ranging from locally to globally. Actors can be defined in terms of areas of application (e.g., flight deck operations, dispatch, maintenance, ATC, ground operations, infrastructure, regulatory, etc.), and in terms of organizational level and necessary interactions (e.g., individual, team, organization, business sector, etc.) [Ref. 32].

Literature is sparse in terms of operational strategies associated with the four capabilities of resilient performance. Lay et al. [Ref. 11] identified strategies employed by organizations within the context of power plant maintenance, including:

- Anticipate knowledge gaps and needs.
- Anticipate resource gaps and needs.
- (Monitor) Support processes of sense-making.
- (Monitor) Support reflective processes.
- (Respond) Manage deployed resources.
- (Respond) Provision of extra resources.
- (Respond) Manage priorities.
- (Learn) Use questions to trigger learning.

These strategies were used by the assessment team as a representative starting point for identifying strategies employed by operators in a civil aviation context, as described in Sections 3.3 and 3.4.

3.3 Identification of Candidate Actions by Operators that Reflect Resilient Performance

Pilot and ATC tower controllers were interviewed by assessment team members using interview protocols designed to elicit specific examples of resilient performance in routine operational situations. This approach focused on identifying behaviors and strategies based on the specific lived experience of the participants in an attempt to focus as closely as possible on work-as-done rather than work-as-imagined. In contrast, previous research on pilot and controller resilient performance [Ref. 33] presented operators in a group setting with descriptions of a disturbance, then asked how they would detect and deal with the situation. While many of the presented disturbances were common, whether a participant had actually experienced each of the disturbances was not reported.

For this assessment, pilot and controller participants were recruited in different ways, which affected the opportunity for data collection from each group. Therefore, the pilot and controller data collections are addressed separately.

3.3.1 Airline Pilot Interviews

Method

The objective for the airline pilot interviews was to obtain data that would allow opportunity to: (a) identify strategies/behaviors that exhibit emergent resilience properties; (b) identify methods/approaches for extracting strategies/behaviors from existing data sources; and (c) identify gaps/opportunities for future data collection.

Participants

Twenty-one airline pilots were recruited to participate. All participants were employed by a major airline operating under Federal Aviation Regulations part 121. Participants were identified based on their availability and willingness to participate and were not remunerated financially. Pilot participants were interviewed at a US airport during their break times between flights. In an effort to be as sensitive as possible to their available time, demographic information about the participants was not obtained. All interviews were conducted under approval from NASA's Institutional Review Board (IRB). Confidentiality was maintained through use of subject numbers not associated with participant names or other personally identifying information.

Materials

Interviews were recorded using commercial off-the-shelf software on a laptop computer using the internal microphone. An assessment team member listened to the recorded files and created a transcribed version using Microsoft Word.

Procedure

Participants were interviewed in an airport conference room with the door closed. First, participants reviewed and signed a consent form explaining the nature of the study and their voluntary participation. Next, the introductory statement and initial question was read aloud by an assessment team member (see Appendix D). Two team members were present during the interviews; one took notes while the other operated the recording equipment and guided the interviews.

Participants were asked to describe a specific unplanned or unexpected event they had experienced in the course of routine operations. This initial question was designed to elicit an experience likely to include examples of resilient performance: anticipating, monitoring, responding, and learning [Ref. 20]. Follow-up questions for each of these areas were available to the interviewers to ensure they targeted those aspects of performance. Participants were not asked to quantify occurrence rates of events. In some cases, depending on time availability, multiple events were captured from a participant. The interview results provided the assessment team with: (1) descriptions of specific instances of pilots adapting, and (2) insights into the pilots' goals, motivations, pressures, and knowledge.

Results

Line-by-line coding of transcribed interviews was used to identify various elements of the strategies framework as described by the pilots. Concept mapping was used to describe interdependencies among these elements: the behaviors, resources, objectives, context elements, and resilience capabilities. An example concept map depicting the action of changing pace of operations for a slower preceding aircraft is illustrated in Figure 3.3.1-1.

Non-linearity in the system is described by the interconnectedness of the nodes in the concept map. In addition, disparate temporal ranges are depicted by the separate times at which learning and monitoring occurred. Some learning that occurred during this situation was not immediately related to it and was stored for later use, depending on the outcome of the event. In contrast, some previous learning was applied to this event from a time before the event happened.

Competing goals were discussed by participants describing the intent of other actors in the system. For example, pilots with the goal of a smooth approach were countered by ATC asking for a steeper than normal approach for traffic flow demands. Although this may not be ideal, it describes a complex domain that cannot always completely satisfy one user. Rather than a fixed ideal, participants described their goal system as a hierarchy, with safety being the foundation and other goals like helping, efficiency, and comfort taking a lower priority. This allows the users to negotiate based on their willingness to concede or the perceived risk that a concession might evoke.

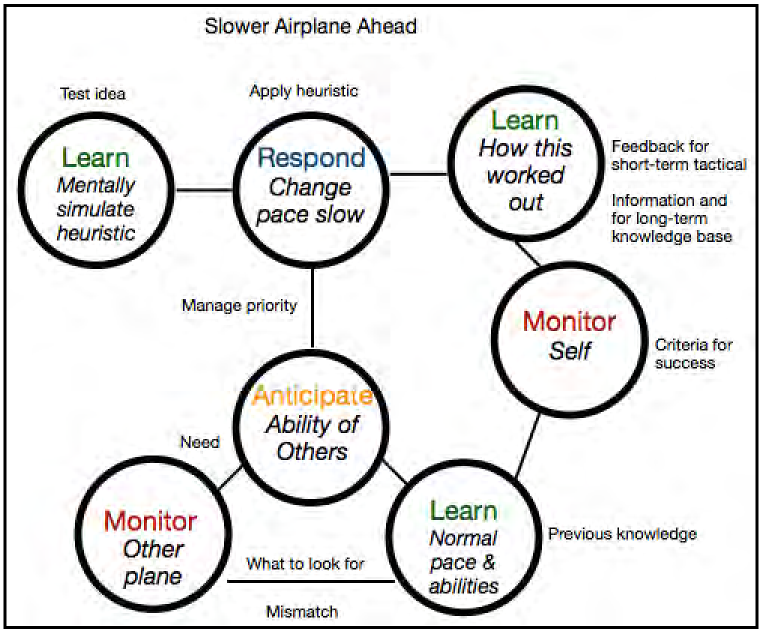


Figure 3.3.1-1. Concept map depicting change in operational pace in response to a slower preceding aircraft.

Several participants noted that sharing actionable information, amending plans so others could benefit, and teaching or compensating for other people was part of their job. These actions may seem counterintuitive in an environment in which airlines compete, but the processes and practices reported had established structures and expectations. This supports the notion that hierarchies of goals likely exist outside of prescribed regulations and policies. For example, pilots reported information over party-line frequencies to inform other aircraft that they had struck a bird. That information was immediately actionable to a plane behind them and actionable to ATC for warning other pilots and airport staff. However, this action did not directly benefit the plane that struck the bird, and, in fact required additional resources to communicate.

Insights from concept maps and assessment team discussions were used to identify strategies for resilient performance linked to the events and behaviors that the participants described (see Table 3.3.1-1). It is noteworthy that many of the reported behaviors associated with participants' ability to anticipate, monitor, and respond require leveraging experience-based information. However, no methods exist to systematically report or capture this information. This is a missed opportunity for developing training, data systems, and procedures whereby operators could systematically benefit from others' lived experiences, not just their own.

Table 3.3.1-1. Identified resilient performance strategies employed by operators in routine aviation contexts.

	Strategy	Description	Reported Behaviors
Anticipate	Anticipate procedure limits.	Predict when the current context inhibits the normal use of a procedure, regulation, policy, or norm.	Anticipate when formal procedure (e.g., STAR) will not work.
	Anticipate knowledge gaps.	Predict whether a crew member or other actor in the system lacks required minimum knowledge.	Anticipate others' intent.
	Anticipate resource gaps.	Compare the level of resources (e.g., fuel, time, workload, etc.) to perceived needs from experience or training.	Anticipate need to "buy time."
			Compare time needed and time available for action.
	Prepare alternate plan and identify conditions for triggering.	Have an actionable plan ready within the time available. Predict available time and what might work.	Request land at alternate airport (e.g., due to weather) or runway.
			Go-around (e.g., if preceding aircraft does not exit the runway.
Monitor	Monitor environment for cues that signal a change from normal operations.	Identify triggering variables that signal something has changed from what was expected.	Monitor for "non-standard" signals/cues.
			Monitor for deviations from normal pace of operations.
			Monitor for deviations from normal control "feel."
	Monitor environment for cues that signal a need to adjust or deviate from current plan.	Identify triggering variables that signal something will not continue to work as planned.	Monitor party-line communications.
			Monitor locations of aircraft in the area.
			Monitor others' workload.
			Monitor for cues (e.g., voice) of flight crew's experience and stress level.
	Monitor own internal state.	Perform self-assessment of physiological state, emotional state, workload, or knowledge.	Monitor own workload.
			Monitor own limits and capabilities.
Respond	Adjust current plan to accommodate others.	Help others in the system by changing timing or other action.	Change speed to accommodate other aircraft.
	Adjust or deviate from current plan based on risk assessment.	Change plan based on monitoring of triggers associated with safety boundaries.	Deviate from procedure based on risk assessment.
	Negotiate adjustment or deviation from current plan.	Work with others to accommodate competing goals and come to a solution that is mutually acceptable to all	Negotiate route change.

	Defer adjusting or deviating from plan to collect more information.	Continue with current plan because acting without critical information could make situation worse.	Defer action until more information is available.
	Manage available resources.	Preserve finite resources by adjusting controllable aspects of the situation.	Divide/take/give tasks to balance workload.
			Outsource tasks to automation or decision aids.
	Recruit additional resources.	Obtain resources locally or externally.	Ask others for assistance/resources.
			Ask others for information and/or clarification.
	Manage priorities.	Change goals, task order, task content, or pace of operation to accommodate resource limitations.	Adjust timing or speed of tasks based on operation pace and workload.
			Balance competing goals of formal expectations (e.g., follow procedures, maintain margins, comfort of passengers/pilots, smooth ride, reduce workload)
			Shed/abbreviate tasks to fit timeline/pace of operations.
Learn	Leverage experience and learning to modify or deviate from plan.	Compare formal expectations and experience to current situation to develop real-time assessment of acceptability or risk.	Predict likelihood of events based on past experience.
			Consider historical occurrences with similar contexts.
			Mentally simulate procedure.
			Use heuristics/rule-of-thumb guidelines.
	Understand formal expectations.	Understand applicability of laws, procedures, policies, and cultural norms.	Know and apply formal expectations (e.g., procedures, regulations, company policies, weather forecasting).
	Facilitate others' learning	Share information with others to increase their immediate understanding and long-term learning.	Teach other crew- or team-members.
			Share actionable information with other aircraft/ATC.

3.3.2 ATC Tower Controller Interviews

As with the pilot interviews, the primary objective for the ATC controller interviews was to obtain data that would allow opportunity to: (a) identify strategies/behaviors that exhibit emergent resilience properties; (b) identify methods/approaches for extracting strategies/behaviors from existing data sources; and (c) identify gaps/opportunities for future data collection.

Method

Participants

Twelve air traffic controllers were recruited to participate in the study. Controller interviews were conducted at NASA Langley Research Center, and participants were off-duty on the day of their interview. All of the controller participants were highly experienced, with an average experience level of 33 years.

All interviews were conducted under approval from NASA's IRB. Confidentiality was maintained through use of subject numbers not associated with participant names or other personally identifying information.

Materials

Interviews were recorded using commercial off-the-shelf software on a laptop computer using the internal microphone. An assessment team member transcribed the recorded files using Microsoft® Word®. Questionnaires were also administered (see Appendix E).

Procedure

Interviews were conducted over three days, with one group of four controllers participating each day. The participants were assigned randomly to one of three groups of four controllers to facilitate data collection. For each group, the session began with participants reviewing and signing an informed consent form explaining the objectives and details of the study and voluntary participation.

Participants were provided with an introductory presentation that included an overview of the concept and principles of resilience engineering including descriptions of each of the four "cornerstones" of resilience performance: anticipating, monitoring, responding, and learning [Ref. 20].

Each participant was subsequently interviewed individually by members of the assessment team using a semi-structured protocol that solicited examples of resilient behaviors that they had observed or experienced as controllers (see Appendix D). Each interview lasted approximately 45 minutes. Participants also completed a written questionnaire while the other controllers in their group were being individually interviewed.

After completion of the individual interviews and questionnaires, all four controllers participated in a group discussion facilitated by assessment team members. The objective of the group discussion was to collect information about strategies used by controllers to promote system resilience, rather than those based on specific events or episodes.

Results

The individual interview protocol used with ATC participants was adapted directly from the protocol developed for pilot participants. Minor changes were made to the wording of the initial probe question to reflect a focus on ATC rather than flight deck operations. Strategies and behaviors indicating resilient performance were identified from interview notes and recordings, and are integrated with data from pilot participants in Table 3.3.2-1.

In responses to the administered questionnaire, all participants indicated that they exhibited resilient performance on the job as air traffic controllers. The result indicated that 83% (N = 10) estimated “at least once per session,” where a “session” refers to each one of the multiple times that a controller works at their position during an 8-hour daily work shift. “At least once every 2 weeks” (N = 1) and “at least once daily” (N = 1) were the other responses.

Results showed that 75% of participants (N = 9) stated that they make traffic management decisions not explicitly specified by policies or procedures (e.g., FAA Order JO 7110.65, facilities standard operating procedures, letters of agreement) “at least once per week” with 58% (N = 7) estimating the occurrence to be “at least once daily.” The participants were further asked, “How many of these decision would you categorize as ‘resilient’ decisions”? The responses were: 75% estimated “more than 50%” (N = 9), and 58% indicated “more than 90%” (N = 7).

Participants were asked to indicate the prevalence of eight behavioral components associated with resilience principles, as identified by Heese, Kallus, and Kolodej [Ref. 34]. Participants’ estimates of the frequencies of these behavior are shown in Table 3.3.2-1.

Table 3.3.2-1. Participant estimates of frequency of behaviors associated with resilience.

	Less Than Once Per Month	At Least Once Per Month	At Least Once Every 2 Weeks	At Least Once Per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session
Finding goal-directed and proactive solutions that require trading for conflicting goals such as capacity, efficiency, and costs			1		4	3	4
Anticipating needs for planning and coordination			1		1	2	8
Using judgment for improvisation of standard operating procedures for safety/efficiency/capacity purposes				2	2	2	6
Inventing work-around procedures and techniques that work better for actual practice	1	2		4	2	1	2
Applying flexibility to increase safety buffers and defensive controlling for buffering capacity, margins, and added safety tolerance				1		4	7
Providing team support and adaptive capacity as required			1		1	2	8
Utilizing resources as required, such as consulting written/printed documentation (manuals, procedures) or electronic information (e.g., Aeronautical Information Publications online, route charts for alternative waypoints)	1			4	2	4	1
Developing strategies for managing workload	1	1		2	3	1	4

Although controllers reported that they exhibit behaviors that support resilient performance on a routine basis, there is little focus on collecting, analyzing, or characterizing these behaviors.

One objective of the focus group discussion was to identify how controllers use event reporting systems and their applicability and utility for obtaining data on resilient behaviors through these systems. All of the participants in the assessment sample stated that they had filed incident reports through one or more safety reporting systems. However, none of the participants stated that their narrative descriptions focused on detailing positive behaviors that demonstrate resilient performance. During focus group discussions, all participants agreed that there should exist a system to collect inputs describing positive incidents, when things go right. The controllers stated that the FAA has an outlet in which controllers can file incidents that reflect "going above and beyond". However the culture within the ATC community that "it was their job" to adapt to routine disturbances, and showing resilient behavior was "what they get paid to do" created barriers to submitting positive event reports. Participants believed that, in the current cultural environment, controllers might be reluctant to file positive incidents except in the case of extraordinary performance (e.g., talking a novice pilot through clouds to land in bad weather).

A barrier to positive event reporting is that reporting systems are structured to capture negative events (i.e., when things go wrong). For example, ASRS provides several Event Assessment codes to categorize type of anomaly (14 categories), primary problem (17 categories), and contributing problem areas (16 categories). There is also a code for the resulting action, but only four categories are provided for ATC: (1) issued advisory/alert, (2) issued new clearance, (3) provided assistance, and (4) separated traffic [Ref. 28, see Appendix B]. The numbers of categories for characterizing the reported events suggest an emphasis on describing what went wrong, not on the actions that controllers might take to resolve the reported anomaly.

Therefore, the assessment team asked participants to list additional actions controllers take in response to anomalies that might be used to supplement the existing categories, resulting in the following list:

- Corrected read-back
- Provided weather information
- Intervened to prevent unsafe situation
- Anticipated potential problem
- Developed strategic plan to avoid a problem
- Adjusted traffic flow
- Cancelled clearance (e.g., takeoff or landing)
- Coordinated support
- Anticipated needs of pilot
- Anticipated flow issues
- Verified pilot intentions
- Repeated transmission for emphasis
- Communicated with professionalism/clarity
- Offered options/alternatives
- Monitored for changes
- Anticipated and adjusted for unexpected event

In the group discussion, controller participants suggested providing guided assistance for furnishing narrative details to ensure that filed reports focused on desired aspects or features of resilient performance.

3.4 Techniques for Exploring Identified Factors in System-Based Data

Although operator-based data (e.g., structured interviews and self-reports) can provide rich data with regard to intentions, goals, pressures, or operator state, recollection-based approaches are subject to the reconstructive nature of human memory [Ref. 35]. Examination of system-based objective data can substantiate subjective accounts and provide quantifiable details that are difficult or impossible to obtain from subjective data alone. Furthermore, system-based data provide a direct link to work-as-done, because the data can be the product of operator actions.

Based on the strategies and behaviors identified through operator interviews (see Table 3.3.1-1), the assessment team considered how these strategies might manifest in aircraft flight data. Two candidate strategies were selected for exploration: 1) “anticipate resource gaps” (Anticipate), and 2) “manage priorities” (Respond). To explore the “anticipate resource gaps” strategy, the assessment team examined pilot behaviors associated with taking preemptive actions to prevent unstable approaches. To explore the “manage priorities” strategy, the assessment team examined pilot behaviors associated with the timing of performing pre-takeoff control surface checks. Each of these analyses is described in detail in the following section.

3.4.1 Exploring “Anticipate Resource Gaps” Strategy in Aircraft Flight Data

Interview participants reported behaviors in which they used their experience to anticipate when resources will approach functional boundaries (e.g., when will I run out of fuel, time, or space to execute a planned maneuver). That is, operators reported proactively seeking to maintain “the cushion of potential actions and additional resources that allow the system to continue functioning and adapting,” known as *margins of maneuver* [Ref. 11]. The assessment team posited that a pilot’s use of this strategy might manifest in objective aircraft flight data as the pilot taking action to preempt an adverse state (i.e., a state indicating that one or more resources had reached their functional boundaries). In the current case study, the assessment team focused on adjustments made by pilots during descent to preempt a high-speed exceedance at 1000 feet.

This approach leverages the idea of searching for degraded states that may arise during operations and detecting when these states are resolved (i.e., the operation is no longer in a degraded state). The assessment team identified degraded states using a machine-learning algorithm called deep temporal multiple instance learning [Ref. 36]. This algorithm was designed to detect states ahead of a predefined known adverse event, as defined/validated by subject matter experts, that has a high probability of predicting that adverse event. These states indicate a precursor to that event if the degraded state is not resolved. DT-MIL (Deep Temporal - Multiple Instance Learning) uses a deep learning neural network to build a model that classifies multivariate time series as resulting in the adverse event or not.

The DT-MIL algorithm was implemented in Python, using Anaconda 2.7 with Keras and Tensorflow modules for parallel processing. The algorithm provided a continuous measurement of the precursor probability that may rise and fall based on the states that it detected throughout the time series. When the precursor probability decreased, this served as an indication that an action may have been taken to shift the operation to a more “nominal state” in which resource margins were maintained. This shift could be quantified by comparing the difference between the degraded state and the new resolved state.

This method was demonstrated using FOQA data. Commercial airlines with FOQA programs use data from flight data recorders to monitor daily operations. These data are analyzed using predefined thresholds to flag and trend known adverse events of interest to the airline. The adverse event used in this example was a high speed exceedance at 1000 feet (ft.). This is one of a handful of FOQA flags that are monitored to identify a category of unstable approaches. A sample of 500 adverse event flights and 500 non-event flights were analyzed. Each flight contained 300 variables, of which 60 were selected using domain knowledge and automated feature selection based on Granger causality [Ref. 37]. The algorithm was randomly split into 50% training, 30% validation, and 20% testing. Because this assessment focused on preemptive actions, the non-event flights were examined for high precursor probabilities. Two examples of flights that resulted in high precursor probabilities followed by the lowering of probabilities are shown in Figures 3.4.1-1 and 3.4.1-2. The x-axis measures distance in nautical miles (NM) from the point at which the aircraft reaches 1000 ft. altitude. The solid blue line is the time series trace for each of the selected parameters that describe the precursor. The black dotted lines indicate the 10^{th}-90^{th} percentiles of the non-event data for each parameter for 0.25 NM binned distances to the event. The bottom right trace is the computed precursor score that DT-MIL provided for each sample of the time series. Samples for which the precursor score was greater than 0.5 are marked with red dots in the parameter traces and are considered high-probability precursors of a high-speed exceedance at the end of the time series. The shaded green region in the precursor score plot represents the event of interest, in which a degraded state was identified and potential for a preemptive action was indicated.

In Figure 3.4.1-1, the primary flight display (PFD) selected speed was higher than in the nominal distribution, which begins to step down at this point in the flight. When the PFD selected speed was reset to a lower value, the aircraft pitched up slightly and returned to nominal range, resulting in a decrease in airspeed. At this point the precursor probability sharply decreased. The algorithm identified a state that, if left uncorrected, had a high probability of leading to the high-speed exceedance. One insight that can be gained from this scenario is that the system is dynamically variable and flexibility is needed to safely accommodate other needs of the system. This may require pilots to drift outside the “normal” operating bounds as defined by typical flights, but the system can continue to safely operate by maintaining the resource margins necessary to adapt to changing and even unexpected demands.

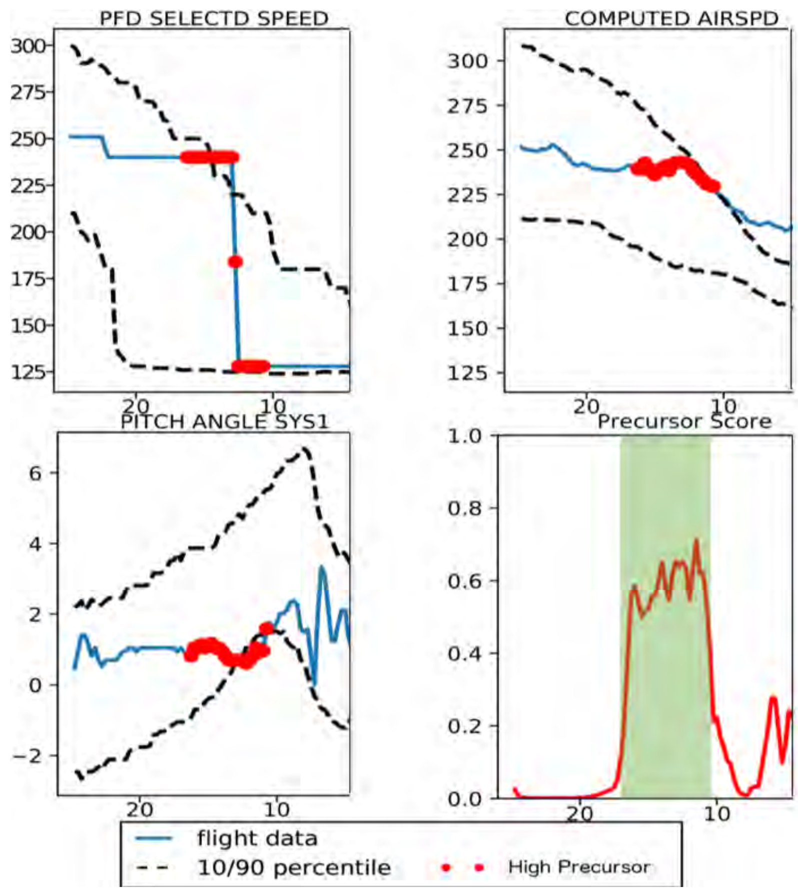


Figure 3.4.1-1.Time series plots for PFD selected speed, pitch angle, computed airspeed, and precursor score are depicted for a flight in which a preemptive action (i.e., resetting speed) was taken to avoid a high-speed exceedance at 1000 ft.

In Figure 3.4.1-2, the descent rate, captured by the vertical speed, was significantly lower than the normal distribution at that point in the flight (i.e., more negative vertical speed indicates faster decent rate in feet/minute). During this period, the airspeed was trending upward toward the upper bound of the nominal distribution. At this point, the pilot slowed the aircraft's descent rate and the airspeed began to hold steady. Although the airspeed remained outside the normal distribution, the transfer of the aircraft's energy from potential (i.e., altitude) to kinetic (i.e., airspeed) reduced the probability of a high-speed exceedance adverse event. Although the airspeed was above the 90th percentile for non-event flights during the event snapshot, it did not trigger the FOQA exceedance flag at 1000 ft. Increasing speed to avoid triggering a high-speed exceedance later in the flight may seem counter-intuitive, but it is common for pilots to employ this technique of trading altitude for speed. When the aircraft's energy is transitioned from altitude to speed, there are more tools available to the pilot to reduce kinetic energy by introducing drag (e.g., through use of speed brakes, lowering landing gear, and/or deploying flaps). These energy-bleeding practices can significantly slow the aircraft in a shorter amount of time and distance, compared with allowing the flight to descend on a shallower glide path to avoid increasing speed.

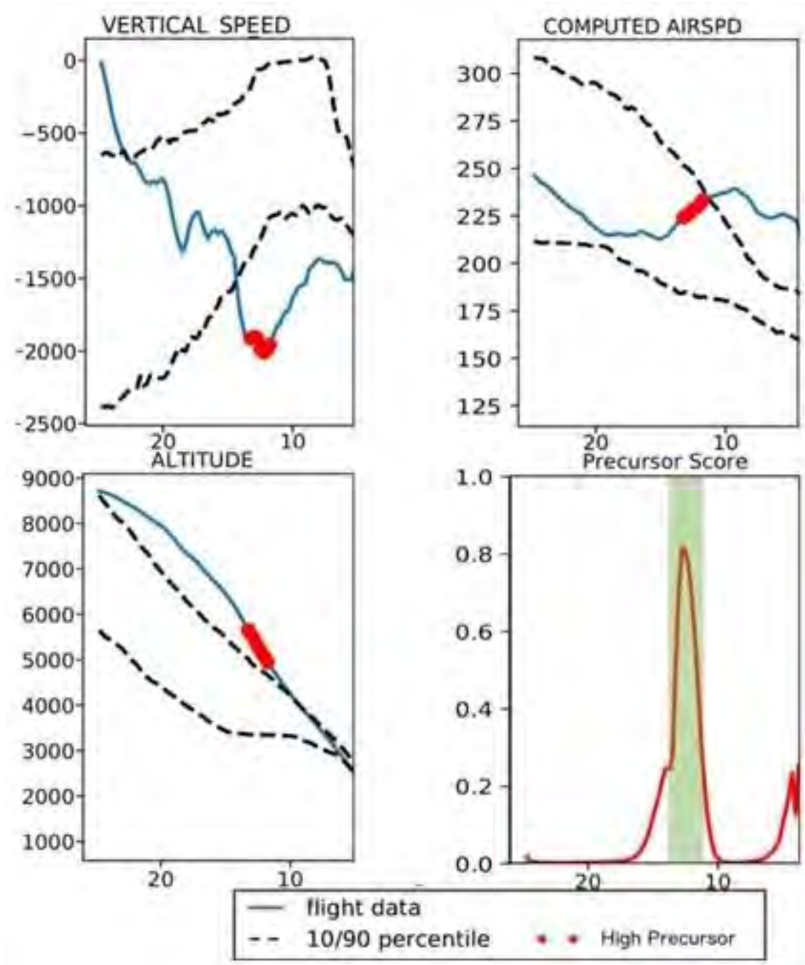


Figure 3.4.1-2. Time series plots for vertical speed, altitude, computed airspeed, and precursor score are depicted for a flight in which a preemptive action (i.e., slowing descent rate) was taken to avoid a high-speed exceedance at 1000 ft.

Both scenarios show situations in which a flight reached a state with an elevated probability of leading to an adverse event, but actions were taken in time to reduce this probability. Crew actions are the result of a combination of factors that include training, experience, intentions, and context. In this example, fundamental flight training provided pilots with knowledge of common available actions to get back on the vertical path. Experience in the aircraft and the ability to mentally simulate its future state was needed to anticipate a required action, choose an appropriate action, and choose the implementation timeframe for the action. Contextual factors such as weather, traffic, and crew coordination also likely influenced what, how, and when action was taken.

FOQA data can provide many quantitative details about operator and vehicle performance, but cannot provide information about the knowledge state, motivation, or broader context for the event. Why was the pilot flying the arrival at a higher than normal airspeed? What cues triggered the pilot to take action? If there were multiple appropriate actions that could have been taken, why did the pilot select that specific action? The answers to these questions could be obtained through observer- and operator-based data to supplement system-based data and provide a more complete understanding of work-as-done.

Although these "precursor" states are not considered unsafe in and of themselves, the algorithm has defined these states as indicative of patterns with a high probability of leading to an adverse event (e.g., an unstable approach). The fact that these flights did not lead to the adverse event indicates the resilience of the system, afforded by the actions of the pilots, to handle these situations and function safely.

3.4.2 Exploring "Manage Priorities" Strategy in Aircraft Flight Data

Interviews indicated that operators manage priorities by adjusting the timing or pace of operations to accommodate resource limitations (e.g., workload). To find evidence of pilots' use of this strategy in FOQA data, the assessment team examined pre-takeoff control surface checks in taxiing aircraft. In this procedure, the pilot checks the aircraft's control surfaces by moving/rotating them to their maximum positive and negative angles. This pre-flight check is performed on the rudder, ailerons, and elevators. While performing these checks prior to take-off is a procedural requirement, the specific timing and spatial location is left to the discretion of the operator. The assessment team reasoned that if pilots were strategically managing priorities, there would be detectable patterns in when or where they decided to perform the control surface check.

Evidence of pre-takeoff control surface checks were identified in FOQA data for departures at Barcelona-El Prat airport by looking for consecutive full-range motion in rudder angle, aileron angle, and elevator angle for aircraft on the airport surface. In this case study, the time and spatial location of the procedure were identified for 980 departures and plotted on an airport map (see Figure 3.4.2-1).

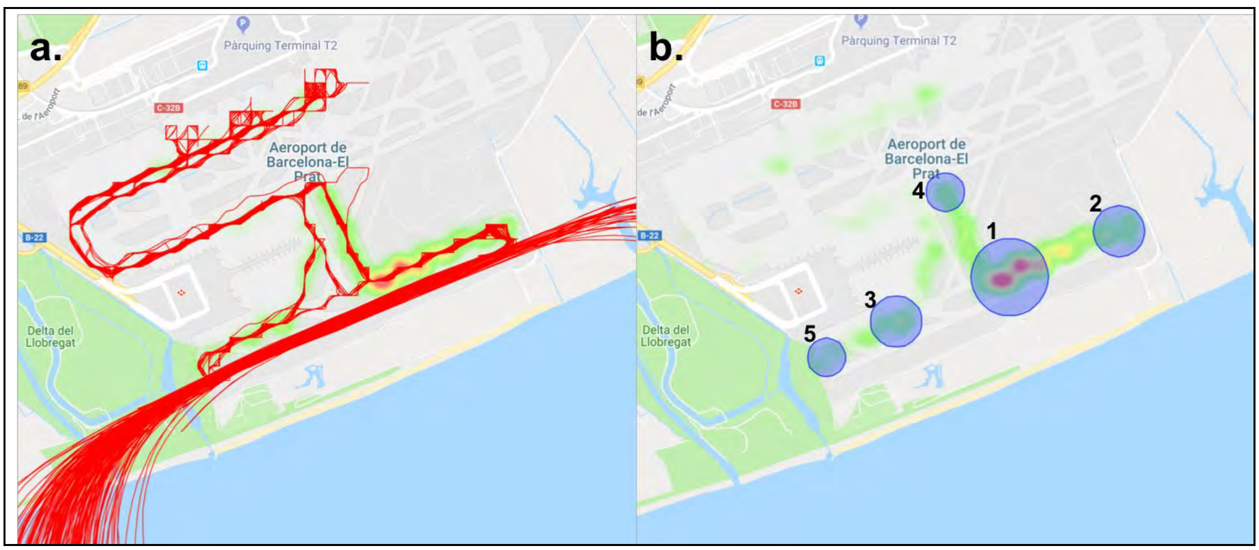


Figure 3.4.2-1. (a) Taxi routes of flights taking off from Barcelona airport. Routes are overlaid on a heat map indicating regions where control surface checks were performed. (b) Numbered regions indicate regions where control surface checks were most commonly performed. Numbering is in descending order of number of flight checks performed per 980 flights.

While there was not a specific time or place along the taxiway where the pilots performed the control surface check, clear patterns emerged. The five areas where the checks were performed most often were identified and examined (see Table 3.4.2-1). Taken together, these areas represented approximately two-thirds (67.7%) of the locations where control surface checks were made. Approximately half of the checks (48.8%) were performed at a single 90-degree intersection, as pilots were turning onto the taxiway parallel to the runway. Four of the five most common check areas were along this taxiway, and all of the top five areas represented intersections where pilots made 90-degree turns. None of the pilots performed the control surface check before starting to taxi.

Table 3.4.2-1. Flights that performed control surface check procedure following turn during taxi to runway. (Region numbers correspond to regions indicated in Figure 3.4.2-1.)

Full turn region no.	No. of flights	% of 980 flights
1	478	48.8
2	73	7.5
3	54	5.5
4	38	3.9
5	20	2.0
Total	**663**	**67.7**

Checking control surfaces is one of several tasks associated with preparing the aircraft for takeoff. To accomplish the control surface check, pilots must estimate how long they have until takeoff, how long each of their other tasks will take, and plan the timing of the check accordingly. Thus, the checks are likely to occur at different locations depending on contextual variables (e.g., traffic, airport familiarity, visibility, etc.). The ability to dynamically interleave tasks based on context is not explicitly trained, but rather a function of individual experience which helps pilots understand how traffic moves around airports, how long the procedures take for them to complete, and specific airport norms and customs.

Although FOQA data cannot directly reveal pilot intentions, the patterns of observed behaviors suggest several possible motivations. For example, pilots might check control surfaces in region 1 or 3 (See Figure 3.4.2-1.) because this is the portion of the taxi where pilots can likely see the runway and the line of aircraft, and they can make a high-confidence estimate that takeoff will happen soon. In addition, taxi workload is reduced here because taxiway navigation to the runway is assured at this point (i.e., only one plausible taxi path to the runway). Performing checks in regions 2 or 5 may be the result of high traffic volume and slower taxi speeds (i.e., waiting in a line of planes). Planes will be released at predictable intervals for takeoff, and the pilots could defer the checks based on the number of preceding airplanes. Performing checks in region 4 could be due to an unusually fast taxi that requires completing tasks earlier than normal to accommodate the pace of operations.

While the specific reasons for the timing of the control surface checks cannot be definitively determined from FOQA data, the existence of discernible patterns in the timing of behavior across pilots performing a task with discretionary timing parameters suggests that the observed performance variance occurred for strategic reasons. Targeted follow-up interviews with pilots who fly at this airport could provide critical details about the intentions, internal states, and knowledge behind the observed behaviors. This example further illustrates the synergistic relationship between subjective and objective data in understanding work-as-done and the strategies that human operators use to accomplish their work both safely and efficiently.

4.0 Findings, Observations, and NESC Recommendations

4.1 Findings

The following findings were identified:

F-1. NASA and industry planning and system design in aviation are based on Safety I principles and methods, focused on predicting and preventing errors.

F-2. Current industry and regulatory safety reporting processes and mechanisms are designed to capture events that degrade safety (e.g., violations, deviations, and non-compliance with rules and procedures; human errors; etc.), but not positive events that bolster safety.

F-3. Operators identified cultural barriers to reporting routine positive behaviors, because adapting to routine disturbances was seen as expected job performance.

F-4. Current observer-based approaches to data collection and analysis (e.g., LOSA, NOSS) do not systematically include resilient behaviors.

F-5. Many of the behaviors reported by pilots and controllers that are associated with their ability to anticipate, monitor, and respond require leveraging experience-based information that is not systematically reported or captured.

F-6. Existing operator behavior taxonomies conflate "positive" operator behaviors with "positive" operational outcomes and "negative" operator behaviors with "negative" operational outcomes.

F-7. Defining safety in terms of "things that go right" enabled new methods for exploring existing data.

F-8. Operators were able to introspect about and provide specific examples of resilient behaviors.

F-9. Evidence of operator strategies that promote resilient performance was identified in objective system-based data.

F-10. Subjective and objective data sources contributed different information toward building an understanding of pilots' and controllers' resilient performance.

F-11. Current approaches for safety data collection and analysis are not designed to integrate data from disparate data sources.

F-12. Subjective data sources are necessary to understand the rationale for actions observed in objective data.

4.2 Observations

The following observations were identified:

O-1. NASA's mishap investigation process is an area in which Safety II principles and methods that focus on behaviors that promote resilient performance could be applied to complement current approaches.

O-2. Information sharing among operators is critical for avoiding known risks and managing resources.

4.3 NESC Recommendations

The following NESC recommendations were identified and directed toward NASA ARMD's Transformative Aeronautics Concepts Program and Airspace Operations and Safety Program:

R-1. Define safety in terms of the presence of desired behaviors as well as the absence of undesired behaviors. *(F-1, F-7)*

- *Justification:* Defining safety only as the absence of undesired behavior creates an incomplete picture of safety, particularly in highly safe systems in which undesired behaviors or outcomes are rare. Examining both desired and undesired behaviors affords more opportunities for performance measurement, increasing sensitivity and confidence in system safety performance.

R-2. Leverage existing data to identify strategies and behaviors that build resource margins and prevent them from degrading. *(F-2)*

- *Justification:* Evidence of operators' resilient performance can be identified from many existing data sources that were designed to capture events that degrade safety.

R-3. Develop organization-level strategies that promote recognition and reporting of behaviors that support resilient performance. (*F-3)*

- *Justification:* New approaches to safety and risk management that focus on sustaining required operations under both expected and unexpected conditions can help organizations recognize the importance of learning from success as well as failure, and overcome cultural norms and organizational practices that fail to recognize or reward operators' adaptations to routine disturbances.

R-4. Develop expert-observer-based data collection tools to capture strategies and behaviors that support resilient performance. *(F-4)*

- *Justification:* Expert-observer-based approaches that leverage "threat and error management" models may not be sensitive to safety-producing actions taken by operators to anticipate and monitor events before threats ever manifest. Integrating approaches to identifying behaviors that support sustaining required operations outside of responding to or managing threats could increase systematic collection of resilient performance indicators.

R-5. Develop methods to collect and analyze operator-reported strategies and behaviors that support resilient performance. *(F-5, F-8)*

- *Justification:* Operators represent the primary source of data about intentions, goals, pressures, and knowledge states that support resilient performance strategies. Methods are needed to understand how operators develop and leverage expertise to support anticipating, monitoring for, and responding to disturbances and opportunities. A better understanding of these processes could create opportunities for developing training, data systems, and procedures whereby operators could systematically benefit from others' lived-experiences, not just their own.

R-6. Develop approaches to expand collection and facilitate analysis of resilient behaviors in adverse event reports. *(F-6)*

- *Justification:* First-hand adverse event reports typically include narrative descriptions that are potentially rich in describing operators' intentions, goals, pressures, and knowledge states. However, adverse event reporting forms do not emphasize or provide structure for reporting of resilient behaviors that may have helped to mitigate or resolve the event. Approaches are needed to provide guided assistance to reporters to ensure capture of resilient behaviors during adverse event reporting. In addition, manual analysis of these reports is slow and labor-intensive. Advances in natural-language processing technologies should be leveraged to develop automated tools that could assist in identification and analysis of resilient behaviors in adverse event narrative databases.

R-7. Refine data analytics approaches for exploiting FOQA data based on identified resilience strategies. *(F-9, F-10)*

- *Justification:* FOQA data represent a valuable source of objective and quantifiable operator performance data. However, few approaches to using these data to identify operator resilient performance have been developed or evaluated. Success of these approaches will depend on integrating information from FOQA analyses about what, when, and where events happened with information from subjective data sources about why and how those events happened.

R-8. Develop a system-level framework for integrating resilient performance data from observer-, operator-, and system-based sources. *(F-10, F-11, F-12)*

- *Justification:* While many individual data sources have been identified as providing valuable information, the collection of this data is often incidental rather than intentional. In some instances, a misalignment between the reasons for data collection and demands of data analysis can lead to practical challenges (e.g., between data de-identification and integration across data sources). Developing a thorough understanding of work-as-done requires a system-level approach for collecting and analyzing the diverse sources of data on the real-world resilient behavior of operators.

5.0 Definition of Terms

Finding	A relevant factual conclusion and/or issue that is within the assessment scope and that the team has rigorously based on data from their independent analyses, tests, inspections, and/or reviews of technical documentation.

Margin of maneuver	The cushion of potential actions and additional resources that allow the system to continue functioning and adapting [Ref. 11].
Observation	A noteworthy fact, issue, and/or risk, which may not be directly within the assessment scope, but could generate a separate issue or concern if not addressed. Alternatively, an observation can be a positive acknowledgement of a Center/Program/Project/Organization's operational structure, tools, and/or support provided.
Safety I	Safety is defined as a condition where the number of adverse outcomes is as low as possible [Ref. 12].
Safety II	Safety is defined as a system's ability to succeed under varying conditions, so that the number of intended and acceptable outcomes is as high as possible [Ref. 12].
Recommendation	A proposed measurable stakeholder action directly supported by specific Finding(s) and/or Observation(s) that will correct or mitigate an identified issue or risk.
Resilience	The ability of a system to sustain required operations under both expected and unexpected conditions by adjusting its functioning prior to, during, or following changes, disturbances, and opportunities [Ref. 2].
Resilience engineering	A paradigm for safety management that focuses on how to help people cope with complexity under pressure to achieve success [Ref. 38].
Sociotechnical system	People and equipment directly dependent on their material means and resources for outputs. The core interface consists of relations between a nonhuman system and a human system [Ref. 39].

6.0 References

1. Hollnagel, E. (2014). Safety-I and Safety-II: The Past and Future of Safety Management. Farnham, UK: Ashgate.
2. Hollnagel, E. (2016). Resilience Engineering. erikhollnagel.com/ideas/resilience-engineering.html. Accessed October 1, 2018.
3. Wald, A. (1980). A method of estimating plane vulnerability based on damage of survivors. *CRC 432*. Alexandria, VA: Center for Naval Analysis (Original work published 1943).
4. PARC/CAST Flight Deck Automation Working Group. (2013). Operational use of flight path management systems. Final Report of the Performance-based operations Aviation Rulemaking Committee/Commercial Aviation Safety Team Flight Deck Automation Working Group. Washington, DC: Federal Aviation Administration.
5. NASA. (2017a). NASA Aeronautics Strategic Implementation Plan (https://www.nasa.gov/aeroresearch/strategy).

6. NASA. (2017b). NASA Embraces Urban Air Mobility, Calls for Market Study. www.nasa.gov/aero/nasa-embraces-urban-air-mobility. Accessed October 1, 2018.
7. NASA. (2018). NASA, Uber to Explore Safety, Efficiency of Future Urban Airspace. www.nasa.gov/press-release/nasa-uber-to-explore-safety-efficiency-of-future-urban-airspace. Accessed Oct 1, 2018.
8. Uber Technologies, Inc. (2016). Fast-forwarding to the future of urban on-demand air transportation. www.uber.com/elevate.pdf. Accessed October 1, 2018.
9. Amalberti, R. 2001. The paradoxes of almost totally safe transportation systems. *Safety Science 37,* 109-126.
10. Nisula, J. (2018). A risk management framework for a complex adaptive transport system. Doctoral thesis: Universite de Toulouse.
11. Lay, E., Branlat, M., Woods, Z. (2015). A Practitioner's Experience Operationalizing Resilience Engineering. In C. Nemeth & I. Herrera (Eds.). Special issue on Resilience Engineering. *Reliability Engineering and System Safety*. 141. http://dx.doi.org/ 10.1016/ j.ress.2015.04.006.
12. Hollnagel E., Wears R.L., & Braithwaite J. (2015). From Safety-I to Safety-II: A White Paper. The Resilient Health Care Net: Published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia.
13. Civil Aviation Authority (2013). Global Fatal Accident Review 2002-2011. www.caa.co.uk/cap1036
14. Woods, D.D. & Dekker, S.W.A. (2000). Anticipating the effects of technological change: A new era of dynamics for Human Factors. *Theoretical Issues in Ergonomic Science, 1*(3), 272-282.
15. Cook, R. I. & Woods, D. D. (1996). Adapting to new technology in the operating room. *Human Factors, 38*(4), 593-613.
16. Rankin, A., Lundberg, J., Woltjer, R., Rollenhagen, C., & Hollnagel, E. (2014). Resilience in Everyday Operations. *Journal of Cognitive Engineering and Decision Making, 8*(1), pp. 78-97. doi:10.1177/1555343413498753.
17. Dekker, S. W. A. (2006). Resilience engineering: Chronicling the emergence of confused consensus. In E. Hollnagel, D. D. Woods & N. Leveson (Eds.), Resilience engineering: Concepts and precepts. Hampshire: Ashgate.
18. Electronic Code of Federal Regulations. (2018). Title 14, Chapter I, Subchapter F, Part 91. https://www.ecfr.gov/cgi-bin/text-idx?rgn=div5&node=14:2.0.1.3.10#se14.2.91_13. Accessed November 21, 2018.
19. Stewart, M., Matthews, B., Avrekh, I., Janakiraman, V. (2018) Variables influencing RNAV STAR adherence. IEEE/AIAA. Proceedings of the 37th Digital Avionics Systems Conference (DASC). London, UK.

20. Hollnagel, E. (2011). “RAG – The resilience analysis grid.” In: E. Hollnagel, J. Pariès, D.D. Woods and J. Wreathall (Eds). *Resilience Engineering in Practice. A Guidebook*. Farnham, UK: Ashgate.

21. Klinect, J.R., Wilhelm, J.A., Helmreich, R. L. Threat and Error Management: Data from Line Operations Safety Audits. In Proceedings of the Tenth International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University.

22. EUROCONTROL/FAA (2011, July). Ensuring Safe Performance in ATC Operations: Observational Safety Survey Approaches, A White Paper. EUROCONTROL/FAA Action Plan 15 Safety.

23. Wiegmann, D. A., & Shappell, S. A. (2003). *A human error approach to aviation accident analysis: The Human Factors Analysis and Classification System*. London, UK: Ashgate.

24. Kouwenhoven, T. (2010). Using positive human factors data. *Operator’s Guide to Human Factors in Aviation.* Retrieved from https://www.skybrary.aero/index.php?title=Using_“Positive”_Human_Factors_Data_(OGHFA_BN)&oldid=59417

25. Tanguy, L., Tulechki, N., Urieli, A., Hermann, E., & Raynal, C. (2016). Natural language processing for aviation safety reports: From classification to interactive analysis. *Computers in Industry, 78,* 80-95.

26. ASRS. (2018a). ASRS Program Briefing. https://asrs.arc.nasa.gov/docs/ASRS_ProgramBriefing.pdf. Accessed November 1, 2018.

27. ICAO/CAST. (2013). Positive taxonomy: Definitions and usage notes. www.intlaviationstandards.org/Documents/Positive%20Taxonomy.pdf. Accessed October 1, 2018.

28. ASRS. (2018b). Aviation Safety Reporting System: ASRS Database Online. asrs.arc.nasa.gov/search/database.html. Accessed October 1, 2018.Boudou, B., Pouliquen, Y., & Ferrante, O. (2006). Towards a human factor “positive” taxonomy for occurrence analysis. *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics.* Seattle, WA.

29. NPR 8621.1C. (2016). NASA Procedural Requirements for Mishap and Close Call Reporting, Investigating, and Recordkeeping. https://nodis3.gsfc.nasa.gov/npg_img/N_PR_8621_001C_/N_PR_8621_001C_.pdf. Accessed Nov 20, 2018.

30. NESC-RP-14-00990. (2017). Best Practices for Organizational Resilience in the International Space Station (ISS) Program.

31. FAA (2018). Aviation Safety Information Analysis and Sharing (ASIAS). www.asias.faa.gov/. Accessed October 1, 2018.

32. Owen, D., Gill, S., Courtney, H., McDonald, N., Liston, P., & Carrera, M. (2017). Exploring resilience in aviation and maritime transport – and how human factors can help. *The Resilience Shift. https://resilienceshift.org/wp-content/uploads/2017/10/026_Exploring-Resilience-in-Aviation-and-Maritime-Transport-And-How-Human-Factors-Can-Help.pdf. Accessed October 1, 2018.*

33. Stroeve, S.H., Van Doorn, B.A., & Everdij, M.H.C. (2013). The human contribution—Analysis of the human role in resilience in ATM. Deliverable D1.2 Resilience2050.eu. http://innaxis.org/resilience/the-project/

34. Heese, M., Kallus, W., & Kolodej, C. (2013). Assessing behavior towards organizational resilience in aviation. Proceedings of the 5th REA Symposium 2013 (pp. 67-74). Soesterberg, Netherlands: Resilience Engineering Association.

35. Schacter DL. 1989. Memory. In Foundations of Cognitive Science, ed. MI Posner, pp. 683–725. Cambridge, MA: MIT Press.

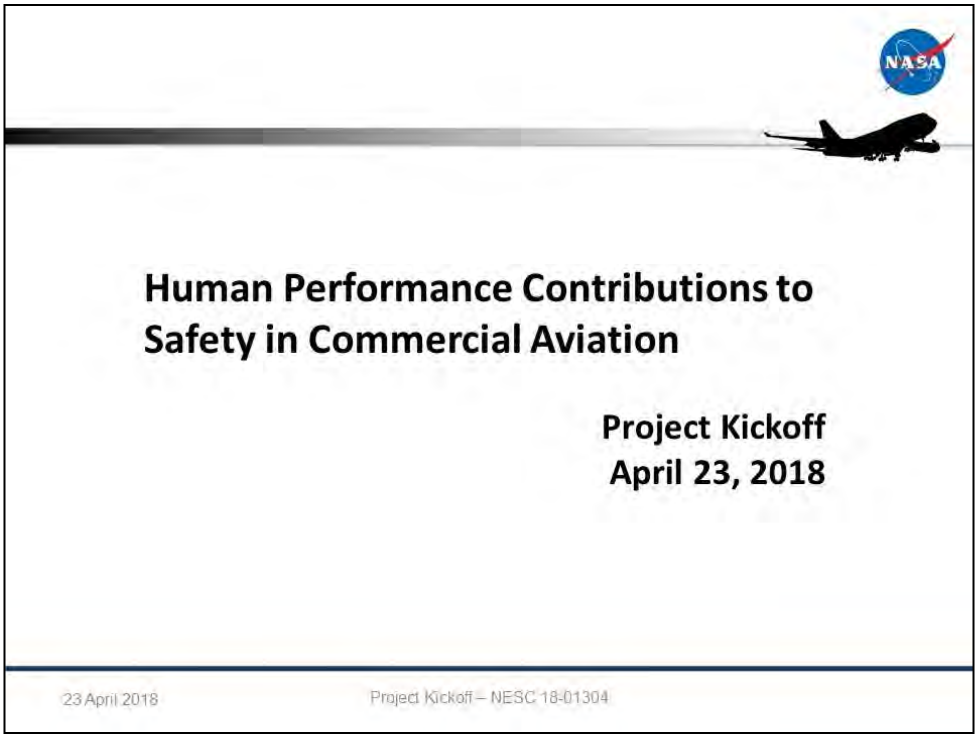
36. Janakiraman. V.M. 2018. "Explaining Aviation Safety Incidents Using Deep Temporal Multiple Instance Learning". In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (KDD '18). ACM, New York, NY, USA, 406-415. DOI: https://doi.org/10.1145/3219819.3219871.

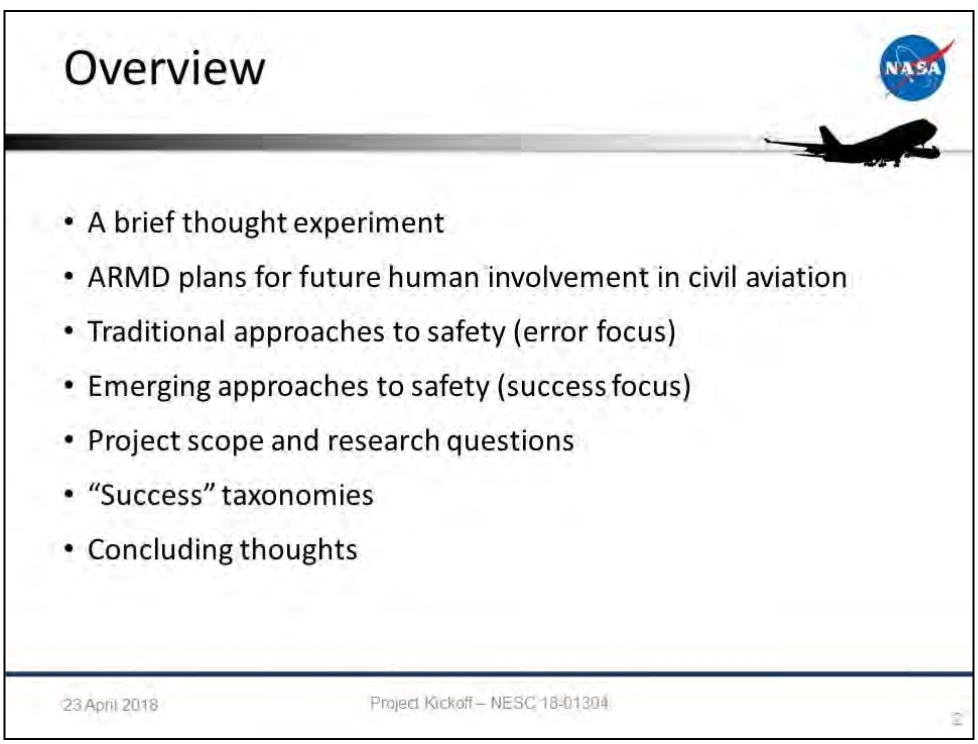
37. Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society, 424–438.

38. Woods, D.D., & Hollnagel, E. (2006). Prologue: Resilience engineering concepts. In *Resilience Engineering: Concepts and Precepts.* E. Hollnagel, D.D. Woods, N. Leveson (Eds.). Boca Raton, FL: CRC Press.

39. Trist, E. (1981, June). "The Evolution of Socio-Technical Systems: A conceptual and an action research program." In A. Van de Ven, W. Joyce (Eds.). *Perspectives on Organizational Design and Behavior*. New York: Wiley Interscience.

Appendix A: Assessment Kickoff Briefing, April 23, 2018, at NASA Ames Research Center





A thought experiment

- Human error has been implicated in 70% to 80% of accidents in civil and military aviation (Weigmann & Shappell, 2001).
- Pilots intervene to manage aircraft malfunctions on 20% of normal flights (PARC/CAST, 2013).
- World-wide jet data from 2007-2016 (Boeing, 2016)
  - 244 million departures
  - 388 accidents

A thought experiment

Outcome

Attributed to Human Intervention	Not Accident	Accident	
No	?	?	?
Yes	20%	80%	?
	?	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

A thought experiment

Outcome

Attributed to Human Intervention	Not Accident	Accident	
No	?	78	?
Yes	20%	310	?
	?	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

A thought experiment

Outcome

Attributed to Human Intervention	Not Accident	Accident	
No	?	78	?
Yes	20%	310	?
	243,999,612	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

A thought experiment

Outcome

Attributed to Human Intervention	Not Accident	Accident	
No	195,199,690	78	?
Yes	48,799,922	310	?
	243,999,612	388	244,000,000

- Human error implicated in 80% of accidents.
- Pilots manage malfunctions on 20% of normal flights.
- 388 accidents over 244M departures.

23 April 2018 Project Kickoff – NESC 18-01304 7

A thought experiment

Outcome

Attributed to Human Intervention	Not Accident	Accident	
No	195,199,690	78	195,199,768
Yes	48,799,922	310	48,800,232
	243,999,612	388	244,000,000

When we characterize safety only in terms of errors and failures, we ignore the vast majority of human impacts on the system.

23 April 2018 Project Kickoff – NESC 18-01304 8

ARMD plans

- ARMD planning activities propose reducing or removing human involvement in civil aviation transport, based on the development of "intelligent machine systems capable of operating in complex environments" (ARMD Strategic Implementation Plan, 2017).
- These plans presuppose that machines will be able to fulfill roles previously performed by humans.

23 April 2018 Project Kickoff – NESC 18-01304 9

Traditional approach to safety

- Traditional approaches to safety automation have focused on characterizing (human) errors and preventing them, but have not considered that humans are also responsible for safety successes.
- Traditional approaches have concluded that the safest path is full machine autonomy, but without analysis of safety success data, this conclusion is premature and does not provide essential design guidance.

23 April 2018 Project Kickoff – NESC 18-01304 10

Evidence of traditional thinking…

… in Urban Air Mobility planning (e.g., Uber Elevate whitepaper)

- "Safety can be measured by a number of dimensions"
  - Injuries, accidents, fatalities
- "Use autonomy technology to significantly reduce operator error"
- "To understand the path to improving safety for urban air transportation, we need to understand the root causes of historical crashes."
- "To fast-forward to the safest possible operational state for VTOL vehicles, network operators will be interested in the path that realizes full autonomy as quickly as possible."

23 April 2018 Project Kickoff – NESC 18-01304 11

Emerging focus on success

- Different approaches yield different insights

Traditional Approach to Safety	Resilience Engineering/Highly Reliable Organizing/Safety II Approach to Safety
Focus on predicting failure probabilities	Focus on preparing for the unpredicted
People are a liability—control, correct	People are an asset—learn, adapt
Variability is a threat—minimize it	Variability is normal—manage it
Focus on incident rates	Focus on learning
Focus on what we don't want: incidents and accidents	Focus on what we want: how safety is created
Procedures are complete and correct	Procedures are always under-specified and must be interpreted and adapted
Systems are well designed, work as designed, and are well maintained	Systems are complex and will degrade; there will always be flaws and glitches
Focus on "who", "when", "where", "what"	Focus on "why" and "how"

23 April 2018 Project Kickoff – NESC 18-01304 12

Emerging focus on success

- Humans' contributions to maintaining safety in complex environments are underappreciated.
- While data on human failures are routinely collected and analyzed, data are rarely collected on human successes.
- Thus, humans' contributions to safety go unmeasured or are minimized relative to human errors.
- This poses risks for future systems intended to replace human involvement or take on roles previously performed by humans.

Project goal

- While the overwhelming majority of flights result in safe and successful outcomes, the contribution of humans to those successes is poorly understood.
- This project aims to find and document "safety successes" made possible by human operators using proven tools, methods, and data-mining techniques.

Project scope

- This 6-month project will conduct a preliminary examination of a range of existing sources of aviation data to identify possible metrics and data gaps for capturing human contributions to safety.
- Methods for exploring and interpreting existing data will be examined, and methods for tapping into new data will be proposed.
- Open research questions include:
  - What actions do humans take in routine flights that contribute to mission safety?
  - What data that have already been collected will help provide insights into human contributions to mission safety?
  - What data that have already been collected could be used to test/validate our insights?
  - What additional data parameters could be feasibly collected that would provide additional insights into human contributions to safety?

Project tasks

- Identify Range of Existing and Available Datasets
- Identify Candidate List of Safety Actions Taken by Pilots in Routine Flight
- Identify Factors in Existing Datasets that Address Candidate Human Safety Actions
- Use Big Data Analysis/Mining Techniques to Explore Identified Factors in Existing Datasets
- Identify Additional Data Parameters that Could Provide Insights into Human Contributions to Safety and Feasibility of Data Collection
- Formulate Recommendations for Using Existing and Future Datasets Based on Analyses and Findings
- Submit Final Report by September 30, 2018

"Positive" taxonomies

Identify safety barriers that prevented the accident from occurring

- Why did this incident not turn into an accident?
- Was there equipment, a decision or a procedure that prevented an accident from occurring?
- Were the results of this occurrence simply a matter of chance or did some prior factor predispose a more positive outcome?
- Were there any positive human factors that mediated the outcome?
- In the case of an accident, could it have been more serious?
- What prevented it from resulting in more serious damage or injuries?

23 April 2018 Project Kickoff – NESC 18-01304 17

"Positive" taxonomies

Aviation Safety Reporting System

The flight crew...

- Became Reoriented
- Diverted
- FLC Overrode Automation
- Executed Go Around / Missed Approach
- Exited Penetrated Airspace
- FLC complied w / Automation / Advisory
- Inflight Shutdown
- Landed As Precaution
- Landed in Emergency Condition
- Overcame Equipment Problem
- Regained Aircraft Control
- Rejected Takeoff
- Requested ATC Assistance / Clarification
- Returned To Clearance
- Returned To Departure Airport
- Returned To Gate
- Took Evasive Action

CAST/ICAO Common Taxonomy Team (CICTT)

- Avoidance maneuver
- Decision to go around/land/reject takeoff
- Crew member assistance/intervention
- ATC assistance/intervention
- Accurate use of training/instructions/SOPs
- Radio procedures/communication
- Environment observation
- Visual/sensory detection/anticipation
- Sound reasoning/problem solving
- Accurate use of documentation
- Hardware/software safety net
- Provenance (Luck)

23 April 2018 Project Kickoff – NESC 18-01304 18

Concluding thoughts

- Data characterizing safety success behaviors will drive design requirements that help assure machines behave safely (rather than just assuring that they do NOT behave UN-safely).
- Without an adequate understanding of human contributions to safety, current proposed strategies for reducing or removing human involvement in civil aviation transport will introduce unrecognized and unknown risks.

23 April 2018 Project Kickoff – NESC 18-01304 19

Appendix B: Existing Taxonomies of Positive Behaviors and Resulting Actions in the Civil Aviation Domain

Source 1: Positive Taxonomy (ICAO/CAST, 2013)

- DECISION
  - Avoidance Maneuver
  - Decision to Go-Around
  - Decision to Land as Precaution
  - Decision to Land on an Unexpected Runway
  - Decision to Reject Takeoff
  - Decision to Return to Departing Point or to Divert
- EXTERNAL INTERVENTION
  - Aerodrome Intervention/Assistance
  - Air Traffic Intervention/Assistance
  - Assistance of an Instructor/Supervisor
  - Passenger Intervention/Assistance
  - Third Party Intervention/Assistance
- HARDWARE SAFETY NET
- PROVIDENCE
- SOFT SAFETY NET
  - Accurate Usage of Documentation
  - Communications
  - Design Requirements
  - Engine Failure Anticipation
  - Environment Observation
  - Logical Problem Solving
  - Use of Training Instructions/SOPs
  - Visual Detection/Anticipation

Source 2: "Result" Event Assessment Codes (ASRS, 2018b)

- GENERAL
  - Evacuated
  - Flight Cancelled / Delayed
  - Maintenance Action
  - None Reported / Taken
  - Physical Injury / Incapacitation

- Police / Security Involved
- Release Refused / Aircraft Not Accepted
- Work Refused

- FLIGHT CREW
  - Became Reoriented
  - Diverted
  - Executed Go Around / Missed Approach
  - Exited Penetrated Airspace
  - FLC complied w / Automation / Advisory
  - FLC Overrode Automation
  - Inflight Shutdown
  - Landed As Precaution
  - Landed in Emergency Condition
  - Overcame Equipment Problem
  - Regained Aircraft Control
  - Rejected Takeoff
  - Requested ATC Assistance / Clarification
  - Returned To Clearance
  - Returned To Departure Airport
  - Returned To Gate
  - Took Evasive Action
- AIR TRAFFIC CONTROL
  - Issued Advisory / Alert
  - Issued New Clearance
  - Provided Assistance
  - Separated Traffic
- AIRCRAFT
  - Aircraft Damaged
  - Automation Overrode Flight Crew
  - Equipment Problem Dissipated

Appendix C: Potential Data Sets for Analysis

C.1 Data Sets

Categories	Publicly Available	Temporal resolution	Source for Obtaining	Data Types	Format	Generating Source
Aircraft Flight Recorder	No	1 Hz	Flight Operational Quality Assurance (FOQA)	Continuous, Binary, Categorical	CSV	System
Radar Track Surveillance	With FAA Approval	1 Min	Traffic Flow Management (TFMS) Aircraft Situation Display (ASDI)	Continuous, Categorical	XML	System
		1/4 Hz	SWIM Terminal Data Distribution System (STDDS)	Continuous, Categorical	CSV	System
		1/12 Hz	SWIM Flight Data Publication Service (SFDPS)	Continuous, Categorical	CSV	System
		1 Hz	Airport Surface Detection Equipment Model-X (ASDE-X)	Continuous, Categorical	CSV	System
	No	Daily/Per flight	Trajectory Analysis/ Report Metrics	Continuous, Binary, Categorical	CSV/ XLS	System
Weather	Yes	1 Hr	Meteorological Terminal Aviation Routine (METAR)	Continuous, Categorical	ASCII	System
	Yes	6 Hr	Terminal Aerodrome Forecasts (TAFs)	Continuous, Categorical	ASCII	System
	Yes	1 Hr	Integrated Terminal Weather System (ITWS)	Continuous, Categorical	ASCII/ XML	System
	Yes	1 Hr	Rapid Refresh (RR)	Continuous	Binary (grib)	System
	Yes	2.5 Mins	Corridor Integrated Weather System (CIWS)	Continuous	Binary (hdf5)	System
	Yes	15 Mins	Convective Weather Avoidance Model (CWAM)	Continuous, Binary	Binary (hdf5)	System
	Yes	2 Hrs	Collaborative Convective Forecast Product (CCFP)	Continuous, Binary	ASCII	System
Narratives	Yes	As reported	Aviation Safety Reporting System (ASRS)	Textual	ASCII	Operator
	No	As reported	Aviation Safety Action Program (ASAP)	Textual	ASCII	Operator
	No	Research Studies	Airlines/Controllers Interviews	Textual, Categorical	ASCII	Operator
FAA Procedures/ Notices	Yes	56 Days	SID/STAR Procedures (Coded Instrument Flight Procedures)/Airnav.com	Continuous, Categorical	ASCII/ PDF	System
	Yes	As reported	Notice To Airmen (NOTAM)	Textual, Categorical	ASCII/ XML	System
Planning	Yes	Per flight/as amended	Flight Plans	Textual	ASCII	System
	No	Per flight	Flight Schedules	Textual	ASCII	System
Voice	Yes	Continuous	Live ATC, FAA Voice Archive (last 45 days)	Audio	MP3	System

Traffic Statistics	Yes	15 mins	Aviation System Performance Metrics (ASPM)	Continuous, Binary, Categorical	CSV	System
Training	No	>1Hz	Flight Simulators	Continuous, Binary, Categorical, Video	CSV, MOV	System
	No	Per flight	Flight Instructor (LOFT)	Textual, Categorical	ASCII	Observer
Audits	No	Limited Audit	Line Operations Safety Audits (LOSA)	Textual	ASCII	Observer
Maintenance	No	Daily/ Per flight	Airlines, Original Equipment Manufacturer (OEM)	Textual, Categorical	ASCII	Operator

C.2 Descriptions and Sources for Data Categories

Aircraft Flight Recorder

Description:

- Onboard aircraft flight data recorder. Typically referred to as FOQA data. Continuously records the following categories: aircraft position/orientation (latitude/longitude, altitude, speed, pitch, roll, yaw, accelerations, etc.), control surface positions (ailerons, elevators, flaps, speed brakes, etc.), auto pilot modes (lateral, vertical, auto throttle modes), and engine parameters (N1/N2 rotor speeds, oil pressure/temp, etc.), environmental (temp, winds).

Sources:

- Airlines
- ASIAS

Radar Track Surveillance

Description:

- 4-D Positional information lat/lon/alt/time. Ground speed is derived from positional updates. Facilities monitored include Center, Terminal Radar Approach Control, airport ground operations. Gate-to-gate trajectories spanning multiple facilities can be stitched together post flight.
- Additional metrics are also computed post flight analysis and are available as reports. These include: go-arounds, deviations, turn to final characteristics, en-route weather avoidance.

Sources:

- FAA System Wide Information Management (SWIM) feed, reports radar hits for flights throughout the NAS. These include:

    - Traffic Flow Management (TFMS) which provides Aircraft Situation Display (ASDI) data.
    - SWIM Terminal Data Distribution System (STDDS),
    - SWIM Flight Data Publication Service (SFDPS),
    - Airport Surface Detection Equipment Model-X (ASDE-X).
    - Post flight gate-to-gate stitched trajectories are available through NASA Sherlock data warehouse, Performance Data Analysis and Reporting System (PDARS), Threaded Track.
- Sources for report metrics include: NASA Sherlock, PDARS reports, Threaded Track key performance indicators.

Weather

Description:

- Reports of surface measurements at airports, wind direction, visibility, weather type (fog, snow, rain, thunderstorms, etc.), humidity, temperature, etc.
- Forecasted weather for the next 24-30 Hrs for a 5 NM radius, centered around airports. Forecasts include wind direction, visibility, cloud cover and ceiling, probability of fog, snow, rain, thunderstorms, etc.).
- Current weather information and predictions using graphical and textual formats. Information includes wind shear and microburst predictions, storm cell and lightning information, and terminal area winds aloft. Anticipated weather conditions are provided as 60-minute forecasts. /cite{ https://www.faa.gov/air_traffic/technology/itws/}
- Estimated winds aloft, temperature, and humidity for 3-d grid across US
- Convective weather measurements 2-d grid across US: Vertical Integrated Liquid, Echo Tops, water phase (frozen, liquid, mixed).
- Tactical convective weather cell polygons for en-route operations.
- Forecast strategic large weather cell polygons for en-route operations.

Sources:

- METAR (National Oceanic and Atmospheric Administration (NOAA)).
- Terminal aerodrome forecast (TAF) (NOAA).
- Integrated Terminal Weather System (ITWS). Product of SWIM feed.
- Rapid Refresh (NOAA).
- Corridor Integrated Weather System (CIWS). Product of Massachusetts Institute of Technology (MIT) Lincoln Labs.
- Convective Weather Avoidance Model (CWAM). Product of MIT Lincoln Labs.
- Collaborative Convective Forecast Product (CCFP). Product of MIT Lincoln Labs.

Narratives

Description:

- Voluntarily submitted aviation safety incident/situation narrative reports from pilots, controllers.
- Subject Matter Expert interviews. Directed questions for specific study.

Sources:

- Aviation Safety Reporting System (ASRS). NASA managed.
- Aviation Safety Action Program (ASAP). Airlines, ASIAS.
- Various research interviews. NASA initiated.

FAA Procedures/Notices

Description:

- Procedural information containing navigational sequence of waypoints, altitude, speed requirements for departure/arrival routes.
- Regularly published notices that contains information regarding special use airspace, equipment outages, runway closures, etc.

Sources:

- Coded Flight Instrument Procedures (CIFP). http://airnav.com, Jeppesen.
- Notice to Airmen (NOTAM), over SWIM feed, https://pilotweb.nas.faa.gov.

Planning

Description:

- Filed flight route that describes a sequence of waypoints and procedures used to navigate flight from origin to planned destination. This includes airport, waypoints, SID, STARs, and airways. Amendments as route is adjusted in flight are also recorded.
- Dispatcher flight scheduling may have historical and future planned flights from origin to destination. Tactical information regarding traffic and weather.

Sources:

- Airlines.

Voice

Description:

- Voice communications between ATC and pilots and includes assigned routes, altitude/speed clearances, holding, and traffic advisories. Contains pilot and controller intent with highly domain specific lexicon that is usually succinct.

Sources:

- Live ATC, FAA Voice Archive.

Traffic Statistics

Description:

- Statistics on airport runway configuration, arrivals/departure rates, taxi delay times, ceiling/visibility, temp, winds, runway configuration across airports.

Sources:

- Aviation System Performance Metrics (ASPM)
- Bureau of Transportation Statistics, https://www.bts.gov/topics/airlines-and-airports-0

Training

Description:

- Simulations of flights and/or interactions with ATC in a controlled environment. Studies can control system behavior (weather, traffic, automation, etc.), and measure variable human response.
- Instructor notes after flight training based on experience during training exercise.

Sources:

- Flight simulators Airlines, NASA, Boeing, Airbus.
- Line Oriented Flight Training (LOFT), Airlines, 3rd party training.

Audits

Description:

- Assessment notes of pilot activity during flight. Looking for positive and negative aspects of the operation.

Sources:

- Contracted services (e.g., The LOSA Collaborative)

Maintenance

Description:

- Records on item/systems that were repaired/replaced. Scheduled vs unscheduled maintenance timelines.

Sources:

- Airlines, OEMs.

Appendix D: Pilot and Controller Interview Protocol

Initial Question: Unplanned and unexpected events happen routinely during operations in the NAS. We are interested in how [pilots/controllers] make adjustments before, during and after these unplanned or unexpected events in order to maintain safe operations. Can you tell me about a specific unplanned or unexpected event that you have experienced in the course of routine operations?

Probe 1 (Anticipate):

- Were there things you were aware of at the start of your [flight/shift] that you thought increased the likelihood that this event might occur during that [flight/shift]?
- How did you know that this event might occur?
- How else might you have been able to anticipate that this event would occur?

Probe 2 (Monitor):

- Were there things that you experienced during that [flight/shift] that you thought increased the likelihood that this event might occur?
- What signaled/indicated to you that this event was about to occur, was occurring, or had occurred?
- How did you know what indicators of this event to look for during your [flight/shift]?
- What other indicators could have alerted you to this event?

Probe 3 (Respond):

- How did you respond to this event?
- How did you know what to do in response to this event?
- If you had not already known what to do to respond to this event, how would you have figured out what to do?

Probe 4 (Learn):

- What did you learn from this event?
- How did what you learned impact the remainder of your [flight/shift] or that operation?
- How did what you learned impact how you prepare for future [flights/shifts] or operations?
- Have you shared what you learned with others in your organization? How did you do that?
- In general, what practices are in place in your organization for [pilots/controllers] to share lessons learned?

Probe 5 (Wrap-up):

- Is there anything further you'd like for us to know about this event that we haven't already discussed?

Appendix E: Controller Questionnaire

NASA is investigating development of continuous real-time monitoring, real-time anomaly and precursor identification tools to help identify, predict, and help prevent emergent risks and hazards to the air traffic system. We are interested in understanding the contribution that human operators, such as air traffic controllers, have on the safety of the air traffic system.

Traditional approaches to safety management focus on preventing the things that can go wrong, using techniques such as accident and incident analysis. Another approach to safety management is emergent risk mitigation by analyzing the things that go right. This NASA project aims to find and document "safety successes" and how people anticipate, monitor for, response to, and learn from unexpected events to exhibit successful, resilient performance.

Resilience can be defined as "the intrinsic ability to adjust functioning prior to, during, or following changes and disturbances, so that required operations can be sustained under both expected and unexpected conditions."

1. How Often Do You Exhibit Resilience to Perform the Job as Air Traffic Controller?

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

2. In your job as Air Traffic Controller, how often do you exhibit the following behaviors?

a) Finding goal-directed and proactive solutions that require trading for conflicting goals such as capacity, efficiency, and costs

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

b) Anticipating needs for planning and coordination

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

c) Using judgment for improvisation of standard operating procedures for safety/efficiency/ capacity purposes

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

d) Inventing work-around procedures and techniques that work better for actual practice

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

e) Applying flexibility to increase safety buffers and defensive controlling for buffering capacity, margins, and added safety tolerance

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

f) Providing team support and adaptive capacity as required

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

g) Utilizing resources as required, such as consulting written/printed documentation (manuals, procedures) or electronic information (e.g., AIP online, route charts for alternative waypoints)

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

h) Developing strategies for managing workload

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

3. How often would you estimate the traffic management decisions you make are not specified as policies and procedures by JO 7110.65 or that of the specific facility (for example, procedural letters of agreement).

Less than Once per Month	At Least Once Per Month	At Least Once Per Every 2 Weeks	At Least Once per Week	At Least Once Daily	At Least Once Per Session	More Than Once Per Session

How many of these would you categorize as "resilient" decisions? ______________ % of decisions

| ---------- | ---------- | ---------- | ---------- | ---------- | ---------- |

0% Less Than 10% 50% More Than 90% 100%

4. **Imagine a new Air Traffic Management system was introduced at the facility where you work/have worked. Assume that you were told that the system was designed to exactly follow the policies and procedures defined by JO 7110.65 and that of the specific facility (for example, procedural letters of agreement). You were also told that you continue to have responsibility but that your job is to now to monitor the system and correct and/or over-ride any decisions made by the system. You may not turn the system off.**

   Please estimate the % of air traffic control decisions the system would make **correctly**:

   a. During peak traffic volume % of decisions made correctly _______
   b. During nominal traffic volume % of decisions made correctly _______
   c. During low traffic volume % of decisions made correctly _______

   Please estimate the % of air traffic control decisions the system would make **safely**:

   a. During peak traffic volume % of decisions made correctly _______
   b. During nominal traffic volume % of decisions made correctly _______
   c. During low traffic volume % of decisions made correctly _______

5. **The Aviation Safety Reporting System (ASRS) provides the following database entry fields for Air Traffic Control:**

   Provided Assistance
   Issued Advisory/Alert
   Issued New Clearance
   Separated Traffic

   We are interested in what additional fields should be added/included in the ASRS system that would characterize positive actions and/or outcomes performed by or as a result of Air Traffic Control. Please provide additional result-oriented action phrases that may describe an action or event that would represent a positive outcome.

6. **What do you consider the most significant emergent risk to resilience today for Air Traffic Management? Why?**

7. **In 20 years, how do you think ATM will be different than it is today?**

8. **What do you think will be the most significant risk to resilience for Air Traffic Management? Why?**

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
11/06/2019

2. REPORT TYPE
Technical Memorandum

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Human Performance Contributions to Safety in Commercial Aviation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Holbrook, Jon B.; Stewart, Michael J.; Smith, Brian E.; Prinzel, Lawrence J.; Matthews, Bryan L.; Avrekh, Ilya; Cordoza, Colleen T.; Ammann, Oliver C.; Adduru, Viraj; Null Cynthia H.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
869021.05.07.08.05

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER
L-21075

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSOR/MONITOR'S ACRONYM(S)
NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NASA/TM-2019-220417

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 1 Aeronautics (General)
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In commercial aviation, large volumes of data are collected and analyzed on the failures and errors that result in infrequent incidents and accidents, but in the absence of data on behaviors that contribute to routine successful outcomes, safety management and system design decisions are based on a small sample of non-representative safety data. Failure to consider human contributions to successful system performance represents a significant and largely unrecognized risk when making policy decisions about human roles and responsibilities. The principal focus of this assessment was to identify current gaps and explore methods for identifying human contributions to safety through analysis of aviation data.

15. SUBJECT TERMS
NASA Engineering and Safety Center; Commercial Aviation; Human Success; Air Traffic Control

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					STI Help Desk (email: help@sti.nasa.gov)
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	64	(443) 757-5802