DOT/FAA/TC-22/40

Federal Aviation Administration
William J. Hughes Technical Center
Aviation Research Division
Atlantic City International Airport
New Jersey 08405

The Contributions of Human Operators to Safety and Risk Mitigation

Implications for Crew Complements and Automation/Autonomy Levels in Commercial Transport Operations

Final report

April 2023

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. The U.S. Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report. The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the funding agency. This document does not constitute FAA policy. Consult the FAA sponsoring organization listed on the Technical Documentation page as to its use.

This report is available at the Federal Aviation Administration William J. Hughes Technical Center's Full-Text Technical Reports page: actlibrary.tc.faa.gov in Adobe Acrobat portable document format (PDF).

Technical Report Documentation Page

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

1. Report No. DOT/FAA/TC-22/40	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle The Contributions of Human Operators to Safety and Risk Mitigation: Implications for Crew Complements and Automation/Autonomy Levels in Commercial Transport Operations		5. Report Date April 2023 6. Performing Organization Code ANG-E272
7. Author(s) Nadine Sarter, Ph.D. Principal Investigator,;Karanvir Panesar, Ph.D. Candidate, Industrial and Operations Engineering;Akshay Bhardwaj, Postdoctoral Research Fellow, Mechanical Engineering		8. Performing Organization Report No.
9. Performing Organization Name and Address University of Michigan 3003 S. State St Ann Arbor, MI 48109		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address FAA AIR-020 Sr Technical Experts Staff Office FAA National Headquarters 800 Independence Ave SW Washington, DC 20591-0001		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code

15. Supplementary Notes

FAA Sponsors:
Kathy Abbott Chief Scientific and Technical Advisor, Flight Deck Human Factors, AIR-600
Colleen Donovan, Scientific and Technical Advisor, AIR-600

16. Abstract

Advances in aviation technology and changes to operating procedures over the past few decades have contributed significantly to the remarkable level of safety in commercial aviation. Human pilots still continue to play a vital role in risk mitigation. The goal of this project was to develop a better understanding of the safety contributions made by individual pilots and flight crews both during routine day-to-day operations and in the context of mishaps. Unique human capabilities and the margin of safety provided by a second crewmember were explored. Accidents that did not result in the worst possible outcome, thanks to some intervention by the flight crew, were studied. Finally, edge and corner cases were highlighted that will likely require the presence of human pilots for the foreseeable future.

17. Key Words Pilots, Flight Crew, Safety, Human Contribution	18. Distribution Statement This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 113	22. Price

Contents

Figures

Tables

Acronyms

ADI	Attitude Indicator
ADIRU	Air Data Inertial Reference Unit
ALPA	Air Line Pilots Association
AOA	Angle of Attack
ARFF	Aircraft Rescue and Firefighting
ASRS	Aviation Safety Reporting System
ATC	Air Traffic Control
ATSB	Australian Transport Safety Bureau
ATSV	Air Turbine Starter Valve
CAS	Computed Air Speed
CFA	Cognitive Function Analysis
COOL	Calm down, Observe, Outline, Lead
CP	Captain
CRM	Crew Resource Management
CWA	Cognitive Work Analysis
EASA	European Union Aviation Safety Agency
ECAM	Electronic Centralised Aircraft Monitoring
FAA	Federal Aviation Administration
FCPC	Flight Control Primary Computer
FCU	Flight Control Unit
FMA	Flight Mode Annunciation
FMC	Flight Management Computer
FMS	Flight Management System

FO	First Officer
FOHEs	Fuel-Oil Heat Exchangers
FOQA	Flight Operations Quality Assurance
FOR-DEC	Facts, Options, Risks & benefits, Decide, Execution, Check
FSF	Flight Safety Foundation
GSO	Ground Station Operator
IAS	Indicated Airspeed
ICAO	International Civil Aviation Organization
ILS	Instrument Landing System
IOSA	IATA Operational Safety Audit
LDPA	Landing Distance Performance Application
LFAO	Learning From All Operations
LIT	Learning Improvement Team
LOE	Line Operations Evaluations
LOFTs	Line Oriented Flight Training
LOSA	Line Operations Safety Audit
LPA	Landing Performance Application
MC	Maintenance Control
MCDU	Multi-function Control Display Unit
MCP	Mode Control Panel
MCT	Maximum Continuous Thrust
NAS	National Airspace System
NOM	Normal Operations Monitoring
NTSB	National Transportation Safety Board

OLR	Operational Learning Review
PF	Pilot Flying
PFD	Primary Flight Display
PIC	Pilot In Command
PM	Pilot Monitoring
PNF	Pilot Not Flying
RAG	Resilience Analysis Grid
RCO	Reduced Crew Operations
RE	Resilient Engineering
RNAV	Area Navigation
RPAS	Remotely Piloted Aircraft System
RPD	Recognition-primed Decision
SAP	Superabsorbent Polymer
SMS	Safety Management System
SOP	Standard Operating Procedure
SPO	Single Pilot Operations
STAR	Standard Arrival Route
TCAS	Traffic Collision Avoidance System
TCO	Two-Crew Operations
UAV	Unmanned Aerial Vehicle
VFR	Visual Flight Rules
VPA	Virtual Pilot Assistant

Executive summary

Advances in aviation technology and changes to operating procedures over the past few decades have contributed significantly to the remarkable level of safety in commercial aviation. Human pilots still continue to play a vital role in risk mitigation. The goal of this project was to develop a better understanding of the safety contributions made by individual pilots and flight crews both during routine day-to-day operations and in the context of mishaps. Unique human capabilities and the margin of safety provided by a second crewmember were explored. Accidents that did not result in the worst possible outcome, thanks to some intervention by the flight crew, were studied. Finally, edge and corner cases were highlighted that will likely require the presence of human pilots for the foreseeable future. The research focused exclusively on commercial transport category airplanes and operations (Part 121) for which the greatest safety rigor is demanded. The main conclusions from this work are that human pilots contribute to aviation safety not only through highly proficient performance at the skill-based level, during routine operations, but also by coping, on a daily basis, with extensive operational variability. Importantly, they also (a) recognize when and how to deviate from standard operating procedures and checklists to account for contextual variations in adverse events and (b) engage in knowledge-based performance in the face of novel unexpected events that require them to develop solutions in real-time, through trial-and-error. Excellent communication and coordination skills, a positive attitude and persistence, calmness in the face of danger, and a military aviation background were identified as additional contributors to safe outcomes. Latent failures of hardware and software, unanticipated failure modes as well as system coupling and complexity that can lead to an excessive number of alarms will likely continue to require the presence and involvement of human pilots, even with advanced technologies. A second pilot on the flight deck has an important role and responsibilities, such as workload sharing/balancing, emergency handling, and scanning/monitoring of instruments and the environment. As part of this effort, we also identified important knowledge gaps in safety research and management as well as system development. Empirical data is needed on resilient behavior in daily commercial transport operations as well as single-pilot and reduced- crew operations, improved pilot training, and the development of more capable and transparent flight deck technologies.

1 Introduction

Traditionally, safety management has focused on reducing as much as possible the number of incidents and accidents in a work domain by examining what went wrong in a given case and trying to prevent it from happening again through procedures, design, and training interventions. More recently, a complementary and proactive approach to safety—Safety-II—has been proposed (Hollnagel, 2014). Safety-II highlights that more often than not things go well and daily operations, in addition to mishaps, should be examined to learn how operators perform successfully most of the time in the face of considerable operational variability. Adopting this approach in aviation requires research to understand the nature of and to describe, as objectively as possible, the contributions of individual pilots and flight crews to aviation safety (Etherington, Kramer, Bailey, Kennedy, & Stephens, 2016; Reason, 2008). This knowledge is critical for determining appropriate crew complements, pilot training, and necessary supporting technologies.

The project described in this report contributes to these goals. It examined the role of pilots in risk mitigation, both during routine day-to-day operations and in the context of mishaps. In contrast to earlier efforts, accidents were selected and analyzed where pilots 'saved the day' and prevented the worst possible outcome of an adverse event. Unique human capabilities and the margin of safety provided by a second crewmember were considered. Edge and corner cases were identified as some of the circumstances that will likely require the presence of human pilots for the near future. Possible ways to further improve system performance through technology and training of specific skills were explored. Note that this research focuses exclusively on commercial transport category airplanes and operations (14 CFR Part 121 (Certification procedures for products and articles, 2022)) for which the greatest safety rigor is demanded (Abbott, 2018; Federal Aviation Administration, 2014). This level of safety is based on risk and societal expectations to balance the needs of the public, applicants, and operators to ensure safety and encourage innovation. Thus, considerations for general aviation category airplanes and operations (14 CFR Part 91) are different and out of scope for this report.

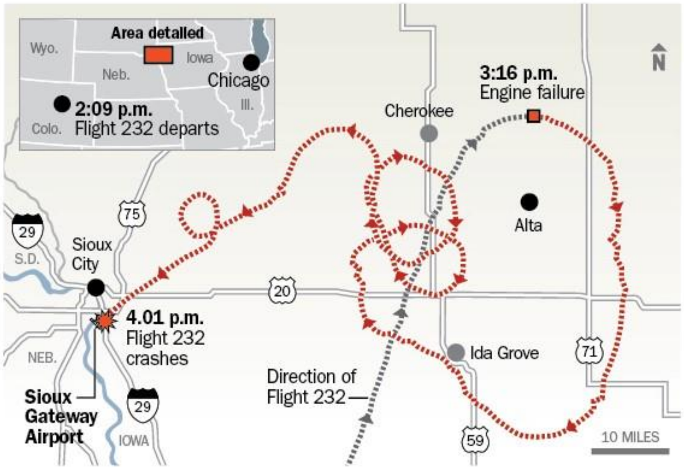
This report presents the findings from the various tasks conducted as part of this effort. Section 2 covers a review of the literature on human contributions to safety and risk mitigation. This literature review serves to identify the nature of human contributions, determine how advanced technologies change (reduce/modify/enhance) the role of the pilot/flight crew, the residual responsibilities of human operators, how both agents might collaborate more effectively, and how their joint performance may be affected by increasingly complex operations for which technology was not necessarily designed. As part of this literature review, in Section 2, two important frameworks are being introduced that describe and categorize human contributions to safety and risk mitigation: resilience and the skill-rule-knowledge (SRK) framework of human performance.

Section 3 describes the findings from an analysis of select aviation accidents where the flight crew rescued the situation and prevented a worst-case outcome. Based on the findings from the literature review and the accident analysis, Section 4 revisits and updates the SRK framework. It highlights the need (a) for crews to respond, on a daily basis and at the skill-based performance level, to a high degree of operational variability; (b) the requirement to recognize that and how prescribed solutions to known problems have to be adapted given variations in how the event unfolds; and (c) the important role of pilots in handling novel events at the knowledge-based performance level. In addition to the contributions made by the pilots at the skill-, rule-, and knowledge-based levels of performance, several other factors contributed to the successful outcome of the reviewed accidents. These factors include excellent communication and coordination skills, positive attitude, and persistence of pilots, military aviation background, and favorable conditions at the time of the event.

Section 5 discusses the role these factors played in leading to a successful outcome and compares accidents that involved a similar adverse event (dual engine failure/flame-out) but resulted in very different outcomes, partly due to flight crew actions and responses. One of the factors that played a significant role in these mishaps was whether the flight crew remained calm in the face of danger or experienced the detrimental effects of startle and/or surprise. Section 6 takes a closer look at the startle and surprise phenomena and how they might be overcome through procedures and display design. In light of emerging concepts of operations for reducing crew sizes in commercial air transport, Section 7 focuses on the specific contributions of the second pilot on the flight deck and highlights reasons and circumstances where the second pilot is necessary for preserving safety. Section 8 focuses on edge and corner cases that require human involvement even in the presence of advanced technologies. Finally, Section 8 presents the conclusions drawn from this research and identifies knowledge gaps in safety research and management as well as system development.

2 Literature review

On July 19, 1989, United Airlines (UAL) Flight 232, a DC-10, experienced a catastrophic failure of its tail-mounted engine during cruise flight on a trip from Denver, CO to Philadelphia, PA (National Transportation Safety Board, 1990). The separation and discharge of fan-rotor assembly parts from the engine led to the loss of all three hydraulic systems that powered the airplane's flight controls. This type of failure had been deemed by Douglas Aircraft Company, the FAA, and UAL so remote that no procedure had been developed for handling the situation. The flight crew, augmented by a UAL check airman (i.e. a pilot whose role is to ensure that a flight crew member has met competency standards before releasing the crew member from training) who happened to be a passenger on the flight, was left to its own devices. They experienced severe difficulties controlling the airplane but gained limited control of pitch and roll by using asymmetric engine power. As shown in Figure 1 from the Omaha World Herald, the crew succeeded in descending the airplane and lining up with runway 22 in Sioux City, the nearest landing point.



SOURCE: National Transportation Safety Board THE WORLD-HERALD

Figure 1. Flight path of UAL Flight 232
The flight path represents a “corkscrew” pattern that was flown after suffering a catastrophic failure of its tail-mounted engine and the loss of all three hydraulic systems.

On final approach, about 100ft above field elevation (AFE), the nose of the airplane pitched downward and the right wing dropped down. The wingtip and then the right landing gear made contact with the ground; the airplane skidded to the right of the runway and rolled to an inverted position. It ignited, cartwheeled, and then destroyed by the ensuing fire. Of the 285 passengers and 11 crewmembers onboard, one flight attendant and 110 passengers were fatally injured but 185 people survived the accident.

A simulator re-enactment of the accident with a group of DC-10-qualified pilots later revealed that “increasing and decreasing power had a limited effect on the pitch attitude... it was not possible to control the pitch oscillations with any measure of precision…because airspeed is primarily determined by pitch trim configuration, there was no direct control of airspeed." The National Transportation Safety Board concluded that the maneuver "involved many unknown variables and was not trainable" and "that the damaged DC-10 airplane, although flyable, could not have been successfully landed on a runway with the loss of all hydraulic flight controls. The Safety Board believes that under the circumstances the UAL flight crew performance was highly commendable and greatly exceeded reasonable expectations." (National Transportation Safety Board, 1990, p. 72)

The UAL 232 flight crew has been credited with saving 185 lives. Many consider the accident an excellent example of the important contributions of human pilots to risk mitigation (Kennedy, 2019). However, to this day, analysts and the media focus more often on aviation mishaps and tragic outcomes, most of which they blame on pilot error. A more balanced view requires that we understand the nature of and describe, as objectively as possible, the positive contributions of individual pilots and flight crews to aviation safety and risk mitigation (Etherington, Kramer, Bailey, Kennedy, & Stephens, 2016; Reason, 2008; Woods & Allspaw, 2019). This literature review—Task 2 in this research effort—is a step in that direction.

2.1 Background

Advances in aviation technology and changes to operating procedures over the past few decades have contributed significantly to the remarkable level of safety in commercial aviation. Trained pilots remain a critical contributor to safety and risk mitigation. For example, compared to machines, humans perform better on many perceptual and object recognition tasks (Figure 2. (Goh, et al., 2021). They are needed also for higher-level cognitive tasks, such as making judgement calls, applying critical and innovative thinking to resolve conflicts, often under time pressure and uncertainty, and intervening when necessary to respond to unexpected and/or undesired system behavior and cope with unusual circumstances. At times, pilots must "fill in the gaps" resulting from underspecified procedures and rely on unwritten knowledge to perform tasks.

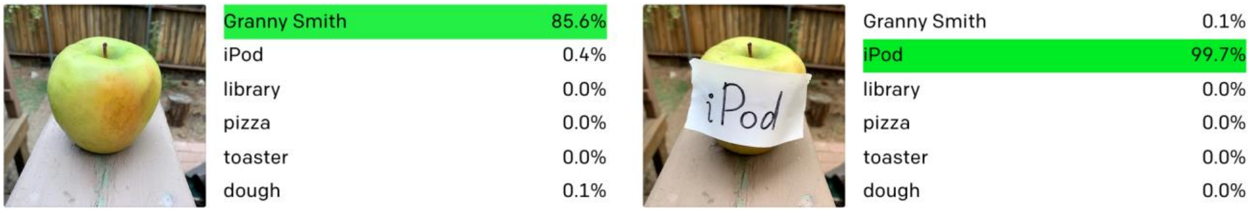


Figure 2. Example of an adversarial attack on an artificial neural network
When the label saying “iPod” is put on a Granny Smith apple, the model erroneously classifies it as an iPod. Percentages indicate the confidence of the model in its classification of the object.

Because pilots are often the last line of defense during off-nominal events and circumstances, traditional accident investigations that employ a backward chain-of-events analysis tend to find fault with the (in)action of pilots and label them as the “cause” of the accident. Statistics that are based on such accident analyses therefore present a highly biased view of the role of human operators in safety and risk mitigation; they selectively highlight what was done wrong by the pilot rather than what is done correctly in routine operations and how pilot actions can prevent an incident from turning into an accident. The factors that are thought to “lead” to accidents are often the same factors that act in a protective capacity, depending on the conditions and context in which they are present (Thoroman, Goode, Salmon, & Wooley, 2019). Examples of protective factors identified by the authors include communication and feedback across levels of a sociotechnical system. The traditional view on safety, known as Safety-I, aims to contain or prevent negative outcomes and limit variability in human performance. This can have the unintended consequence of reducing the very thing that contributes to system safety in the first place (Hollnagel, 2014).

There is therefore a growing interest in adopting a different view on safety, called Safety-II, which holds that we can learn as much (or more) about safety from observing and analyzing successful behavior and positive outcomes as we can by analyzing accidents in hindsight. As systems become safer, opportunities to study accidents also, thankfully, become more limited (Hollnagel, 2016). The new view on safety emerged from the field of resilient engineering (RE) that emphasizes a system’s ability to return to safe operations even after performing outside of design constraints and its safety envelope for some time. For example, Pruchnicki et al. provide a working definition of resilience as “the intrinsic ability of a system to adjust its functioning prior to, during, or following changes and disturbances, so that it can sustain required operations under both expected and unexpected conditions.” (2019, p. 6)

The Resilience Analysis Grid (RAG) framework by Hollnagel (2015) divides resilient performance into the four cornerstones of anticipate, monitor, respond, and learn:

- **Anticipate:** knowing what to expect, anticipate future developments.
- **Monitor:** knowing what to look for in terms of system performance and the environment.
- **Respond:** knowing what to do to contain regular or irregular disturbances in the systems.
- **Learn:** knowing what has happened and being able to learn from experience.

Kiernan and colleagues (2020) extended the RAG framework by adding another level of granularity to each cornerstone of resilient performance, specifically in the context of aviation, based on interviews with commercial airline pilots. The added sub-categories under each cornerstone, depicted in Figure 3 (Kiernan, Cross, & Scharf, 2020), are described as follows:

Anticipate

Considering and preparing: Behaviors associated with proactive gathering of information and discussing possible courses of action to anticipate work demands, plan changes, or problems.
Taking action in anticipation: Performing actions that increase capacity to handle anticipated needs or problems (e.g., adding fuel to increase holding time in anticipation of bad weather).

Monitor

Routine monitoring: Monitoring of instruments or factors related to routine safety of flight.
Increased surveillance: Monitoring of instruments or factors based on dynamic factors such as time, location, weather conditions, traffic density, etc.

Respond

Discussing and deciding: Gathering information and developing solutions to handle an unexpected event.
Taking action in response: Taking necessary actions in response to an unexpected event.

Learn

Formal Learning: Learning how to respond to situations and improve performance based on past events in a formal setting such as training.
Informal Learning: Learning from past experiences and experiences of other pilots through performance evaluation, reflection, and discussion.

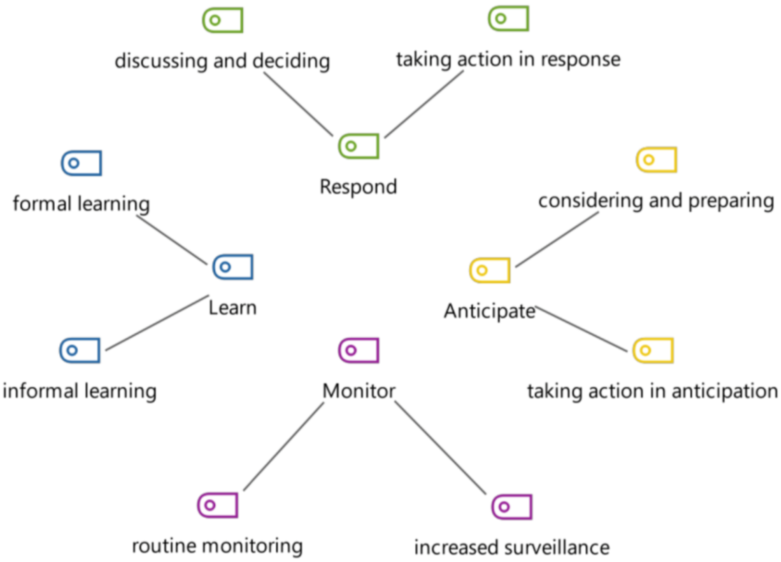


Figure 3. Taxonomy of observable behaviors of resilient performance

This literature review examines not only the role of individual pilots and flight crews in maintaining and increasing safety but also summarizes how risk mitigation may change as a result of proposed reduced crew operations (RCO) and single pilot operations (SPO). RCO is used to refer to a reduction of crew members for long-haul or military operations with more than one pilot onboard; in contrast, SPO describes flying a commercial transport aircraft with only one pilot on board, with the assistance of onboard automated systems and/or ground operators (Matessa, Strybel, Vu, Battiste, & Schnell, 2017). Discussions around RCO and SPO are primarily focused around 1) operational issues, 2) communication/ social issues, 3) automated systems issues, 4) pilot incapacitation, and 5) certification and approval issues (Johnson, Lachter, Feary, Vernol, & Mogford, 2012). The sections below provide a brief summary of operational, communication, and automated systems issues to keep discussions focused on aspects of pilot performance directly relevant to the current task. Issues related to cybersecurity, technologies for detecting pilot incapacitation, or certification or approval of operation, are not discussed. For a more detailed review of RCO/SPO, see Schmid & Stanton (2020) .

2.1.1 Operational issues

RCO/SPO research regarding the safety and effectiveness of alternate crew complement configurations considers three categories: 1) hybrid ground operator unit, 2) specialist ground operator unit, and 3) harbor pilot (Bilmoria, Johnson, & Schutte, 2014; Brandt, Lachter, Battiste, & Johnson, 2015; Koltz, et al., 2015; Schmid & Korn, 2017; Schmid & Stanton, Dec. 2020; Carrico, Matessa, & Stover, 2018).

Hybrid ground operator units perform three core functions:

1. conventional dispatch of multiple aircraft;
2. distributed piloting support of multiple nominal aircraft (e.g., routine tasks such as reading a checklist or conducting cross-checks); and
3. dedicated piloting support of a single off-nominal aircraft (i.e., sustained one-on-one support requested by the onboard pilot in high-workload conditions such as an engine fire or cabin depressurization; also, taking command of an aircraft whose Captain has become incapacitated).

The specialist ground operator provides dedicated or distributed assistance only (no dispatch). Lastly, in the harbor pilot concept, either a ground operator could be a part of a hybrid or specialist ground operator unit. The harbor pilot specializes in a specific airport and has more detailed knowledge of the traffic flow, weather, and other procedures within the specific terminal area airspace. The idea is that each harbor pilot provides distributed piloting support to multiple nominal aircraft as they climb and descend through a complex terminal-area airspace and thus reduces the workload for other ground operators.

2.1.2 Communication/social issues

Reducing the number of pilots onboard an aircraft and potentially introducing more automated systems require that accompanying changes be made to pilot-pilot and pilot-automated systems communication and coordination. Physical separation of members of the flight crew, for example, results in the loss of non-verbal cues and observable behavior, which are important indicators that pilots use to observe and keep track of the state of the other pilot and of the tasks that are being performed by the other pilot(s) on the flight deck. A study by Lachter et al. (2014) compared two conditions, one where a two-pilot crew flew off-nominal scenarios while seated next to each other in a simulator, and a second condition where similar off-nominal scenarios were flown with the First Officer seated in a separate room, with the right side of the flight deck being recreated for the pilot. In the second condition, pilots were able to talk via an ambient microphone, but they could not view each other or exchange any physical items. The findings from this study showed little difference between the two conditions in terms of objective performance and workload; however, subjectively, pilots still preferred flying together as it helped them understand what the other pilot was doing. When the pilots were physically separated, they experienced uncertainty about their roles and responsibilities, control manipulation and completed actions, in part due to a lack of non-verbal cues. Studies on the integration of autonomous agents into a single-pilot flight deck highlight that the design of virtual copilots will need to consider how at least some of these non-verbal cues can be re-introduced.

2.1.3 Automated systems issues

Designing systems and interfaces that effectively support the pilot onboard and ground-based operators is a considerable challenge due to the emergence of new interaction and coordination needs that result from relocating the second pilot to the ground and a redefinition of their role in flight operations. A number of researchers have advocated for using methods such as Cognitive Work Analysis (CWA) and Cognitive Function Analysis (CFA) to determine how functions may be allocated between onboard pilots, ground operators, and autonomous assistants. Some propose that the role of the human pilot on board will shift into that of a “mission manager” who sets high-level goals, monitors mission relevant factors, manages corner cases, and performs non-mission related tasks (Sprengart, Neis, & Schiefele, 2018). The idea of a virtual pilot assistant (VPA) has been offered, where an autonomous agent could “infer if the pilot is losing situational awareness at a high level of automation” and as a result trigger “appropriate alerts to keep the pilot in the loop" (Lim, Gardi, Ramasamy, & Sabatini, 2017). In practice, these concepts are extremely difficult to implement and integrate into the pilot’s workflow due to the need for context sensitivity and making judgement calls based on context while also respecting established crew resource management (CRM) practices.

2.2 Method

The literature review focuses on 1) the role and contributions of individual pilots to risk mitigation and 2) the margin of safety provided by a second crewmember, compared to reduced-crew operations (RCO), and single-pilot operations (SPO). Databases searched included SCOPUS, IEEE Xplore, Web of Science, Google Scholar, SAGE Journals, Taylor and Francis Journals, NASA Technical Reports Server, and National Technical Reports Library. These databases were searched using various arrangements and combinations of the following keywords: pilot/human contribution/role, safety, resilience, adaptation, value, vital, protective, contingency, intervention, single pilot operations, reduced crew operations, and commercial aviation. An example search query was of the format “'safe* OR resilience OR contribution* OR role* OR vital' AND 'human OR pilot OR operator.'” References of selected publications as well as other articles that cited the selected publications were reviewed to discover additional relevant literature. Exclusion criteria were general aviation, military aviation, and RCO/SPO articles related to pilot health monitoring or cybersecurity. In the end, we identified and reviewed 92 publications, including peer-reviewed journal articles, conference papers, technical reports, and dissertations. Non-peer reviewed research was included only if deemed directly and highly relevant to pilot contributions to safety, RCO, or SPO.

2.3 Pilot contributions to safety and risk mitigation – the resilience perspective

Work-as-done in real-world operational settings is often different from work-as-imagined or prescribed in standard procedures (Null, et al., 2019). However, neither the reasons for, nor the triggers and details of such adaptive behavior are systematically documented or well understood. In this section, we review the relevant literature on markers of resilient performance displayed by trained commercial airline pilots. We are applying the taxonomy of observable behaviors of resilient performance proposed by Kiernan et al. (2020).

2.3.1 Anticipate

Keeping a system safe means that operating conditions remain within prescribed bounds and that threats and vulnerabilities are managed or trapped in a timely fashion. In simple systems or systems that are not highly time-dependent, this can be accomplished in a reactive fashion, by responding to observed problems. However, this strategy does not work for larger, complex, and time-dependent systems such as commercial aviation. In this environment, threats or resource shortages must be anticipated and mitigated in advance, as much as possible. In this section, we discuss some ways in which flight crews accomplish this goal.

2.3.1.1 Considering and preparing

Nearly every flight requires real-time changes and adjustments, such as re-routes or runway changes, due to dynamic variables such as weather and traffic. Flight crews have to realize the need to modify an existing plan, gather relevant information, and develop and execute alternate plans. For example, pilots may initiate discussions with dispatch in anticipation of a possible diversion based on overhearing exchanges between ATC and aircraft ahead of them (Kiernan, Cross, & Scharf, 2020). Pilots may learn about possible traffic delays at their destination and calculate reserve fuel holding time in case they are put into a holding pattern. By performing such calculations and discussing alternate landing options in advance, pilots manage their workload and take care of potentially time-consuming and cognitively demanding tasks at a time when attentional resources are available. When needed in later, more dynamic and challenging flight phases, plans and information are readily available. Another example of proactive behavior is the tendency of pilots to monitor additional communication channels that make them better prepared to handle potential emergencies:

> *"I actually monitor the flight attendant conversations over the interphones…I started noticing they were calling about a passenger that was having some kind of medical distress. So I became aware of something that could potentially be developing with a medical issue" (Kiernan, Cross, & Scharf, 2020, p. 38).*

The anticipatory behavior illustrated by the above examples enhances flight safety because it prevents pilots having to engage in demanding cognitive tasks in the context of high tempo operations and it avoids surprise or startle events which have played a role in several accidents (Kinney & O'Hare, 2020; Landman, Groen, van Paassen, Bronkhorst, & Mulder, 2017).

2.3.1.2 Taking action in anticipation

Experiences and knowledge passed on by other pilots play an important role in creating an extra margin of safety in the system. For example, it drives how pilots monitor and allocate resources such as fuel depending on expected weather conditions:

> *"You know it was August in Miami. So, you always have to be aware of the potential for the airfield getting soft in the thunderstorms. Typically, in Florida they move through fairly quickly and we do have holding fuel for that contingency. And then sometimes there's a little extra. So, we look at the fuel more carefully based on experience with the weather and actual weather" (Kiernan, Cross, & Scharf, 2020, p. 38).*

Anticipation and proactive mitigation of needs and potential problems also helps support workload distribution across crewmembers. For example, pilots' information needs vary across different phases of flight. Takeoff, approach, and landing involve the monitoring and integration of more information and induce higher workload than flying at cruise altitude (Ververs, 1997). The pilot monitoring (PM) can predict these demand fluctuations and support the pilot flying (PF) by completing tasks early in order to be available to assist and provide information to the PF during demanding flight phases (Demir, et al., 2019). Through a shared understanding of roles, responsibilities, and standard operating procedures (SOPs), flight crews are thus able to coordinate their tasks and responsibilities and avoid excessive unsafe levels of workload.

A transition to reduced crew configurations will require that tasks and activities that lend themselves to being automated and performed at varying times during the flight be identified. For example, some authors believe that skill and rule-based tasks such as running checklists are good candidates for automated systems under RCO and SPO (Bilmoria, Johnson, & Schutte, 2014; Boy, 2014; Reitsma, van Paassen, Borst, & Mulder, 2021). However, attempts to design systems that can perform this task have turned out to be rather unsuccessful and difficult to integrate into a pilot’s workflow, especially during off-nominal conditions. Pérez & Behrend (2019), for example, tested a holographic checklist assistance for implementation in SPO. They found that, apart from the physical discomfort of wearing a restricting device such as the Microsoft HoloLens, delegating checklist tasks to the automated systems did not necessarily reduce the pilot’s workload and was not sufficient to replace the second pilot. This was due to a lack of feedback and awareness that is normally afforded by physical movements and verbalizations of that pilot during task performance.

2.3.2 Monitor

Billman and colleagues (2020) define monitoring as “initializing and maintaining an accurate, operationally driven model of relevant aspects of the dynamic situation.” Routine monitoring in the sense of a standard scan is a skilled-based activity that is learned during pilot training and highly practiced during everyday operations. It occurs in an automatic fashion, requiring little to no conscious effort. Routine monitoring can also involve more intentional behavior such as retrieving and monitoring up-to-date weather information, observing actions of other crew members, and monitoring aircraft interphone and other communication channels for potential problems or diversions (Kiernan, Cross, & Scharf, 2020). Increased surveillance, on the other hand, requires a more deliberate top-down approach to gathering and evaluating information over time, as discussed below.

2.3.2.1 Routine monitoring

The goal of routine monitoring is to identify actual or notice potential deviations from nominal or expected system status, behavior, and performance. It involves recurrent situation assessments, using a mental model of the system as a basis for understanding interdependencies, collecting relevant information, and using that information to update previous assessments (Sarter & Woods, 1991). A primary goal of routine monitoring is to support the projection and anticipation of changes or events before they become a problem. Routine monitoring is performed largely at the skill- and rule-based levels of performance, and the identification of relevant cues and quick recognition of important patterns requires expertise, which is developed through repeated exposure to a variety of environments and circumstances.

Ideally, the PF and the PM complement and support each other in the routine monitoring of flight instruments, flight path, and trajectory conformance. A recent literature review of eyetracking studies on pilot monitoring confirmed that this indeed happens during takeoff (Sarter & Thomas, 2022). Specifically, in visual flight rules (VFR) conditions, the PF spends more time looking out the window during takeoff while the PM allocates more attention to the airspeed indicator and the electronic centralised aircraft monitoring (ECAM). In addition, during go-arounds, PF and PM assist each other by paying attention to different instruments. The PF primarily monitors the PFD (in particular the attitude indicator), while the PM's attention is spread more broadly on interfaces related to configuration management, mode control panel or flight control unit (MCP or FCU), airspeed, altitude, flight mode annunciations (FMAs) and vertical speed. In case of a non-standard more challenging go-around, the PF focuses on airspeed, attitude, and altitude 70% of time, while the PM engages with the multi-function control display unit (MCDU). However, on approach, PF and PM show fairly similar, rather than complementary scanning strategies. They both spend the majority of their time looking out the window, followed by fixating the primary flight display (PFD). When excluding the out-the-window view from the analysis, both the PF and the PM spend more than 50% of the dwell time on the attitude indicator (ADI) during manual approaches. During coupled approaches, the dwell time on the ADI significantly decreases while it more than doubles for the horizontal situation indicator (his) for both pilots. In the absence of a PM, i.e., during single-pilot operations, scanning patterns become more dispersed as the single pilot needs to spend more time on secondary instruments. Pilots compensate by monitoring primary instruments less and, during approaches, they transition more often between flight deck instruments and the outside view, with significantly shorter dwell times on the outside.

Increased Surveillance

Whereas routine monitoring is a highly practiced behavior, increased surveillance requires a more effortful attention-demanding top-down approach to information gathering. It involves understanding which cues are relevant to performing a particular task or meeting a specific goal. For example, if one pilot lacks knowledge or experience with a particular task, procedure, or equipment, the second pilot may be able to adapt their own monitoring strategy to compensate for the deficiency:

> *"from a CRM standpoint, the captain told me he needed me to help him by backing him up because he was not as current in the airplane as I was at that time." (Young, 2020, p. 143)*

This example demonstrates that, even though one pilot may be relatively unfamiliar with an aircraft, the flight crew as a whole is able to compensate by changing how closely one pilot monitors the actions of the other. Similarly, awareness and appropriate understanding of a situation emerges from the joint actions of the entire flight crew rather than a single pilot. For example, if one of the pilots is unsure of the altitude for which the crew was cleared, he/she can ask the other pilot for confirmation. This could be viewed either as a negative, i.e., a failure of the first pilot or as a positive, i.e. the successful mitigation of a potentially dangerous situation through collaboration and redundancy (Weber & Dekker, 2017).

Another example of how context shapes pilots' monitoring behavior is described by the First Officer (FO) of a Boeing 757:

> *"As you descend in mountainous areas at night, you are aware that there is terrain that you might not be able to see. That knowledge makes you more aware of the importance of being on course, because your course is what provides you clearance through the terrain"* (Young, 2020, p. 270)

In this case, the FO recognizes that if certain important cues are unavailable (i.e., the presence/location of mountains), their monitoring strategy must change, and attention must be re-directed to other cues that can act as a substitute to accomplish the goal of staying clear of terrain. In both examples, pilots show a capacity to absorb disturbances or vulnerabilities in the system based on their knowledge of the system, the environment, and other agents.

In a recent whitepaper, ALPA (2019) argues that the transition to SPO is unlike previous reductions in crew complement. The latter reduced redundancy on the flight deck but SPO eliminates it. There is some evidence in the SPO literature that, in fact, this may be a major issue in off-nominal conditions. Faulhaber (2019) conducted a study to evaluate workload changes between two-pilot and single-pilot crews and discovered that workload was higher in SPO for turbulence and in abnormal conditions but not during a baseline scenario. Even though pilots did not perceive workload as a major challenge, performance decrements were observed. The SPO condition resulted in one crash landing due to a failure to extend landing gear. In the two-pilot condition, the PF made similar errors, but these errors were caught by the PM. These results highlight the value of the second pilot as an error checking mechanism.

Another recent eye tracking study that compared pilots' monitoring strategies in single pilot vs. two-crew operations revealed that the PF, under SPO, may be susceptible to spending more time monitoring secondary displays (e.g., flaps, landing gear, electronic centralized aircraft monitor) that tend to be monitored by the PM in current two-crew operations (Faulhaber, Friedrich, & Kapol, 2020). This could result in degradation in overall monitoring performance, which was confirmed in a series of studies conducted by Bailey and colleagues (2017). For example, in an off-nominal scenario involving unreliable airspeed indications, Etherington et al. (2017) found that SPO and RCO crews were less likely (33% and 56%, respectively, compared to 67% of the two-pilot crews) to detect an indicated airspeed (IAS) disagree warning light before other, more salient, warnings were issued, such as an overspeed clacker.

2.3.3 Respond

Responding to adverse events involves complex problem solving, decision making, and taking appropriate actions to mitigate the threat. Often times, these actions must be performed in limited time, with missing or conflicting information, and under uncertainty, thus requiring knowledge-based performance, expertise, inference, and judgement.

2.3.3.1 Discussing and deciding

Making decisions involves gathering and integrating information in real time, generating possible options and alternatives, and selecting an action to take. Carroll et al. (2019) discovered that, during the information gathering and integration stage, pilots routinely encounter conflicting information from multiple sources. Examples of information conflicts include weather conflicts between onboard radar and Air Traffic Control (ATC) or Next Generation Weather Radar, and conflicting traffic information between the Traffic Collision Avoidance System (TCAS) and ATC. Pilots view such information conflicts as a part of their regular workflow, i.e., they are accustomed to evaluating, comparing, and integrating potentially conflicting information from multiple sources with varying levels of integrity to decide on an appropriate plan of action. Pilots decide which source(s) of information to trust based on factors such as the recency of information, reliability of source, and their personal knowledge about the strengths and weaknesses of each source.

Decision making in complex environments also often requires coordination with other people. Here, too, pilots demonstrate the unique ability to consider multiple perspectives, evaluate tradeoffs, and come to a mutual agreement that satisfies all parties involved. For example, when discussing possible alternates for a diversion caused by approaching thunderstorms, a pilot states:

> *"Between the two or three of us with dispatch, we continually monitored the weather and tried to make the best possible decision. Like I said, I was more for going to Dulles, which was open at the time, and they were both like, 'Yeah, but Dulles, they've got that thunderstorm there, close by, and they're predicting it's going to move in. I think BWI is clear and a million, and we're going to be pretty safe going in there'. . . It was actually a little bit closer than Dulles, although either one of them were super close. Between the three of us, we gathered the information, made a decision we were all comfortable with. I was comfortable going with Baltimore" (Kiernan, Cross, & Scharf, 2020, p. 38).*

2.3.3.2 Taking action in response

Pilots occasionally need to respond to novel or unexpected events, such as equipment/automated systems failures or route and last minute runway changes. Actions or procedures that address these events may not be specified, or may require intentional deviation from SOPs. According to the traditional Safety-I view; such non-compliance or violations should be avoided at all costs to prevent incidents and accidents. However, pilots may not always have a choice; they sometimes need to deviate from standard procedures in the interest of the efficiency and safety of flight. Their ability to recognize this need and adapt to both anticipated and unanticipated events should be viewed as a positive.

For example, only 12% of flights fully adhere to the lateral and vertical profiles prescribed by area navigation standard terminal arrival route (RNAV STAR) procedures (Stewart, Matthews, Janakiraman, & Avrekh, 2018). Holbrook et al. (2020) highlight that more than 40% of cases of non-adherence reported to ASRS were *intentional* deviations from the course, altitude, and speed profiles. Many of these deviations were initiated by ATC or the flight crew to account for factors such as unexpected aircraft performance, automated systems issues, traffic or runway changes, and flaws inherent in RNAV procedure design. Holbrook et. al. found these adjustments and behaviors of resilient performance are not rare, one-off occurrences but were estimated by pilots to be a necessary part of roughly half of all RNAV arrival operations into KCLT, Charlotte/Douglas International Airport.

Similar adaptations and innovations were reported by Etherington and colleagues (2016) who conducted a series of flight simulator studies comparing current two-crew and proposed reduced-crew and single-pilot operations. The studies involved routine scenarios and off-nominal conditions. In a simulated rudder failure scenario, for example, some pilots developed a new way to apply the 40-pound force required to control yaw by jamming their foot under the rudder pedal. This allowed the PF to focus on manually flying the aircraft while the PM helped run checklists and communicate with ATC. Such techniques are not part of standard training procedures and require out-of-the-box, creative thinking.

In RCO and SPO, the absence of a second pilot on the flight deck can be expected to increase the workload for the pilot in command (PIC) who needs to continue flying the aircraft while performing other duties such as running checklists. Empirical studies have shown that it takes longer to start the correct checklist following an adverse event in SPO and RCO, compared to two-pilot crew operations (Etherington, Collins, Kramer, Bailey, & Kennedy, 2017). ATC communication was shown to disrupt checklist flow for SPO configurations significantly. In RCO, the post-rest debriefing is critical to bringing the resting pilot up to speed when they rejoin the flight during an ongoing emergency. Failure to mention important events or pieces of information could lead to wrong assumptions/actions. In one particular case, a failure of the PF to mention which checklist items had been completed resulted in the rejoining pilot to erroneously assume completion of some items (Bailey, Kramer, Kennedy, Stephens, & Etherington, 2017).

2.3.4 Learn

For an environment as complex and dynamic as commercial aviation, continuous learning is critical to maintaining the ability to handle unexpected events and cope with disturbances in the system. Pilots learn through a combination of formal training (e.g., classroom and simulator training) and informal training (observing more senior pilots, discussing with fellow pilots critical incidents and lessons learned). Both formal and informal learning practices greatly contribute to pilots' knowledge of the airplane and overall aviation system and their ability to recognize and respond to a wide range of contingencies.

2.3.4.1 Formal learning

In a qualitative study by Kiernan and colleagues (2020), all pilots pointed to simulator training and past real-world experience as their main sources of recognizing and handling off-nominal events:

> *"Every year we train in the simulator for all kinds of different problems"*
> *(Kiernan, Cross, & Scharf, 2020, p. 39)*

> *"I think there is pattern matching that goes on. I think I find in other emergency situations I have handled in my career there's pattern matching"*
> *(Kiernan, Cross, & Scharf, 2020, p. 39)*

Training equips pilots with the necessary skills, rules, and knowledge to perform manual flying tasks, manage their flight path, and engage in problem solving and decision-making. Over time, as pilots are exposed to a wide variety of scenarios, they become able to quickly recognize relevant cues, features, and subtle patterns associated with a previously experienced problem (McNeese, Salas, & Endsley, 2001). This ability, in turn, enables them to engage in recognition-primed decision-making, a form of decision making that relies heavily on quick situation assessments and the activation, execution and evaluation of if-then rules for known problems. When faced with problems that are novel, unpredictable, and rapidly changing, pilots switch to a more effortful analytical approach to decision making which requires (shared) mental models to diagnose and develop solutions through trial and error.

2.3.4.2 Informal learning

Pilots also benefit from on-the-job informal training by observing more experienced colleagues on the flight deck (Cummings, Stimpson, & Clamann, 2016) and by being exposed to a wide variety of challenges and conditions. This allows them to build a repertoire of strategies and often-unwritten knowledge. An example of how such experience can contribute to safety is described by Kiernan et al. (2020):

> *"Experience because many airports that have construction anywhere near the end of the runway, have frequently had their instrument or glide slope and localizer antennas interfered with by construction or vehicles driving right in front of them... Personal experience, since I was a private pilot, you just land the airplane" (Kiernan, Cross, & Scharf, 2020, p. 40).*

In current two-crew operations, informal training takes place between the two pilots through casual conversations and post-flight debriefings. SPO operations will severely limit the extent to which this type of training can occur. Schmid & Stanton (2019) suggest that an alternate apprenticeship-style training could be developed by allowing a pilot to observe single-pilot operations before transitioning to performing SPO duties themselves.

In summary, pilots contribute to safety and risk mitigation through a range of resilient behaviors.

- They anticipate and prepare for potential problems;
- monitor for actual deviations or notice potential deviations from nominal or desired system status;
- respond to adverse events by applying known or developing new solutions, which, in some cases, necessitates deliberate deviations from standard procedures; and
- learn through training and real-world experiences and observations and thus develop a wide repertoire of skills, knowledge, and mental models.

In the context of high-tempo and off-nominal situations, accomplishing these tasks requires two pilots who complement and offload one another.

2.4 Pilot contributions to safety and risk mitigation – the SRK framework

In addition to describing pilot contributions to safety using the above taxonomy of resilient performance, the Skills-Rules-Knowledge (SRK) framework (Chauvin, Lardjane, Morel, Costermann, & Langard, 2013) (Figure 4) can be employed to establish when human pilots are needed the most to mitigate risks (Rasmussen, 1983). The SRK framework describes how a task can be performed at three different performance levels.

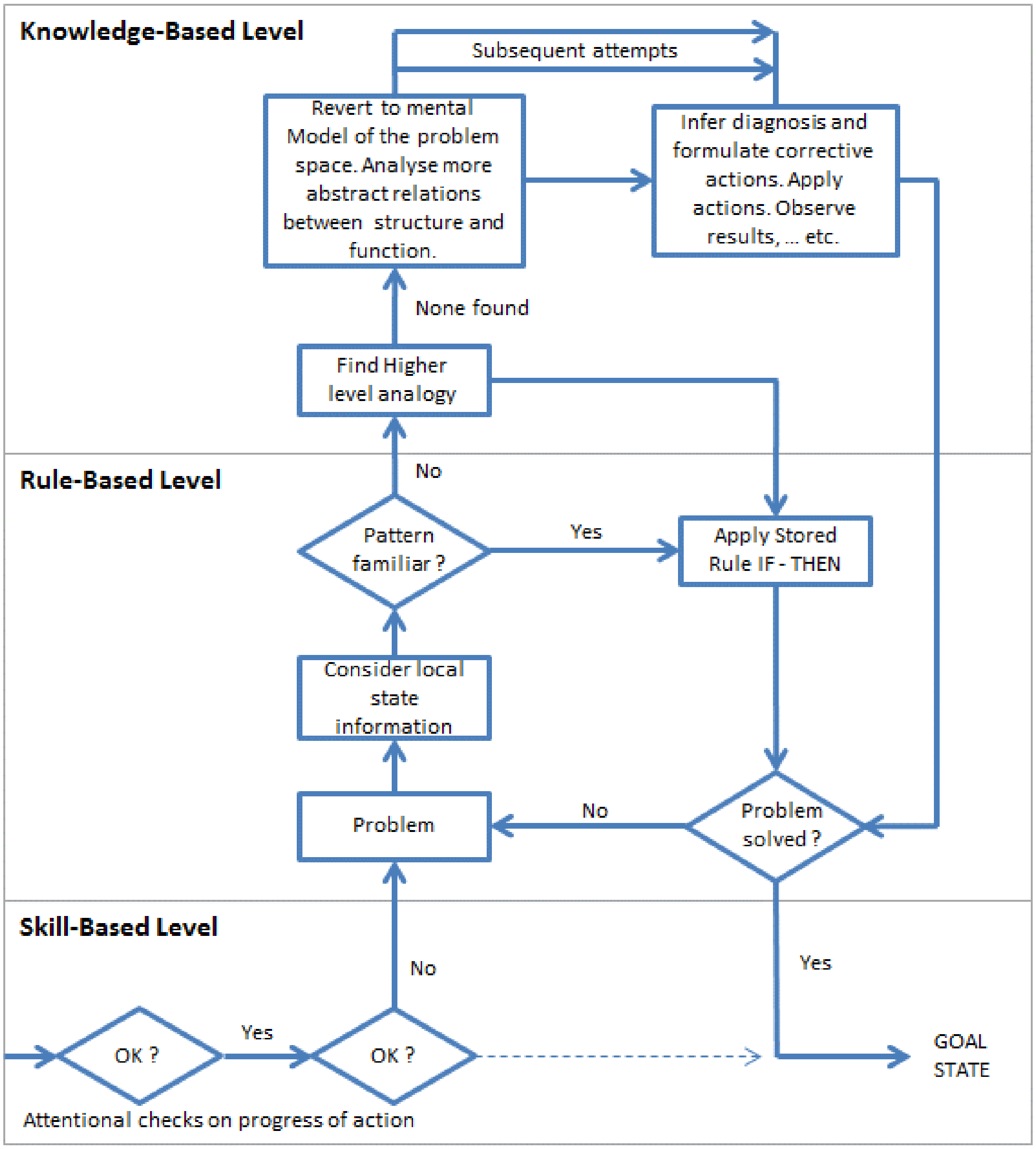


Figure 4. The SRK framework

Skill-based performance involves highly practiced routine behavior that can be executed with little or no conscious attention required. A well-trained pilot manually flying an aircraft in VFR conditions is a good example of performance at this level. In contrast, rule- and knowledge-based performance is required to cope with non-routine events and off-nominal circumstances. Rule-based performance is triggered when a known, previously encountered problem (e.g., windshear or an engine failure) is experienced. It involves the recall and application of stored if-then rules learned in training or created based on past occurrences. Rule-based and even more so skill-based tasks have been automated to quite an extent, as required inputs to the system can be specified and sensed and associated responses are rather well defined. However, when faced with a novel problem for which no response is prescribed, as in the case of UAL flight 232, pilots need to engage in knowledge-based performance. At this performance level, they use their knowledge of flying and aircraft systems to reason through the problem, and develop and test possible solutions through trial-and-error.

Knowledge-based performance is the most cognitively demanding of the three performance modes; it involves effortful thinking, reasoning, and judgement to cope with potential gaps in system design, resource shortages, and unanticipated conditions. In the case of UAL flight 232, the loss of all three hydraulic systems had been deemed highly unlikely and had never been experienced before. No checklist existed for how to deal with the situation. The pilots had to develop their own solution to the problem ‘on the fly’. They managed to gain some control of the aircraft with engine power alone and ended up saving 185 lives. As Captain Haynes stated: “Up there, we found ourselves in a whole new world. None of us had ever been in that type of situation. No simulator exercise had prepared us. Loss of hydraulics just was not supposed to happen. We had no procedures to follow.” (Orlady & Orlady, 2016, p. 359).

Human pilots continue to perform tasks at all three performance levels but their contributions to safety become more critical as we move from skill- to rule- to knowledge-based performance. Many skill-based tasks like manual flight control are good candidates for automated systems, which can perform them with high precision and reliability, without being affected by factors such as vigilance decrement and fatigue. The best opportunity for effective collaboration between human and machine agents may exist at the rule-based level of performance. Rule-based tasks require some interpretation and pattern recognition but also involve well-defined rules and procedures that can be executed by automated systems. Knowledge-based performance, on the other hand, will likely remain in the hands of pilots for the foreseeable future because they require expertise-based intuition, judgement, and creativity. In the face of highly variable and diverse problems with vaguely defined goals, high uncertainty, and incomplete information, the human’s ability to develop, hypothesize, and test multiple solutions provides a critical safety net.

Could modern technology have performed as well as the pilots on flight 232? Does it possess the skills and knowledge required to cope with the kind of novel situation faced by the flight crew? Can it play the role of teammate in reduced crew operations? The following section discusses these questions.

2.5 Human pilots – why they are still vital after all these years

Proponents of SPO and RCO believe that the development and introduction of new highly advanced technologies can make it possible to achieve the same level of safety in reduced crew operations as demonstrated by current two-pilot crews. Findings from empirical studies to date do not support this claim; it is important to keep in mind, however, that these experiments have examined the effectiveness of existing, less advanced systems. These studies show that while perceived as being potentially useful, the current level of sophistication of these tools falls far behind that of a human pilot, and their interface design is often poor (Cover, Reichlen, Matessa, & Schnell, 2018; Dao, et al., 2015; Lachter, et al., 2014; Lachter J. , Brandt, Battiste, Matessa, & Johnson, 2017; Lachter J. , et al., 2014; Ligda, et al.; Lim, Gardi, Ramasamy, & Sabatini, 2017; Liu, Gardi, Ramasamy, Lim, & Sabatini, 2016). For example, Lachter and colleagues (2014) tested a single-pilot and a two-crew configuration with and without collaboration tools (video feed of pilot and ground operator, mechanism for tracking responsibility, actions, and acknowledgements, and shared charts/displays). Pilots generally preferred the baseline current two-crew operations because the value of the human-autonomy teaming (where humans and machines share authority to pursue common goals (Lyons, Sycara, Lewis, & Capiola, 2021)). The use of collaboration tools was hindered by factors such as usability issues and a clunky interface that made it difficult to use touch gestures, the placement of the video feed of the remote pilot outside of peripheral vision, and speech-to-text algorithms that were talking over and interrupting pilot conversations.

Other concerns voiced regarding automated systems replacing a second human pilot are the inability of the automated systems to be proactive, follow CRM etiquette, have sufficient procedural knowledge, perform cross-verification, and respectfully correct the pilot in command when needed (Cummings, Stimpson, & Clamann, 2016). Pilots feel that current tools are not context-sensitive, lack the ability to make judgements, and do not support dynamic allocation of tasks based on roles, as opposed to direct assignment and micromanagement of responsibilities (Cover, Reichlen, Matessa, & Schnell, 2018; Geiselman, Johnson , & Buck, 2013).

Limitations of modern flight deck technology have been noted irrespective of the issue of crew complement. In this section, we will discuss the strengths and limitations of both human pilots and machines. For example, machines are excellent at solving known and well-structured problems and generating mathematically optimal solutions rapidly out of a set of alternatives (Roitblat, 2020). Deep learning has made major leaps forward in recent decades in improving object detection and recognition but, as illustrated in Figure 2, they can still be tricked fairly easily into misclassifying objects using simple techniques such as adding stickers to stop signs, introducing noise to images, and scaling/rotating the input images (Heaven, 2019). In contrast, while not infallible, humans can identify previously recognized objects in a wide range of conditions and contexts. In addition, humans have the ability to use analogues and metaphors to understand and deal with new situations based on experience (Klein, 2017).

In a recent article, Toews (2021) discussed four main fundamental shortcomings of current technology that necessitate the continued presence of human pilots on the flight deck. Automated and AI systems (1) lack common sense, (2) do not learn and adapt on a continuous basis, (3) do not understand cause-and-effect relationships, and (4) cannot reason ethically. The lack of common sense in machines can be explained by the fact that humans possess an immense body of knowledge about how the world works. As Toews points out common sense is built up over the course of our lives as we "develop persistent mental representations of the objects, people, places and other concepts that populate our world—what they're like, how they behave, what they can and cannot do." In contrast, AI mostly develops insights by uncovering statistical relationships in vast amounts of raw data, but it does not possess semantically grounded representations of objects. This makes it impossible for a machine to develop a deep, meaningful understanding of the world and cope with unforeseen events and circumstances. Attempts have been made to codify common sense into a set of rules to be shared with technology (Metz, 2016) but the promise and success of this approach continues to be a matter of debate. One reason these efforts struggle is that for nearly every rule, there is some exception or variation that would need to be captured, which leads to an ever-growing database that needs to be created and fed to the machine by humans.

The second limitation of existing technology highlighted by Toews (2021) is the way AI tends to learn. It is trained on an existing dataset to learn a task. At the end of training, the parameters of the model are fixed, and the system is deployed to generate insights based on a new data set. The model needs to be retrained and adjusted whenever new data or contexts emerge. This approach is referred to as batch-based training and deployment. It differs significantly from how humans learn and adapt. They face a continuous stream of new data and changing circumstances that they manage to incorporate dynamically into their knowledge and understanding of the world. They engage in what Toews refers to as "train and deploy in parallel and in real-time." As a result, humans have the ability to engage in knowledge-based performance in the context of novel problems or events. They draw from a deep and context-sensitive understanding of the system and its interdependencies (i.e., their mental model), and they have the capability to frame and reframe problems as they evolve, and new information becomes known. Machines, on the other hand, are limited to working with algorithms and representations created by designers and engineers during development and often fail under situations that are even slightly different from the conditions under which they were trained. Addressing this brittleness is not simply a matter of feeding the algorithms more data and trying to identify as many edge-cases as possible because there will always be situations or combinations of conditions that are completely novel and unanticipated (Dickson, 2020; Wood, 2021).

Current AI outperforms humans at uncovering correlations – associations in data – but it does not understand the causal mechanisms that explain emerging patterns. Yet, causal reasoning is critical for the ability of humans to make sense of and shape the world. For example, assume a correlation is observed between the color of cars and their accident rate, with red cars being involved in more accidents than any other color car. Does that mean we should avoid buying a red car? Not really, as various causal relationships may underlie the observed correlation. It could be that our visual system is not as good at perceiving the distance and speed of red objects and therefore we tend to misjudge the speed and distance of approaching red cars and collide with them more often. The correlation may have nothing to do with the color itself. It could just be that people who prefer red cars are more thrill seeking than the average driver. Both hypotheses, individually and jointly, may explain the higher accident rate but only the first should make us avoid buying a red car. As Pearl & Mackenzie (2018) point out, the key to building truly intelligent machines "is to replace reasoning by association with causal reasoning." Recent work in AI has begun to address this challenge (Gerstenberg, Goodman, Lagnado, & Tenenbaum , 2021; Madan, Ke, Goyal, Schölkopf, & Bengio, 2021) but it remains unsolved at this time.

Finally, machines currently cannot reason ethically. They have no sense of morality and human values and cannot determine the ethical significance of statements, decisions, or actions – the so-called alignment problem which needs to be resolved before entrusting machines with safety-critical decisions. As Christian puts it:

> *"As machine-learning systems grow not just increasingly pervasive but increasingly powerful, we will find ourselves more and more often in the position of the 'sorcerer's apprentice': we conjure a force, autonomous but totally compliant, give it a set of instructions, then scramble like mad to stop it once we realize our instructions are imprecise or incomplete—lest we get, in*

> *some clever, horrible way, precisely what we asked for.”* (Christian, 2020, p. 19)

Attempts to address this challenge (e.g., (cooperative) inverse reinforcement learning (Russell, 2019)) try to build systems that observe human behavior, try to figure out what human values the behavior is based on and then align themselves with those values.

Another ability possessed by humans but not modern technology is emotional intelligence (not discussed by Toews (2021)). Emotional intelligence has been defined as “the subset of social intelligence that involves the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions” (Salovey & Mayer, 1990, p. 189). Emotional intelligence involves 1) appraisal and expression of emotion, 2) regulation of emotion, and 3) utilization of emotion. The first two capabilities refer to both the expression and regulation of one’s own emotions and emotions of others, and the third capability manifests itself in processes such as flexible planning, creative thinking, redirected attention, and motivation. Concepts like emotion, confidence, and motivation are often overlooked as contributors to safety and human performance in complex systems, despite the fact that perseverance through extreme challenges and system breakdowns in multiple aviation accidents has been attributed to qualities like “realistic optimism” and “cheerful confidence” (Reason, 2008). Consider, for example, the level of calmness and optimism demonstrated by Captain Moody of British Airways Flight 09, which experienced a loss of all engines due to volcanic dust caused by an eruption of Mount Galunggung:

> *“Ladies and Gentlemen, this is your Captain speaking. We have a small problem. All four engines have stopped. We are doing our damnedest to get them going again. I trust you’re not in too much distress.”*
>
> – Captain Eric Moody, British Airways Flight 09

The Captain and crew not only showed a sense of calm and humor during a potentially fatal catastrophic event, but they were also confident and persistent in repeatedly attempting engine restart procedures while considering a backup solution of ditching the aircraft on the water in the Indian Ocean. Eventually, the crew was able to get three engines running again and landed safely despite being visually impaired by the dust and the lack of guidance from an instrument landing system (ILS).

2.6 Conclusion

In summary, the findings from the literature review show that pilots continue to play a critical role in ensuring flight safety, especially in the context of unexpected and/or novel events requiring knowledge-based performance. Pilots' ability to make sound judgments derived from experience rather than study. Their capacity for causal reasoning and their ability to continually learn and incorporate new information into their model of the world allows them to anticipate potential problems, make sense of a given situation, adapt their behavior to changing circumstances, and develop innovative solutions to never before experienced off-nominal events. These events, in particular, call for the presence of a second pilot who can offload the PF by taking over tasks and thus manage workload and by providing monitoring and operating redundancy.

The following section illustrates pilot contributions to safety by analyzing accidents where the flight crew rescued the situation and prevented the worse possible outcome of an adverse event. The findings from this analysis will be used to create an updated version of the original SRK framework that was introduced in this chapter.

3 Analysis of select aviation accidents to identify contributions of flight crews

3.1 Accidents where pilots rescued the situation and prevented a worst-case outcome

Maintaining and improving the safety of aviation operations requires careful analysis of a variety of data. Traditionally, the focus in aviation safety has been on reviewing incidents and accidents to learn, after the fact, "what went wrong" and help prevent similar events from re-occurring (an approach referred to as Safety-I). More recently, it has been recognized that it is equally important to analyze routine operational data to identify contributors to the very high level of aviation safety, i.e., to "learn what goes right", and to detect, early on, potential trends and concerns before they lead to unwanted and potentially dangerous events. This approach, called Safety-II, is taken in several ongoing efforts, such as ENAIRE's Normal Operations Monitoring (NOM), American Airlines' Learning and Improvement Team (LIT), and Cathay Pacific Airways' Operational Learning Review (OLR). Similarly, Flight Safety Foundation's (FSF) 'Learning From All Operations (LFAO)' concept (Flight Safety Foundation, 2021) calls for a fundamental shift from the traditional reliance on mishap data to using information about normal operations for safety improvements. FSF highlights safety management systems (SMS) that have been established by most aviation organizations to identify hazards and proactively manage risks still focus on “the absence of safety, rather than its presence”. As a result, what is being learned is valuable but limited in its applicability and timeliness. Instead, LFAO aims to understand how work actually is done and how personnel cope successfully with challenges and operational variability that they encounter.

In this section, we first present the findings from task 3 in this research effort, an analysis of select aviation accidents to identify positive contributions made by flight crews. This task complements Safety-I and Safety-II in that it combines the use of traditionally relied upon accident data with the goal to learn “what went right” during the event (Table 1). Fourteen accidents were analyzed where the outcome of the accident was better than might have been expected, due to the positive contributions made by pilots/flight crews.

Table 1. Task 3 in the overall safety management space

	Flight Crew Actions	
	‘What Went Wrong’	**‘What Went Right’**
Routine Operations		**Safety II**
Off-Nominals and Accidents	**Safety I**	**Task 3**

3.1.1 Method

3.1.1.1 Accident selection

Fourteen accidents where the outcome of the accident was better than might have been expected were identified primarily from the FAA ‘Lessons Learned’ database and the ALPA airmanship awards. The database search was limited to accidents that (a) involved commercial transport airplanes, (b) occurred after 1990, and (c) where the outcome of the mishap was rather positive despite unfavorable circumstances. A small number of accidents that occurred between 1985 and 1990 were included only because they were extremely relevant to highlighting the contributions of flight crews to the safety of flight. Some recent accidents, while relevant, were not included due to the lack of an official accident account, documentary, and/or investigation report (e.g., Ural Airlines Flight 178 on August 15, 2019; Volga-Dnepr Airlines Flight 4066 on November 13, 2020). The 14 accidents that were ultimately selected for full analysis are listed in Table 2. These mishaps were analyzed based on information contained in official accidents investigation reports (e.g., reports by the National Transportation Safety Board and the Australian Safety Transportation Board), peer-reviewed journal articles, disaster-investigation documentary videos, pilot interviews, and relevant website articles.

We used the SRK framework (Rasmussen, 1983) (Figure 4) introduced in section 2.4 to analyze the accidents listed in Table 2 below and categorize the nature of pilot contributions to safety. As a reminder, the SRK framework describes how a task can be performed at three different performance levels.

Table 2. List of accidents identified for analysis

Date	**Accident Flight**	**Description**
Apr 28, 1988	Aloha Airlines Flight 243	Explosive decompression due to metal fatigue
May 24, 1988	TACA Flight 110	All-engine failure due to rain/hail ingestion
Jul 19, 1989	United Airlines Flight 232	Loss of all hydraulics
Dec 27, 1991	Scandinavian Airlines Flight 751	Dual-engine failure due to ice ingestion
Aug 24, 2001	Air Transat Flight 236	Loss of both engines due to fuel leak
Sep 28, 2007	American Airlines Flight 1400	In-flight engine fire
Jan 17, 2008	British Airways Flight 38	Loss of thrust during final approach
Oct 07, 2008	Qantas Flight 72	Uncommanded pitch-downs during flight
Apr 13, 2010	Cathay Pacific Flight 780	Engine failure due to fuel contamination
Nov 04, 2010	Qantas Flight 32	Catastrophic engine failure
May 11, 2015	ExpressJet 4291	Failure of all air data computers mid-flight
Sep 30, 2017	Air France Flight 066	Uncontained engine failure
Feb 13, 2018	United Airlines Flight 1175	Partially contained engine failure
Apr 17, 2018	Southwest 1380	Single-engine failure during climb

Accident Analysis: The Skills, Rules, and Knowledge (SRK) Framework

Skill-based performance involves highly automated routine behaviors that can be executed smoothly, with little or no conscious thought or attention to the task at hand. Such behaviors use a very efficient and dynamic internal world model that is largely based on feedforward instead of feedback control. Thus, skill-based behaviors are not explicitly goal-oriented and operators engaged in skill-based performance are often unable to describe how they performed the task and what information they used. An example of skill-based performance is a well-trained pilot manually flying an aircraft in routine VFR conditions.

A switch from skill- to rule-based performance occurs when an operator encounters some kind of off-nominal event or circumstance. Rule-based performance involves the recall and conscious application of stored if-then rules and procedures by operators performing an unfamiliar task and in response to non-routine but previously experienced problems (e.g., windshear or a single-engine failure). Such rules and procedures are learned in training and/or developed based on operational experience. They are communicated formally (e.g., a checklist) or informally (e.g., verbal exchanges between pilots). At the rule-based performance level, the goal is not necessarily explicitly formulated but the operator is usually able to describe the rules and procedures used. An example of rule-based performance is completing a checklist after experiencing an engine failure.

Finally, knowledge-based behaviors involve conscious real-time planning and problem solving to cope with novel events for which no rules are available from previous experience. In knowledge-based performance, the goal is explicitly formulated based on an analysis of the environment and overall aims of the operator. Several plans are developed and are tested against the goal physically (by trial and error) or conceptually (by mental simulation of the considered plan) before selecting a final plan. Knowledge-based performance is the most cognitively demanding of the three performance levels and relies heavily on an accurate mental model of the system.

An example of this type of performance was seen during the accident of UAL flight 232, a DC10 that experienced a loss of all three hydraulic systems resulting from an uncontained engine failure. This type of event had been deemed highly unlikely and had never been experienced before. When the pilots lost control of the aircraft’s pitch and roll, they had to develop their own solution to the problem ‘on the fly.’ They managed to regain some control of the aircraft by using asymmetric engine power and ended up saving the lives of 185 passengers and crew. As Captain Haynes stated: “Up there, we found ourselves in a whole new world. None of us had ever been in that type of situation. No simulator exercise had prepared us. Loss of hydraulics just was not supposed to happen. We had no procedures to follow.” (Orlady & Orlady, 2016, p. 359)

The analysis of the 14 accidents was performed in three stages. First, official accident reports along with relevant journal articles, disaster-investigation documentary videos, pilot interviews, and website articles were reviewed to identify flight crew behaviors that contributed to the safe outcome. Second, the behaviors were categorized as skill-, rule-, or knowledge-based performance. Third, the two researchers conducting the analysis crosschecked each other's work to ensure that the flight crew actions were categorized appropriately.

The following sections present the 14 accidents in chronological order. Each section starts with an executive summary of the accident. Following each summary, flight crew actions that went beyond routine skill-based performance and the mere execution of checklists and prescribed procedures at the rule-based performance level are listed. Notably, these actions involved the adaptation to existing rules or procedures and the development and implementation of solutions to novel problems at the knowledge-based performance level. Factors that were not under the control of the flight crew but helped the pilots 'save the day' are also identified. In a small number of cases, the flight crew made significant contributions to risk mitigation but also committed errors. For example, in the case of Aloha Airlines flight 243, the NTSB found that "the flight crew's use of a target speed of 280-290 KIAS and speedbrakes in the descent after the structural separation indicated they did not consider the appropriate emergency descent checklist which states, in part, that if structural integrity is in doubt, airspeed should be limited as much as possible and high maneuvering loads should be avoided." (National Transportation Safety Board, 1989, p. 71). Given the focus of this project on positive contributions by flight crews, these erroneous actions and assessments will not be included in the analysis.

3.1.2 Aloha Airlines flight 243

3.1.2.1 Executive summary

On April 28, 1988, a Boeing 737-200 operated by Aloha Airlines Inc., as flight 243, experienced an explosive decompression and structural failure at 24,000 feet, while enroute from Hilo to Honolulu, Hawaii. Approximately 18 feet of the cabin skin and structure aft of the cabin entrance door and above the passenger floorline separated from the airplane during flight (Figure 5).

Figure 5. Aloha Airlines Inc., flight 243 aftermath

Following the rapid decompression, the flight crew put on oxygen masks and the Captain immediately began an emergency descent, extended the speedbrakes, and activated the passenger oxygen-supply system. Given the ambient noise in the flight deck, the pilots could not hear each other and thus used hand signals to communicate during the initial part of the descent until noise levels dropped to acceptable levels when voice-communication was again possible.

The FO lowered the landing gear during the approach but the nose gear position-indicator light did not illuminate. Manual nose gear extension was performed but the nose gear green indicator light still did not illuminate, even after two attempts. The Captain informed ATC that they would land without the nose gear extended and requested available help and rescue equipment. The Captain subsequently noticed a yawing motion when advancing thrust levers during approach and determined that the left engine had failed. He placed the No. 1 engine start switch in the “flight” position in an attempt to restart the engine but there was no response. Post-accident investigation revealed that the left engine became inoperative because the engine control cables separated because of an increase in cable tension caused by the cabin floor deformation.

A stable descent profile was established 4 miles out on final approach, and the Captain successfully landed the aircraft on runway 02 at Maui's Kahului Airport. There were 89 passengers and 6 crewmembers onboard. One flight attendant was swept overboard during the decompression and was presumed to have been fatally injured; seven passengers and one flight attendant received serious injuries.

The NTSB (1989) determined that the probable cause of this accident was the failure of the Aloha Airlines maintenance program to detect the presence of significant disbonding and fatigue damage, which ultimately led to failure of a lap joint and the separation of the fuselage upper lobe.

Contributing to the accident were the following:

- the failure of Aloha Airlines management to properly supervise its maintenance force;
- the failure of the FAA to evaluate the Aloha Airlines maintenance program and to assess the airline's inspection and quality control deficiencies;
- the failure of the FAA to require Airworthiness Directive 87-21-08 inspection of all the lap joints proposed by Boeing Alert Service Bulletin SB 737-53A1039;
- the lack of a complete terminating action (neither generated by Boeing nor required by the FAA) after the discovery of early production difficulties in the B-737 cold bond lap joint, which resulted in low bond durability, corrosion, and premature fatigue cracking.

3.1.2.2 Flight crew actions

Adaptations to rules or procedures

- Only gentle turns to minimize impact on the damaged aircraft
- Bring landing gear down early during approach to maintain altitude
- Decide to use flaps 5 and IAS 170 knots for approach and landing because the aircraft became less controllable with higher flaps settings and speeds below 170 knots. This was based on the Captain’s observation that the airplane was "shaking a little, rocking slightly and felt springy.”
- Decide not to overfly the airport to confirm landing gear down (as should be done “by the book”) due to concerns of airframe integrity

Knowledge-based behaviors:

- Use hand signals to communicate during the initial part of the descent when wind, oxygen masks, and other ambient noises in the flight deck made it impossible to communicate verbally

3.1.2.3 Factors not under pilot control

- Nose gear was indeed locked during the touch-down even though neither the flight crew nor the ATC knew it was
- Very low winds (atypical for Maui area)
- Damage to aircraft was across the top instead of bottom (airframe structure members were therefore in tension rather than compression, which could cause buckling)

In addition, it is worth noting that the flight crew remained calm and maintained control of the aircraft in the face of a novel unexpected failure, as highlighted by the NTSB in their Safety Recommendation to the Federal Aviation Administration, dated July 21, 1989:

> *“The magnitude of the accident was well beyond any anticipated emergency scenario… The flightcrew’s success in managing the multiple emergency situations and recovering the aircraft to a safe landing speaks well of their training and airmanship.” (NTSB, 1989, p. 12)*

3.1.3 TACA flight 110

3.1.3.1 Executive summary

On May 24, 1988, Flight 110 (Figure 6) operated by TACA Airlines experienced extreme weather while traveling from Goldson International Airport, Belize to New Orleans International Airport. During the descent from FL 350 for an IFR arrival to New Orleans, the flight crew noted green and yellow returns on the weather radar with some isolated red cells, to the left and right of the intended flight path. Before entering clouds at 30,000 ft, the Captain selected continuous engine ignition and activated the engine anti-ice system. The crew selected a route between two severe weather cells displayed in red on the weather radar.

Figure 6. TACA flight 110 emergency landing

As the crew got closer to the red cells, heavy rain, hail, and turbulence were encountered. The high turbulence and heavy hail made it extremely difficult for the flight crew to read the flight instruments. At approximately 16,500 ft, both engines flamed out. The APU was started, and AC electrical power was restored while descending through approximately 10,600 ft. However, attempts to windmill restart the engines were unsuccessful. Neither engine would accelerate to idle power and advancing thrust levers increased the engines beyond limits. The Captain shut down the engines to avoid overheating and catastrophic failure.

Based on the amount of turbulence and their descent rate, the flight crew determined that options suggested by ATC for landing on nearby lakes or at a nearby airport were not feasible as the aircraft was already descending through an altitude of 3,000 ft. Since no airports were within range of the aircraft for an emergency landing, the flight crew considered landing on a nearby highway, as recommended by ATC, but decided against it due to concerns about colliding with traffic. After emerging from the clouds at around 5,000 ft, the Captain spotted a canal surrounded by buildings (~15 miles away) and considered it a potential ditching location as it would allow for a straight-in approach. While the aircraft was aligning with the canal, the First Officer spotted an adjacent levee as an alternative, better landing option. After a last-minute course correction, the flight crew proceeded to make an emergency landing on the levee without further damage to aircraft. All 45 passengers survived, with only one person sustaining an injury.

The post-accident investigation revealed that the aircraft had entered level 4 thunderstorms. The NTSB (1991) determined the probable cause(s) of this incident to be a double-engine flameout due to water ingestion, which occurred because of an inflight encounter with an area of very heavy rain and hail. A contributing factor was the inadequate design of the engines and the FAA water-ingestion certification standards, which did not reflect the precipitation rates that can be expected in moderate or severe thunderstorms.

3.1.3.2 Flight crew actions

Adaptations to rules or procedures

- Apply engine failure and restart checklist to both engines (no checklist was provided for dealing with a dual-engine failure)
- Apply braking in a controlled manner after touchdown due to uncertainty about the softness of the ground (i.e., not apply full braking, which could result in damage to aircraft and/or injuries to passengers)

Knowledge-based behaviors

- Identify and select a nearby canal as safest possible location to ditch the aircraft.
- Actively engage in revising current solution and develop new alternatives based on new information. At an altitude of around 1,500 ft, the FO identifies a flat grass area right next to the canal as an alternative landing spot. Despite being narrower and shorter than the canal, the levee was deemed a safer landing spot.
- Perform a sideslip maneuver to quickly reduce air speed and align with the levee (performing a sideslip maneuver on a large commercial airliner was unheard of at the time and outcome of performing such a maneuver on a Boeing 737 was uncertain).

3.1.4 United Airlines flight 232

3.1.4.1 Executive summary

On July 19, 1989, flight 232 operated by United Airlines departed from Denver, Colorado to Philadelphia, Pennsylvania with a scheduled stop in Chicago, Illinois. There were 285 passengers and 11 crewmembers on board the DC-10 aircraft. Approximately one hour and seven minutes after departure and at a cruise altitude of FL370, flight 232 experienced a catastrophic failure of the No. 2 tail-mounted engine. The flight crew heard a loud explosion that was initially presumed to result from a decompression. The First Officer took manual control of the aircraft, and the Captain and Second Officer examined the engine instruments. They determined that the No. 2 engine had failed and initiated engine shutdown procedures. The crew was unable to shut down the engine due to a jammed throttle and fuel lever but eventually succeeded by actuating the firewall shutoff valve to cut off fuel supply to the damaged engine. The separation and forceful discharge of the No. 2 engine's fan rotor assembly led to the loss of all three hydraulic systems that powered the airplane's flight controls.

Attempts to restore hydraulic power by activating the air-driven generator (ADG) were unsuccessful. Without functioning hydraulics systems, the flight crew experienced severe difficulties controlling the airplane. Despite full left aileron and full-up elevator inputs, the aircraft was in a descending right turn. The Captain immediately throttled down the No. 1 engine to prevent the aircraft from rolling over and established a wings level attitude. In addition, the crew had to counter up and down oscillations, known as phugoids that were induced by trim settings for a speed different from what could be achieved using only two engines and due to the aircraft's inability to maintain level flight because of damaged flight control surfaces.

After gaining partial control of the aircraft, the crew contacted Minneapolis Center to request emergency assistance. They accepted the Center's recommendation to land at the Sioux City airport. Around this time, a flight attendant advised the Captain that a UAL DC-10 training check airman, who was traveling as a passenger, had volunteered his assistance. The Captain immediately invited the airman to the flight deck and asked him to operate the throttles. The check airman attempted to use engine power to control pitch and roll.

The crew jettisoned fuel to the level of the automatic system cutoff, leaving 33,500 pounds, and executed the emergency manual landing gear extension procedure. After making visual contact with the airfield about nine miles out, the flight crew aligned with runway 22 for approach. The flaps and slats remained retracted. Based on experience with no flap/no slat approaches, the assisting pilot knew that power would have to be used to control the airplane's descent. He used the First Officer's airspeed indicator and outside visual cues to determine the flightpath (Figure

7) and the need for power changes. In the last 20 seconds before touchdown, the airspeed averaged 215 KIAS, and the sink rate was 1,620 feet per minute. Smooth oscillations in pitch and roll continued until about 100 feet above the ground, just before touchdown, when the right wing dropped rapidly, and the nose of the airplane began to pitch downward. The airplane touched down on the threshold slightly to the left of the centerline on runway 22 and subsequently crashed during the attempted landing at Sioux Gateway Airport. Out of the 285 passengers and 11 crewmembers onboard, one flight attendant and 111 passengers were fatally injured.

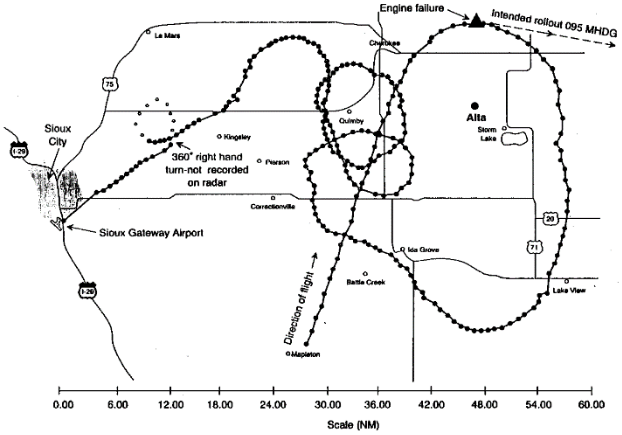


Figure 7. United Airlines flight 232 flightpath

The NTSB (1990) determined that the probable cause of this accident was the inadequate consideration given to human factors limitations in the inspection and quality control procedures used by United Airlines' engine overhaul facility. This resulted in the failure to detect a fatigue crack originating from a previously undetected metallurgical defect located in a critical area of the stage 1 fan disk that was manufactured by General Electric Aircraft Engines. The subsequent catastrophic disintegration of the disk resulted in the liberation of debris in a pattern of distribution and with energy levels that exceeded the level of protection provided by design features of the hydraulic systems that operate the DC-10's flight controls.

3.1.4.2 Flight crew actions

Adaptations to rules or procedures

- Use firewall shutoff valve as an alternative method to cut fuel supply to number two engine when it was not possible to do so due to jammed/binding fuel lever

- Communicate with ATC to figure out position of aircraft relative to Sioux City airport without an operational VOR (very high frequency omni-directional range navigation aid) station

Knowledge-based behaviors

- Contact San Francisco Aero Maintenance (SAM) facility to explore potential solutions with trained personnel and engineers (none were suggested)
- Use differential thrust on the remaining two engines to regain partial control of aircraft heading and altitude and prevent aircraft from rolling over (due to tendency of right wing to dip)
- Navigate to alternate airport with only right turns (airplane had a right turning tendency) – i.e., fly a ‘corkscrew’ pattern
- Use remaining two engines to counter phugoid oscillations (cyclic up and down oscillations induced by displacement of aircraft from level flight) by adding power on both engines simultaneously when the aircraft nose pitched down and reducing power when the nose pitched up. This maneuver allowed the aircraft to descend from cruise altitude (FL370) to the Sioux City airport.
- Experimentation and learning by “feel” what inputs to make to maintain control of the aircraft

3.1.4.3 Factors not under pilot control

- Additional Onboard Pilots: A DC-10 check airman was on board the aircraft as a passenger and joined the flight crew on the flight deck to help control the aircraft. The check airman had learned of the 1985 crash of Japan Air Lines Flight 123 which had been caused by a catastrophic loss of hydraulic control. Thinking about that accident, he had wondered if it was possible to control an aircraft using throttles only, and he had practiced this skill on a simulator prior to this flight.
- Favorable weather: clear day
- Favorable location: over flat lands of Iowa (i.e., not over densely populated cities or mountainous terrain), open corn fields near airport
- Favorable time: the time of emergency (late afternoon, around 4pm) coincided with shift changes for emergency services in Sioux city and surrounding communities, which

allowed for double the number of staff available to help injured passengers; 285 trained national guard personnel also happened to be available on this particular day of the month, ready to help on ground

3.1.5 Scandinavian Airlines System (SAS) flight 751

3.1.5.1 Executive summary

Scandinavian Airlines System (SAS) flight 751 took off from Stockholm/Arlanda airport on December 27, 1991. It had landed at Stockholm/Arlanda the previous day and had been parked outdoors overnight. Prior to takeoff, the aircraft was de-iced. At the moment of lift-off, the Captain heard an unusual noise, which he could not identify. The noise, which was recorded by the Cockpit Voice Recorder (CVR), resembled a low humming sound. Approximately 25 seconds into the flight, the right engine started to surge. Using the engine instruments, the pilots diagnosed a malfunction of the right engine, and the First Officer said "...think it's a compressor stall." The Captain stated that, because of the vibrations and the rapid changes in the display, he had difficulty reading the engine instruments. He throttled back the engine but this did not have an effect. The engine surges continued until the engine completely failed and stopped delivering thrust just 51 seconds after the first sign of trouble. Fourteen seconds later, the left engine also started to surge. The pilots did not notice the problem until this engine also failed completely. This was followed by an engine fire alarm thirteen seconds later. The First Officer extinguished the fire of the left engine.

A uniformed SAS Captain on the flight realized that the crew was experiencing major problems. He hurried to the flight deck and asked if he could be of any help. The First Officer asked him to complete the emergency/malfunction checklist, and the Captain instructed him to start the auxiliary power unit (APU). The crew prepared for an emergency landing. Approximately 420 m above the ground, and still in the clouds, the assisting pilot started gradually extending the flaps. The flaps were fully extended approximately 30 seconds later, at an altitude of approximately 300 m above the ground.

The aircraft broke out of the clouds at an altitude of 300 meters, and the Captain decided that a large field far to the right could not be reached. He elected instead to try to land in a smaller field in roughly the direction of flight. During the approach to the field, the Captain turned approximately 25 degrees to the right to avoid houses located beyond the intended landing site. Seventeen seconds before the aircraft struck the ground, the First Officer asked whether they should lower the landing gear. This was answered by the assisting Captain with the call: "Yes, gear down, gear down."

On final approach to the selected field, the aircraft collided with treetops. During an interview following the accident, the Captain indicated that this was intentional in order to lower the rate of descent. The tail of the aircraft struck the ground first. On impact, a major part of the right wing was torn off, and the aircraft continued sliding across the ground for approximately 110 meters before it came to a stop. The fuselage broke into three pieces but no fire broke out (Figure 8). Except for four passengers, everyone on board was able to exit the aircraft on their own. One passenger incurred a disabling back injury.

Figure 8. SAS flight 751 fuselage after landing

The Accident Investigation Board determined that the accident was caused by SAS instructions and routines, which were inadequate to ensure that clear ice was removed from the wings of the aircraft prior to takeoff. This resulted in the aircraft taking off with clear ice on the wings that came loose during takeoff, and climbout, and was then ingested by the engines. The ice caused damage to the fan stages of the engines, which led to surging and destroying both engines.

Contributory causes were determined to be that the pilots were not trained to identify and correct engine surges. In addition, automatic thrust restoration (ATR) – which was unknown within SAS – was activated and increased engine throttles without the pilots' knowledge. According to interviews done with the pilots, they believed that experiences and knowledge gained through Air Force training (e.g., landing outside of airports, managing tasks on your own, recalling by-heart items rather than relying solely on checklists) were critical in managing the emergency situation.

3.1.5.2 Flight crew actions

Adaptations to rules or procedures

- Pull fire extinguisher handle immediately after onset of left engine fire warning because of imminent danger of fire (rather than first waiting 10 seconds according to procedures). This action saved the lives of many people on board because the engines were severely damaged (Martensson, 1995).
- Turning 25 degrees to the right during final approach to minimize damage to people and infrastructure on the ground by avoiding houses beyond landing site.

Knowledge-based behaviors

- Select a nearby open field surrounded by trees for emergency landing. After exiting clouds below 1000ft, the Captain noticed a light spot surrounded by (snow-covered) green. A larger open field was also considered as a potential landing site but not selected because it was too far away given the low altitude of the aircraft.
- Gradually extend flaps to reduce aircraft speed and prevent stall. The SAS emergency checklist for MD-80 did not provide a configuration for speed and flap positions for approach and landing with both engines out. The assisting Captain had created his own checklist to deal with a dual engine failure scenario based on the DC-9. He successfully applied those steps in this situation to manage the approach speed.
- Use trees as a "pillow" to slow down descent before landing in an open area. In an interview for the Mayday Air Crash Investigation series, the Captain noted observing that the pine trees looked "soft" from above and thought that he could use the trees "almost like a pillow."

3.1.5.3 Factors not under pilot control

- Additional Onboard Pilot: A SAS Captain onboard the airplane hurried to the flight deck and helped by completing the emergency/malfunction checklist and starting the auxiliary power unit (APU). He also gradually extended the flaps in preparation for the emergency landing.

- No fire broke out despite the large amount of fuel spilled on crash site.

3.1.6 Air Transat flight 236

3.1.6.1 Executive summary

On August 24, 2001 at 00:52 UTC, Air Transat Flight 236, an Airbus 330, departed Toronto, Canada for Lisbon, Portugal with 13 crew and 293 passengers on board. The plane (Figure 9) took off with 104,500 lbs. (12,600 gallons) of fuel, including 5% reserve fuel and an additional 16% for potential in-flight rerouting and tankering (i.e., moving fuel from one tank to another to save on fuel costs). A fuel leak developed approximately four hours into the flight through a fracture that had developed in a fuel line to the right engine. At the time of the accident, standard procedures required that the flight crew periodically check fuel on board for discrepancies. The pilots completed these fuel checks six times during the flight but no anomalies were detected as the leak started slowly before growing rapidly.

Figure 9. Air Transat plane

At 05:03 UTC, twenty-five minutes after the fuel leak began, the crew observed low oil temperature and high oil pressure readings for the right engine, which were communicated to the dispatcher at the company’s Maintenance Control Centre (MCC) at Mirabel Quebec, Canada. Neither the MCC engineers nor the flight crew was able to make sense of the anomalies. At 05:33 UTC, the crew received a fuel imbalance warning, with 6,600 lbs. of fuel missing from the right wing. Three minutes later, the crew-initiated procedures to balance the fuel by transferring fuel from the left-wing tank to the right. However, the transferred fuel was lost through the fractured fuel line. At 05:45 UTC, after discovering that more than 14,000 lbs. of fuel were missing, the crew initiated a diversion to the Azores.

The right engine flamed out shortly thereafter at 06:13 UTC. At 06:26 UTC, when the aircraft was about 65 nautical miles from the airfield, the left engine also flamed out. The crew considered that a ditching at sea was a possibility in case the aircraft could not reach the Azores. However, assisted by radar vectors from air traffic control, the crew reached the Lajes airport after a 19-minute 75-mile unpowered glide and carried out a visual approach in good visibility and weather. The aircraft landed on runway 33 at 06:45 UTC. After the aircraft came to a stop, small fires started in the area of the left main-gear wheels, but these fires were immediately extinguished by crash rescue-response vehicles. The Captain ordered an emergency evacuation almost immediately after the plane came to a stop; 16 passengers and 2 cabin-crew members received injuries during the emergency evacuation. The aircraft suffered structural damage to the fuselage and to the main landing gear.

The investigation determined that the double-engine flameout was caused by fuel exhaustion, which was precipitated by a fuel leak that developed in the right engine as the result of the use of mismatched fuel and hydraulic lines during the installation of the hydraulic pump.

3.1.6.2 Flight crew actions

Adaptations to rules or procedures

- The Captain did not perform the FUEL LEAK – LEAK NOT FROM ENGINE procedure because this would have required him to descend the aircraft to 20,000 feet. If there was indeed a leak, he would be losing fuel anyway, and he thought that, by descending to 20,000 feet he would give up altitude and performance margin in a situation where fuel remaining was critical.
- During approach, the Captain made a series of quick turns to slow the aircraft down before landing.
- During landing, following the initial bounce, the nose of the aircraft rose significantly. Because he did not want to become airborne a second time, on the second touchdown, the Captain did not flare and applied and held maximum braking.

Knowledge-based behaviors

- Neither pilot had ever encountered a fuel leak or an unexplained low fuel quantity either in training or in flight. Because the total fuel quantity was reducing at an unexplainable high rate, the Captain decided to use up the fuel from the right tank before it was lost, and selected the right wing boost pumps ON and the left-wing boost pumps OFF to establish a crossfeed from the right wing tank to the left engine.

- Flying an unfamiliar approach to Lajes during nighttime with no engines, very limited electronics and instruments, limited pitch control, no flaps, no spoilers, and no reverse thrust. Neither pilot had ever trained on or performed a landing in these circumstances.

3.1.6.3 Factors not under pilot control

- The flight was rerouted by ATC to be about 60 miles south of original route across the Atlantic, which put the aircraft within gliding distance to Lajes.
- The aircraft departed from Toronto with an extra 12,100lbs of fuel, which put the flight within range of Lajes airport.

3.1.7 American Airlines flight 1400

3.1.7.1 Executive summary

On September 28, 2007, American Airlines flight 1400, a McDonnell Douglas DC-9-82, experienced engine problems while still on the ground at the Lambert-St. Louis International Airport (STL), Missouri. The left engine would not start. It took two attempts, with the assistance of maintenance, until they were able to start the engine manually. During departure climb, at around 1:13pm Central Daylight Time, they received an engine fire warning for the left engine. During the return to STL for an emergency landing, the nose landing gear failed to extend. The aircraft lost all power, and the pilots were unable to start the APU. They tried to lower the landing gear manually but the air traffic controller informed them that it was not extended. The crew initiated a go-around, during which the Captain asked an off-duty company pilot to come to the flight deck to assist. The Captain also requested and received permission from ATC to land on runway 30L, which was 2,000 feet longer than the runway (30R) initially assigned by ATC. The First Officer performed the Emergency Gear Extension checklist, and while the landing gear indications did not illuminate, the pilots heard a noise suggesting that the nose gear was extended. This was subsequently confirmed by ATC. After the aircraft landed, aircraft rescue and firefighting (ARFF) personnel responded to the engine fire. Originally, the Captain had not planned to evacuate the airplane and have it be towed to the terminal. However, while the ARFF was responding to the fire, fuel spilled out of the engine area and the incident commander recommended deplaning all passengers (2 flight crewmembers, 3 flight attendants, and 138 passengers) on the runway. No occupant injuries were reported, but the airplane sustained substantial damage from the fire (Figure 10).

Figure 10. American Airlines flight 1400 fire damage

The NTSB determined that "the probable cause of this accident was American Airlines maintenance personnel's use of an inappropriate manual engine-start procedure, which led to the uncommanded opening of the left engine air turbine starter valve (ATSV), and a subsequent left engine fire, which was prolonged by the flight crew's interruption of an emergency checklist to perform nonessential tasks. Contributing to the accident were deficiencies in the maintenance procedures used by American Airlines maintenance personnel that were not in accordance with the airline's written manuals and guidelines." (National Transportation Safety Board, 2009, p. 68)

3.1.7.2 Flight crew behaviors

Adaptations to rules or procedures

- Ultimately land the plane based on the sound of the nose gear extension despite no illumination of landing gear lights (which was subsequently confirmed by ATC to be down)
- Decide to land on the runway which was 2,000 feet longer than the runway initially assigned by air traffic control to ensure safe landing

Knowledge-based behaviors

- Decide to execute a go-around during the first attempt because the landing gear had not extended, and the airplane was too close to the airport to extend the nose landing gear manually. The Captain also stated that he did not want to attempt a landing without the nose gear extended without briefing the flight attendants and passengers. NTSB (2009) concluded that this was good thinking because a nose-gear-up landing would have been difficult to perform given the airplane's altitude and the lack of time to prepare to perform it.

3.1.7.3 Factor not under pilot control

- Additional Onboard Pilot: An off-duty pilot was deadheading on the plane and significantly reduced pilots' workload by helping with flight attendant and passenger communications, and troubleshooting the nose landing gear situation.

3.1.8 British Airways flight 38

3.1.8.1 Executive summary

On January 17, 2008, while on approach to London (Heathrow), the autothrottles on flight 38 (a Boeing 777-236ER) commanded an increase in power from both engines. Initially, the engines responded but at 720 feet AGL and about 57 seconds before touchdown, the thrust of the right engine reduced. Seven seconds later, the left engine also experienced a reduction in power. Both engines continued to produce thrust above flight idle, but less than the commanded thrust. About 27 seconds before touchdown, the First Officer noticed that the airspeed was decreasing to below the expected approach speed of 135 kts. The flight crew attempted to identify the cause for the loss of thrust. The engines failed to respond to demands for increased thrust even when the throttles were moved manually to full power. When the airspeed reached 115 kts, the 'airspeed low' warning was triggered, along with a master caution aural warning. The airspeed stabilized for a short period. In an attempt to reduce drag and stretch the glide, the Captain retracted the flaps from flaps 30 to flaps 25. This decision allowed the plane to miss the ILS beacon within the airport perimeter, thus avoiding more substantial damage.

In the last few seconds before impact, the Captain attempted to start the APU and when he realized that a crash was imminent, he transmitted a 'MAYDAY' call. The First Officer pulled back on the control column but the aircraft still hit the ground approximately 1,000ft (330m) short of the paved runway surface (Figure 11). The airplane was evacuated. Of the 152 people on board, 47 sustained injuries, one serious. No fatalities were recorded.

Figure 11. British Airways flight 38 landing

The investigation of the accident determined that the reduction in thrust was due to restricted fuel flow to both engines. Ice accretion within the fuel system was identified as the probable cause (Department for Transport, 2010). The flight had crossed over Mongolia, Siberia, and Scandinavia at an altitude between FL 348-400, in temperatures between −65 °C (−85 °F) and −74 °C (−101 °F). The flight crew, aware of the cold conditions outside, had monitored the temperature of the fuel, with the intention of descending to a lower altitude if any danger of the fuel freezing arose. This did not become necessary as the fuel temperature never dropped below −34 °C (−29 °F). However, while the fuel itself did not freeze, small quantities of water in the fuel did. During the final stages of the approach into Heathrow, increased fuel flow and higher temperatures released this frozen slush, which had likely adhered to the inside of the fuel lines it back into the fuel. It flowed forward until it reached the fuel-oil heat exchangers (FOHEs) where it caused a restriction in the flow of fuel to the engines.

3.1.8.2 Flight crew behaviors

Knowledge-based behaviors

- Decide to retract the flaps from FLAP 30 to FLAP 25 on final approach to reduce the drag and increase the distance to touchdown (which allowed the aircraft to clear the aerial antenna array preventing more structural damage to the aircraft)

The accident report highlights that "On the final approach to land the flight crew were presented with an operational situation, a double-engine rollback at a low height, which was unprecedented." (Department for Transport, 2010, p. 140)

3.1.8.3 Factors not under pilot control

- The emergency evacuation benefited from the lack of a post-crash fire and the fact that the cabin and aircraft structure remained largely intact.
- The loss of thrust happened about 30s before landing instead of happening earlier which would have resulted in disastrous consequences.

3.1.9 Qantas Flight 72

3.1.9.1 Executive summary

On October 7, 2008, Qantas flight 72 (an Airbus A330) departed Singapore on a scheduled passenger flight to Perth, Western Australia. While at the cruise altitude of FL370, the Air Data Inertial Reference Unit (ADIRU) 1 started providing intermittent, incorrect AOA values to other aircraft systems. Soon after, the autopilot disconnected, and the crew received numerous ECAM warnings and caution messages concerning irregularities with the autopilot and inertial reference system, as well as contradictory audible stall and overspeed warnings.

The Captain took manual control of the aircraft, very briefly re-engaged the autopilot, but then returned to manual flight for the remainder of the trip. Because the Captain was unsure of the reliability of the information on his primary flight display (PFD) (airspeed and altitude indications on his PFD were fluctuating), he decided to use the standby instruments and the First Officer's PFD while flying the aircraft.

Two minutes after the ADIRU 1 had first started providing wrong data, the aircraft suddenly pitched nose down, experiencing −0.8 g, reaching 8.4 degrees pitch down and rapidly descending 650 feet. The pilots were able to return the aircraft to the assigned cruise altitude within 20 seconds. But approximately 2 ½ minutes later, the aircraft pitched down again, this time causing an acceleration of +0.2 g, a 3.5 degree down angle, and a loss of altitude of 400 feet. Again, the flight crew was able to climb back to the aircraft's assigned level flight 16 seconds later.

The resulting forces from the pitch-down maneuvers were sufficient for unrestrained (and even some restrained) passengers and crew to be flung around the cabin or crushed by overhead luggage, as well as crashing with and through overhead compartment doors (Figure 12). In all, one crew member and 11 passengers suffered serious injuries, while eight crew and 99 passengers suffered minor injuries. The flight diverted to Learmonth, Western Australia, where medical assistance was provided to the injured passengers and crew.

Figure 12. Quantas flight 72 aftermath

The Australian Transport Safety Bureau (ATSB) investigation determined that the in-flight upset occurred due to the combination of a defect in the flight control primary computer (FCPC) software of the Airbus A330/A340, and a failure mode affecting one of the aircraft's three air data inertial reference units (ADIRUs). The software defect meant that, in a very rare and specific situation, multiple spikes in angle of attack (AOA) data from one of the ADIRUs could result in the FCPCs commanding the aircraft to pitch down.

3.1.9.2 Flight crew behaviors

Adaptations to rules or procedures

- Use the First Officer's PFD for flying after deducing that the Captain's PFD was unreliable

Knowledge-based behaviors

- Take manual control of the aircraft after autopilot 1 disconnected, instead of engaging autopilot 2. This action enabled the Captain to respond more quickly to subsequent pitch-downs.
- Conduct a series of wide left orbits to maintain the aircraft's speed below 330kts, lose altitude for landing, and prevent any potential problems associated with another unexpected pitch-down event. The Captain reported that he descended cautiously in order to prevent any potential problems associated with another unexpected pitch-down event.

3.1.10 Cathay Pacific flight 780

3.1.10.1 Executive summary

Cathay Pacific flight 780, an Airbus A330, departed Surabaya, Indonesia, for Hong Kong International Airport on April 13, 2010 at 01:20 UTC. During climb, the flight crew noticed some abnormal engine pressure ratio (EPR) fluctuations on both engines. Shortly after levelling off at FL390 and again two hours after departure, the ECAM displayed a "ENG 2 CTL SYS FAULT" message. Both times, the flight crew contacted maintenance control (MC) to discuss the situation. As all other engine parameters were normal, MC and the flight crew decided that it was safe to continue the flight to Hong Kong.

During the descent to Hong Kong, at FL230, two ECAM messages—ENG 1 CTL SYS FAULT and ENG 2 STALL—appeared. The second message indicated an engine compressor stall. The flight crew completed the required ECAM actions and set the engine No. 2 thrust lever to IDLE and the engine No. 1 thrust lever to maximum continuous thrust (MCT). The flight crew contacted Hong Kong ATC and declared a "pan-pan". They requested the shortest possible route to and priority landing at the airport. When the aircraft was approximately 45 nm southeast of the airport and about to level off at 8,000 ft, an ECAM message "ENG 1 STALL" was displayed. The crew carried out the actions for a No. 1 engine compressor stall and declared a "mayday". The Captain then performed a test of the engine responses. The No.1 engine's fan speed spooled up to about 74% N1, while the No.2 engine remained below idle speed. Combined, the two engines provided sufficient thrust to reach Hong Kong. As the flight approached the airport, the Captain tried to reduce thrust from the No.1 engine but the engine output did not respond to movement of the thrust lever and remained at 74% N1.

As a result, 11 minutes after declaring the emergency, the aircraft landed hard at a groundspeed (Figure 13) of 231 kts (95 kts above the normal touchdown speed for the aircraft). Only the No.1 engine's thrust reverser deployed, forcing the crew to bring the aircraft to a stop using manual braking. Five of the eight main tires were deflated, and there was fire and smoke coming from the landing gear. The Captain ordered an emergency evacuation of passengers during which 57 passengers and six cabin crew were injured.

Figure 13. Cathay Pacific flight 780 landing

The accident analysis revealed that the loss of thrust control on both engines was caused by contaminated fuel, which contained particles of superabsorbent polymer (SAP) introduced into the fuel system when the aircraft was fueled at Surabaya.

3.1.10.2 Flight crew behaviors

Adaptations to rules or procedures

- Stay as high as possible for as long as possible to maintain sufficient glide performance if needed.

Knowledge-based behaviors

- Very slowly introduce fuel into # 1 engine to try to restart it. This was the only method that worked after multiple failed attempts to restart both engines.
- Figure out a thrust setting (74%) that minimized engine surges and popping noises but still provided enough thrust to hold altitude and perform a landing at the airport.
- Perform a high-speed hard landing with unresponsive thrust control (engine #1 locked at 74%)

3.1.11 Qantas flight 32

3.1.11.1 Executive summary

On November 4, 2010, Qantas Flight 32, an Airbus A380, departed Changi Airport, Singapore on a scheduled passenger flight to Sydney, Australia. About 4 minutes after take-off, while climbing through 7,000 ft, the aircraft suffered an uncontained engine failure. The flight crew heard two ‘bangs’ and a number of warnings and cautions were displayed on the ECAM. Initially, the ECAM displayed a message warning of turbine overheat in the No.2 engine. That warning was followed soon after by a multitude of other messages relating to a number of aircraft system problems. The First Officer also reported observing an ECAM warning of a fire in the No. 2 engine (Figure 14) that was displayed for 1-2 seconds, before the ECAM reverted to the overheat warning. As part of the turbine overheat procedure, the crew elected to shut down the No. 2 engine. During the shutdown procedure, the ECAM displayed a message indicating that the No. 2 engine had failed.

Figure 14. Simulation of Quantas flight 32 engine fire

As the aircraft remained controllable, and there was ample fuel on board, the flight crew decided that the best option would be to enter a holding pattern to try to diagnose the problem and decide on a course of action. They were cleared to hold 30nm East of Changi Airport. There were five pilots on the flight deck, including two check airmen. This allowed the Captain to concentrate on flying and managing the aircraft while the other pilots monitored and responded to ECAM messages. The second officer also went into the cabin to assess the damage visually. He observed damage to the left wing and fuel leaking from the wing.

On completion of the ECAM procedures, which took about 50 minutes, the flight crew assessed the status of the aircraft. The First Officer and one of the check airmen entered information about the plane's status into the landing-distance performance application (LDPA) to calculate the required distance for a landing 50 tons over the maximum landing weight. The LDPA was unable to calculate a landing distance for this condition. The crew then removed the constraint of landing on a wet runway, as they knew that the runway was dry. This resulted in the LDPA indicating that landing on runway 20C at Changi airport was feasible with 330ft of runway to spare.

The flight crew progressively configured the aircraft for the approach and landing. Because of the damage to the aircraft, extending the landing gear required use of the emergency manual extension procedure. The Captain set engines No. 1 and 4 to provide symmetrical thrust and controlled the aircraft's speed with the No. 3 engine. During the approach, the autopilot disconnected twice. When the autopilot disconnected for the second time (at about 800 ft), the Captain decided to manually fly the aircraft for the remainder of the approach.

The aircraft came to a stop 150m from the end of the runway. Fuel continued to leak from the left-wing tank. To minimize the risk associated with this leak, the airport emergency services doused the engine with water and foam. Knowing the fire risk was being managed and had decreased, the crew decided that the safest course of action was to have passengers disembark via stairs on the right side of the aircraft. There were no reported injuries to the 440 passengers and 29 crew on board.

The ATSB investigation found that the engine failure was caused by shrapnel from the engine puncturing part of the wing and damaging the fuel system, causing leaks and a fuel-tank fire. One hydraulic system and the antilock braking system were disabled, and flaps and the controls for the No.1 engine were damaged. The failure was the result of an internal oil fire within the Rolls-Royce Trent 900 engine that led to the separation of the intermediate pressure turbine disc from its shaft. The ATSB found that the oil pipe cracked because it had a thin wall from a misaligned counter bore that did not conform to the design specification.

3.1.11.2 Flight crew behaviors

Adaptations to rules or procedures

- Conduct ECAM procedures and checklists for more than 50 consecutive failures (turbine overheat, degraded flight controls, emergency gear extension, evacuation, etc.). This required the flight crew to actively suppress a flood of cascading alerts and warnings (137 displayed on ECAM, 37 not displayed) to focus on aircraft subsystems that were operational or most relevant
- Decide to stop fuel transfer between wings due to uncertainty about the integrity of the fuel system. The flight crew discussed the impact of ECAM procedures before carrying

out relevant checklists. Fuel-related ECAM procedures, for example, were more closely scrutinized due to damage to fuel system (observed by the second officer from the cabin)

- Perform landing calculations using landing performance application (LPA) for an abnormal landing configuration (inoperative wing leading edge lift devices, reduced braking function, reduced number of operational spoilers and inactive left engine thrust reverser) not accounted for by the software.
- Decide to do precautionary evacuation rather than emergency evacuation because fire risk was minimized by emergency services. This action prevented potentially fatal injuries that could result from conducting an emergency evacuation using slides

Knowledge-based behaviors

- Configure outboard engines 1 and 4 to provide symmetric thrust and use inboard engine (produces less yaw effect) for fine-grained speed control. Because roll control and ailerons were compromised, maintaining appropriate yaw and heading were critical to flight safety. This configuration allowed the Captain to keep the aircraft within safe margins.
- Recognize that brake temperature after landing is much higher than normal and identify it as a major fire risk. This allowed the flight crew to communicate with ground firefighters to put foam on the landing gear and brakes.

3.1.11.3 Factors not under pilot control

- Additional Onboard Pilots: 5 pilots on the flight deck (Captain, First Officer, Second officer, a check Captain, and a training check Captain)
- Ample time and fuel to deal with failures, warnings, and ECAM procedures

3.1.12 ExpressJet 4291

3.1.12.1 Executive summary

On May 11, 2015, ExpressJet Flight 4291—an Embraer ERJ-145—took off from Houston's George Bush Intercontinental Airport on its way to San Luis Potosí, Mexico. The flight departed just prior to midnight, and the crew faced a challenging flight ahead as extreme weather covered most of the flight path. After maneuvering around a patch of inclement weather, the crew initially leveled off at 34,000 feet but contacted Air Traffic Control (ATC) to request a higher altitude in hopes of providing passengers with a smoother ride. After climbing to FL360, the crew flew around additional areas of deteriorating weather. After around 125 miles on their new flight path, Capt. VanHoose and First Officer Moser (Figure 15) began to see anomalies in their basic flight instruments. Noticing an IAS flag on the Primary Flight Display that indicated a difference in air data, they realized that they had lost both of their air data computers – meaning that the basic flight instruments were unreliable.

Figure 15. Capt. VanHoose and First Officer Moser

The Captain flew the aircraft and tasked the FO to run the appropriate troubleshooting procedures. With multiple faults presenting, the FO ran the Quick Reference Handbook procedure associated with the most likely starting point, an unreliable airspeed. Disconnecting the autopilot and disabling the flight director and yaw damper, the Captain flew the aircraft manually as the FO called out power settings and continued through the checklist. With weather conditions worsening and mountainous terrain along their route, combined with the current condition of the aircraft, the crew determined that continuing to Mexico was not an option and requested assistance from ATC.

Using all resources available, the crew was able to find acceptable conditions in San Antonio, Texas, and elected to divert. Shortly after turning back to the north and deviating around thunderstorms in the area, the aircraft encountered moderate turbulence, aircraft icing, and was struck by lightning. The flight crew determined that this did not result in any additional loss to their already minimal number of flight instruments. Throughout the descent, instruments on the Captain side at times began to regain some data and then fail again, while instruments on the FO side were void of information regarding air data. The crew coordinated with ATC and asked the controller to call out their ground speed and altitude in order to cross-reference instrument indications that were intermittent and to maintain positive control.

Once established on final, the pilot flying was able to make visual contact with runway 4, and an overweight landing was made. Relying on their manual flying skills and combined years of experience, the crew safely landed the 50-passenger jet at San Antonio International Airport shortly before 2:00 a.m. local time.

3.1.12.2 Flight crew behaviors

Adaptations to rules or procedures

- Divert to San Antonio Texas (SAT) airport based on favorable location/weather and availability of additional options regarding customs and operations capabilities at SAT instead of attempting to divert to a “weather entrenched” alternate.
- Follow procedures for unreliable airspeed as listed in QRH but additionally cross-reference ground speed and altitude with ATC to stay on track and maintain control.

3.1.13 Air France flight 066

3.1.13.1 Executive summary

On September 30, 2017, an Airbus A380 operated by Air France, was carrying out a scheduled passenger flight from Paris (France) to Los Angeles (USA). About 4 hours into the flight, the flight crew heard an explosion, immediately followed by severe vibrations. The “ENG 4 STALL” and then the “ENG 4 FAIL” messages nearly simultaneously appeared on the ECAM.

The Captain engaged the autopilot and asked the First Officer to complete the required ECAM actions. The Captain then started Air France’s facts, options, risks & benefits, decide, execution, check (FOR-DEC) method for handling an incident. He observed that, from the time of the failure and for around 1min 30s, the computed air speed (CAS) had decreased from 277kts to 258 kts while level flight at FL 370 was maintained. The Captain decided to descend to the drift-down altitude calculated by the flight management system (FL 346) in an effort to maintain a constant airspeed in level flight. However, this did not work, and so he continued descending step-by-step, through FL 360, FL 350, FL 330, FL 310, and finally leveling off at FL 290 where he was able to maintain a constant speed (290 kt) by keeping the remaining three engines in MCT. He decided to continue the descent to FL 270 in order to prevent overexerting the engines. The speed stabilized at 279 kt. After that, the crew diverted to Goose Bay airport (Canada) where they landed without any further incident.

A visual inspection of the engine found that the fan, along with the air inlet and fan case had separated in flight, leading to slight damage to the surrounding structure of the aircraft. The in-depth investigation of the accident concluded that a crack started in an area called a macro zone. It was introduced during forging of the fan hub. The crack formed 1,850 cycles into the part's 15,000-cycle life and expanded over the next 1,650 due to dwell fatigue until the part failed. The failure left only a small part of the fan hub attached to the Airbus A380's No. 4 engine, offering investigators a key early clue.

Figure 16. Air France flight 066 Airbus A380 No. 4 engine

3.1.13.2 Flight crew behaviors

Adaptations to rules or procedures

- Decided to continue descending in a stepwise fashion until constant speed could be achieved after realizing that the plane failed to achieve constant speed at the driftdown level (FL370) calculated by the flight management system (FMS).
- Decided to descend further from FL 290 to FL 270 after realizing that the three remaining engines were working at MCT to be able to maintain airspeed at that altitude.

Knowledge-based behaviors

- Select Goose Bay airport for landing versus a closer airport (Kangerlussuaq) to avoid flying near the mountains at the closer airport due to compromised engine performance.

3.1.13.3 Factor not under pilot control

- A third pilot was present in the crew rest station who helped the active flight crew perform FOR-DEC and assess the damage to the engine.

3.1.14 United Airlines flight 1175

3.1.14.1 Executive summary

On February 13, 2018, United Airlines flight 1175, a Boeing 777-222, experienced an in-flight separation of a fan blade as well as portions of the inlet and fan cowl of the No. 2 (right) engine (Figure 17) over the Pacific Ocean, enroute to Honolulu, Hawaii. Shortly before starting their descent from their cruise altitude of FL 360, the flight crew heard a loud bang, followed by violent shaking of the airplane and warnings of a compressor stall. The flight crew shut down the affected engine, declared an emergency, initiated a drift-down descent, and proceeded to Honolulu where they completed a single-engine landing without further incident. There were no injuries to the 374 passengers and crew onboard and the airplane received minor damage. On July 18, 2019, the flight crew was awarded the Superior Airmanship Award by the Airline Pilots Association for safely landing the plane.

Figure 17. United Airlines flight 1175 No. 2 (right) engine damage

The NTSB (2020) determined the probable cause of this incident to be the fracture of a fan blade due to Pratt & Whitney's continued classification of the TAI inspection process as a new and emerging technology that permitted them to continue accomplishing the inspection without having to develop a formal, defined initial and recurrent training program or an inspector certification program. The lack of training resulted in the inspector making an incorrect evaluation of an indication that resulted in a blade with a crack being returned to service where it eventually fractured.

3.1.14.2 Flight crew behaviors

Knowledge-based behaviors

- Take pictures and videos of the failed engine to assess the engine's condition, evaluate the risk, and take appropriate steps.
- Notice the vibrations on control and deduce that the debris from engine separation had struck the stabilizer (which was indeed found to be damaged). This knowledge helped the pilots carefully fly and land the airplane.

3.1.14.3 Factor not under pilot control

- There was a third pilot on the flight deck, a jump seat rider, who helped the two flying pilots by reporting engine condition from the cabin

3.1.15 Southwest 1380

3.1.15.1 Executive summary

On April 17, 2018, Southwest Airlines flight 1380, a Boeing 737-7H4, departed from LaGuardia Airport, Queens, New York, en route to Dallas Love Field. While climbing through FL320 to the assigned cruise altitude of FL380, portions of the left engine inlet and fan cowl separated from the airplane, and fragments from the inlet and fan cowl struck the left wing, the left-side fuselage, and the left horizontal stabilizer (Figure 18). One fan cowl fragment affected the left-side fuselage near a cabin window and damaged the window, which resulted in a rapid depressurization. Immediately afterward, the No. 1 (left) engine's fan and core speeds decreased, the cabin altitude warning horn sounded, and the flight crew felt significant vibrations of the airplane. This was followed by an uncommanded roll to the left, to a maximum angle of 41.3°. The First Officer, who was the pilot flying at the time, rolled the aircraft back to wings level within 11 seconds of the explosion.

NTSB_Newsroom @NTSB_Newsroom · 22h
NTSB investigators on scene examining damage to the engine of the Southwest Airlines plane.

Figure 18. Southwest 1380 airplane damage

The pilots reduced the right engine power to idle and began an emergency descent in accordance with the emergency descent checklist. During the descent, the Captain requested vectors to the closest airport. The air traffic controller suggested Harrisburg International but the Captain decided to divert to Philadelphia International Airport instead, at the recommendation of the First Officer who had looked at a map and determined that Philadelphia was a close suitable airport. As prescribed by the Southwest Airlines Flight Operations Manual in the event of an engine shutdown or failure, the Captain took control of the aircraft and the FO switched to handling communications with ATC. The flight crew reported initial communications difficulties because of the loud noises, distraction and wearing oxygen masks, but as the aircraft descended, communications improved.

At one point, the First Officer stated, “check your speed,” and the Captain stated that she was trying to slow down the airplane on purpose. The flight data recorder later showed that the airspeed had decreased from 272 to 232 knots during a 40-second period. During a post-accident interview, the Captain stated that she flew slower than the emergency descent checklist speed to reduce the severity of the airframe vibration. This action aligned with the note provided in the emergency descent checklist that “if structural integrity is in doubt, limit speed as much as possible and avoid high maneuvering loads.”

The First Officer contacted the flight attendants who informed him that a passenger had been partially sucked out of the damaged cabin window. While initially intending to perform a long final approach, the Captain decided to expedite the approach after hearing about the injured passenger. During this time, the Captain also instructed the FO to complete the Engine Fire or Engine Severe Damage or Separation checklist. The aircraft was cleared to land on runway 27L.

The Captain chose to use 5° of flaps for the landing due to concerns about controllability. The airplane landed safely at Philadelphia about 17 minutes after the engine failure occurred.

Of the 144 passengers and 5 crewmembers aboard the airplane, 1 passenger received fatal injuries, and 8 passengers received minor injuries. The airplane was substantially damaged. Investigators determined that the left engine failure occurred when one of the fan blades fractured at its root due to a low-cycle fatigue crack that initiated in the dovetail (part of the blade root). The impact of the separated fan blade with the fan case imparted significant loads into the fan cowl. The left side of the fuselage near the location of the missing cabin window (row 14) had impact damage and witness marks that were consistent with the size and shape of the inboard fan cowl aft latch keeper and surrounding structure.

3.1.15.2 Flight crew behaviors

Adaptations to rules or procedures

- Conduct emergency descent at a speed below the recommended emergency descent speed (V_{MO}) due to airplane vibrations.
- Expedite approach upon receiving new information about a passenger being sucked out of a window. During a post-accident interview, the Captain stated that she had initially requested a long final approach to allow time to accomplish checklists but then decided to expedite the approach due to the passenger injury.
- Decided not to carry out a step in the emergency descent checklist that called for the use of speed brakes for a faster descent in order to prevent additional stresses on the airframe.
- Choose to land with 5° flaps configuration rather than a flaps 15 landing configuration specified by the One Engine Inoperative checklist due to concerns about controllability of the aircraft. During a post-accident interview, the Captain reported that she was experiencing “lots of drag” on the flight controls during the descent. The Captain also reported that to determine the approach speed with 5° of flaps, she considered the 160-knot airspeed for flaps 15 (the recommended landing flap configuration for a single-engine landing, according to the B737NG Aircraft Operating Manual) and added 20 knots to attain an approach airspeed of 180 knots (the SWA 737NG QRH did not provide guidance for a single-engine landing with a flap setting of 5°.)

3.2 Summary of flight crew contributions to safety and risk mitigation

In the above accidents, flight crew behaviors that went beyond proficient skill-based and rule-based performance prevented an adverse event or failure from resulting in a worst-case outcome.

These behaviors fall into two main categories: (1) adapting existing rules or procedures to account for operational/contextual variations on an adverse event and (2) knowledge-based performance, in the sense of developing solutions to a novel problem in real time when no known rules or remedies exist.

3.2.1 Adapt existing rules to account for operational/contextual variations

In some of the accidents, the failure/adverse event was not entirely new and unexpected. Procedures and checklists existed to address the problem which had been experienced before (often in training/ simulation) but in a different context and/or in a different form (e.g., single- versus dual-engine failure). In these cases, pilots 'saved the day' by realizing the need to vary or adapt the solution to preempt inefficiencies or potentially harmful consequences of following the available rule or procedure. Examples of such scenarios include:

- Air Transat 236: Choose to land with 5° flaps configuration due to concerns about controllability of the aircraft, rather than the flaps 15 landing configuration specified by the One Engine Inoperative checklist.
- Qantas 32: Decide to stop fuel transfer between wings due to uncertainty about status of fuel system and to reduce the possibility of running out of fuel. This was based on observations of the second officer of a fuel leak in the left wing.
- Air France 066: Decide to perform a step-down descent to a lower altitude after realizing that the plane failed to achieve a constant airspeed at the drift-down level calculated by the FMS.
- Aloha Airlines Flight 243: Decide to use flaps 5 and IAS 170 knots for approach and landing because the aircraft became less controllable with higher flaps settings and speeds below 170 knots. This was based on the Captain’s observation that the airplane was "shaking a little, rocking slightly and felt springy”.

3.2.2 Develop solutions to novel problems and events in real time—knowledge-based performance

In cases where pilots encountered a novel unanticipated problem for which no solutions in the form of if-then rules were available, airmanship and their mental model of the system/aircraft enabled them to develop a (partial) solution in real time. In such scenarios, goals were explicitly formulated by the pilots. In some cases, several plans were developed and tested against the goal physically (by trial and error) or conceptually (by mental simulation of the considered plan) before selecting a final plan. Following are some of the examples from the accident analysis that demonstrate such behaviors:

- Scandinavian Airlines Flight 751: Captain uses trees as a “pillow” to reduce the plane’s sink rate before landing.
- United Airlines Flight 232: Use differential thrust on remaining two engines to control heading and pitch.
- British Airways Flight 38: Decide to retract the flaps from FLAP 30 to FLAP 25 during approach to increase the distance to touchdown, which allowed the aircraft to clear the ILS aerial antenna array preventing more structural damage to the plane.
- Cathay Pacific Flight 780: Slowly introduce fuel into engine to try to restart it again after multiple failed attempts on both engines.
- TACA Flight 110: Perform a sideslip maneuver to quickly reduce air speed and align with the levee – a maneuver not normally used on large commercial airliners; the outcome of performing such a maneuver on a Boeing 737 was uncertain.

4 Adaptation of the SRK framework

Based on the findings from the literature review and the accident analysis, we updated the original SRK framework to account for pilot contributions to safety and risk mitigation that go beyond the traditional descriptions of the three performance levels. In addition to highlighting the importance of knowledge-based performance, illustrated in several of the accidents, the revised SRK model, modified from Chauvin et al. (2013), (Figure 19), now points out the need for flight crews to adapt, at the skill-based level, to considerable variability in day-to-day operations (e.g., weather, ATC), as emphasized by Safety-II, and the need to modify existing rules or procedures to account for contextual variations of an adverse event, a behavior that 'saved the day' in many of the 14 cases. The updated framework also includes references to the related concept of, and steps involved in recognition-primed decision (RPD) making (Klein, 2008). RPD is relevant to this effort as a model of how experts make quick, effective decisions when faced with complex situations. It highlights that experts engage in pattern matching to determine whether the problem they face resembles a prototypical situation they have encountered before. If a match is found, reasonable actions that were successful in the past are quickly identified and taken, a process similar to applying known if-then rules at the rule-based performance level in SRK. If no matching situation can be identified, the operator is forced to shift to the more effortful process of developing a solution in real-time, partly through mental simulation, similar to knowledge-based performance.

At the rule-based performance level, pilots contribute to flight safety in ways that go beyond the mere execution of checklists and procedures in response to adverse events. Instead, they adjust and adapt prescribed actions based on the unfolding of an event, and the context in which it takes place. This requires that they first *recognize* the deviation from the typical event or pattern. According to RPD, experts are able to do this by matching the situation at hand with prototypical situations they have encountered before. Compared to novice problem solvers who have a limited repertoire of experiences, experts are particularly adept at noticing early on the cues or events that violate expectations. They recognize a basic pattern, minor variations, and identify and carry out a course of action without having to generate and analyze a full set of options.

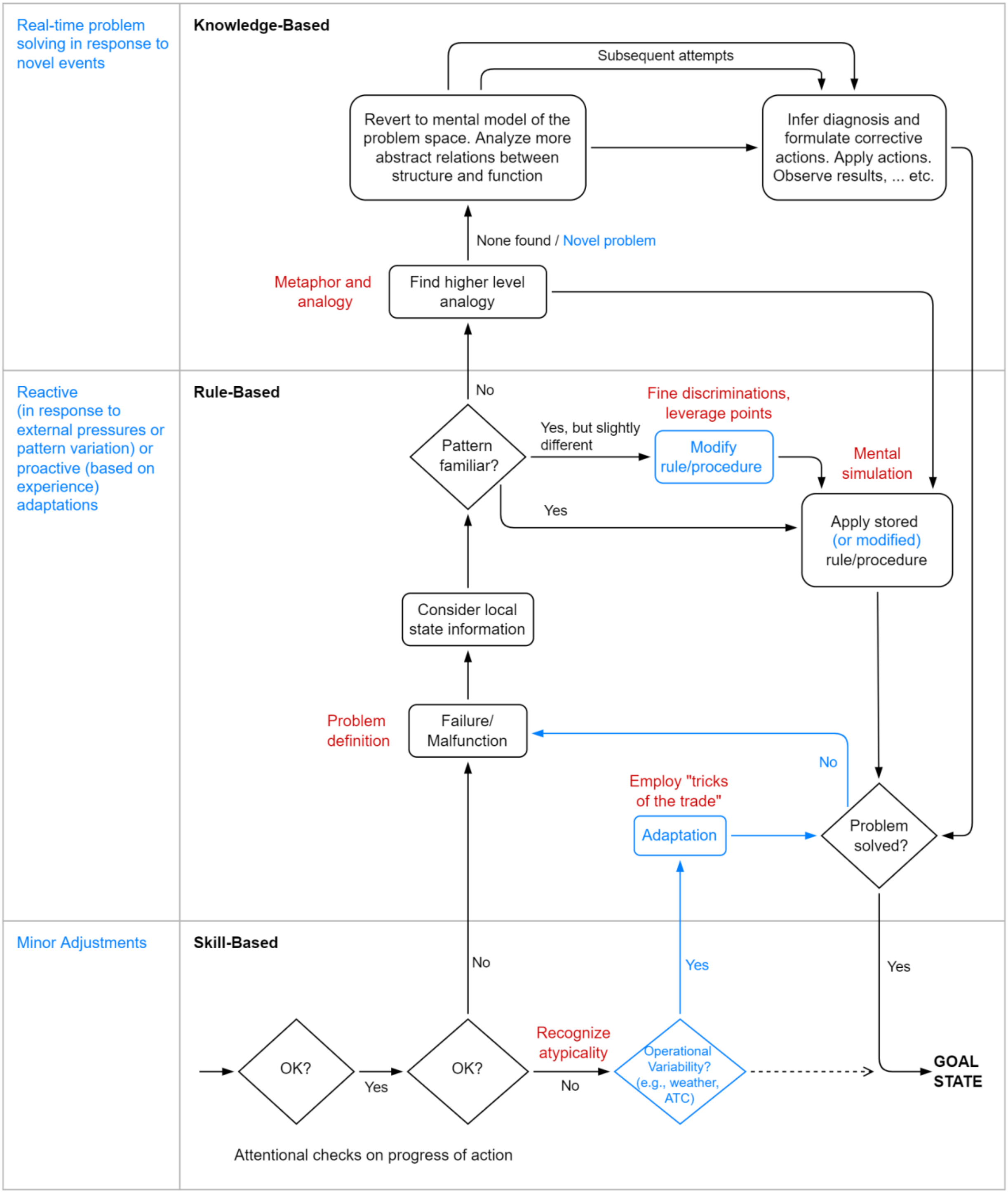


Figure 19. Revised SRK model
Changes to the SRK model shown in blue. Red text shows related concepts from the RPD model

Adaptations (Figure 20) are needed in response to operational variability, such as severe or unusual weather events or challenging air-traffic control situations and instructions. For example, Stewart et al. (2018) report that only 12% of flights fully adhere to the lateral and vertical profiles prescribed by area navigation standard terminal arrival route (RNAV STAR) procedures. Holbrook et al. (2020) highlight that more than 40% of cases of non-adherence to RNAV STAR procedures reported to ASRS were *intentional* deviations from the course, altitude, and speed profiles. Many of these deviations were initiated by ATC or the flight crew to account for factors such as unexpected aircraft performance, automated systems issues, traffic conflicts, runway changes, and flaws inherent in RNAV procedure design. According to observations and interviews conducted by the American Airlines' Department of Flight Safety (2021), 42% of flights experience factors such as weather, ATC, or aircraft mechanical/automated systems issues that require adaptations of standard rules or procedures. These adjustments are thus not rare, one-off occurrences, but rather a relatively frequent part of flight operations that pilots need to respond to in the interest of maintaining safe operations.

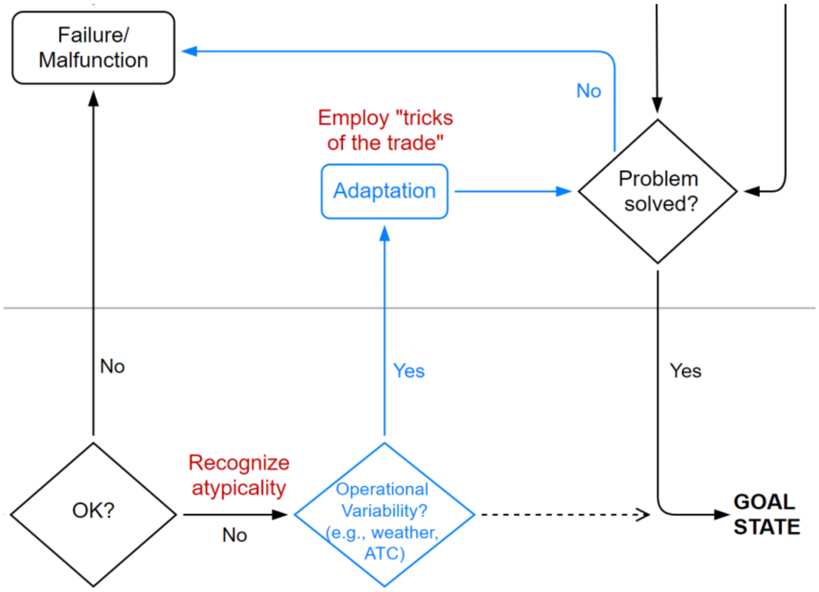


Figure 20. Adaptions in response to operational variability

The Safety-II perspective pioneered by Hollnagel (2014) recognizes that these adaptations happen as a part of everyday operations and contribute to increased safety. Adopting the Safety-II perspective requires that existing frameworks, methods, and taxonomies used to observe/discuss pilot behaviors and analyze accidents evolve to focus more on positive outcomes and contributions by flight crews (Flight Safety Foundation, 2021). The adapted SRK model highlights this shift by including adaptations at the rule-based level of performance. These adaptations can be reactive—in that they are triggered by external stimuli or conditions (e.g., unexpected weather events, ATC requests)—or proactive, meaning that steps are taken by the flights crew in anticipation of a potential problem or off-nominal condition. The latter is driven by the pilots' experience and mental model of the aircraft and the airspace system. Kiernan and colleagues (2020) provide a number of examples where pilots' past experiences and knowledge allow them to pre-plan and prepare for potential upsets during flight:

> *"Once we got up with the Washington Center frequency that was starting to do the traffic delays we had a plan in place so we knew once we got into holding we'd already calculated that we could hold for about 20-25 minutes before we'd have to go to our alternate." (Kiernan, Cross, & Scharf, 2020, p. 37)*

> *"You know it was August in Miami. So, you always have to be aware of the potential for the airfield getting soft in the thunderstorms. Typically, in Florida they move through fairly quickly and we do have holding fuel for that contingency." (Kiernan, Cross, & Scharf, 2020, p. 38)*

In some cases, adaptations are needed not in response to operational variability, but to cope with failures or malfunctions that occur in unusual circumstances or involve unique aspects. For example, on Air France flight 066, which experienced an uncontained engine failure, the flight crew performed a step-down descent to a flight level that was different from that suggested by the FMS. This was necessary and appropriate because the FMS did not consider the additional drag introduced by the severely damaged engine. The step of modifying or fine-tuning an existing rule or procedure before applying it was added to the adapted SRK model. It highlights that although a problem may not be altogether novel or unexpected (which would require a transition to knowledge-based performance); it may differ in some aspects from previously experienced cases and thus require adjustments and corrections rather than rote execution of standard procedures or checklists. According to the RPD model, experts are able to make fine discriminations and detect small but potentially impactful differences between situations based on their large repertoire of experiences and then use leverage points as "fruitful starting points in the construction of novel courses of action" (Klein & Wolf, 1998).

5 Additional factors contributing to safety

In addition to safety contributions made by pilots at the skill-, rule-, and knowledge-based performance levels, several other factors seem to have played a role in bringing about the rather successful outcome of the 14 accidents. These include the presence and involvement of additional pilots on the flight, the excellent communication and coordination skills of the flight crew, the attitude and persistence of pilots, the military aviation background of some crew members, and favorable conditions at the time of the event (e.g., weather or the altitude at which the adverse event occurred).

Half of the fourteen accidents highlight that even two pilots may not always be able to handle the workload imposed by an adverse event and/or develop a solution to an unknown problem. In these cases, the involvement of additional pilots who happened to be on the flight (check airmen, pilots deadheading/commuting, and pilots on board as passengers) was required. For example, on American Airlines Flight 1400, the Captain asked an off-duty pilot to come to the flight deck to help him and the First Officer cope with their workload. The off-duty pilot took care of the communication with flight attendants and passengers, helped troubleshoot the nose landing gear situation, and confirmed that all hydraulic pressure was lost. On Air France Flight 066, the pilots asked the First Officer who was in the crew rest station to come to the flight deck to help perform the FOR-DEC technique for processing the engine failure and assess the damage to the engine by visually inspecting it from the upper deck of the airplane. On United Airlines Flight 232, the presence of an additional highly experienced pilot (Captain Denny Fitch) who was able to develop and execute a plan for controlling the flight path using the differential engine thrust likely prevented the situation from degrading further and increased everyone’s chance of survival. Captain Per Holmberg provided critical assistance to the pilots on Scandinavian Airlines Flight 751 by running checklists and providing directional guidance. Qantas Airlines flight 32 benefitted greatly from the contributions made by three additional highly experienced pilots who were on the flight deck in their role as check airmen and who contributed to the decision making process and help manage workload.

In the majority of accidents, the flight crew demonstrated excellent communication and coordination skills. The pilots avoided panicking which allowed them to jointly discuss and diagnose the problem(s) at hand in a timely and rational manner and take the appropriate actions to deal with them. For instance, the flight crews on Qantas 32 and United 232 faced an unprecedented number and severity of warnings and failures but remained calm and demonstrated excellent joint problem solving. Captain Denny Fitch, the United 232 pilot, mentioned in a post-accident interview that the flight crew did not hesitate to provide him full authority over throttles once they had agreed on how to proceed and that there was a “complete trust” between the flight crew ((Director) Sinyi & (Writer) Kazazian , 2012, April 13). Moreover, on crew coordination in United 232, the Captain Alfred Haynes reported:

> *“If we had not worked together, with everybody coming up with ideas and discussing what we should do next and how we were going to do it, I do not think we would have made it to Sioux City”* (Haynes, 1991, p. 5)

Another characteristic of many of the pilots involved in the 14 accidents is a positive, ‘can-do’ attitude and a persistent resolve to rescue the situation. For example, in a post-accident interview, Denny Fitch of United 232 described his and the flight crew’s attitude as follows:

> *“You will get this done, you will do it. I will not accept failure, I will not accept anything less the best. Even if that’s the way I die” ((Director) Sinyi & (Writer) Kazazian , 2012, April 13).*

Similar optimism and confidence were demonstrated by the Captain of Cathay Pacific Flight 780. The flight crew had initially planned to ditch the aircraft in the South China Sea, (understanding that this could be fatal). However, the Captain decided to try to restart the engine one more time (after multiple failed attempts) using very slow and incremental throttle inputs, which resulted in the left engine coming back online and allowed the pilots to land at the destination airport. Such qualities of “realistic optimism” and “cheerful confidence” have been previously attributed to perseverance through extreme challenges and system breakdowns in multiple aviation accidents (Reason, 2008).

Many pilots in the above accidents had some form of military background or related experience. For example, Captain Schultz of Southwest 1380 served as a Naval Aviation Instructor, Captain Haynes of UA232 served four years in the Marine Corps, and Captain Crespigny of Qantas 32 and First Officer Hayhoe of Cathay Pacific 780 both served with the Royal Australian Air Force. It is possible that military training and experience enabled these pilots to handle a higher workload and remain calm in the face of danger. However, more accidents would need to be analyzed in order to establish military experience as a contributor to aviation safety.

Finally, favorable environmental conditions at the time of the event helped prevent a worse outcome in some cases. These include weather and visibility (e.g., unusually low winds on the day of the Aloha Airlines accident) and a favorable time and place of occurrence (e.g., daytime rather than night; at high altitude or close to an airport). Discussing United Flight 232, Captain Alfred Haynes (1991) recognized luck as “factor number one” (including excellent weather conditions and flat terrain), above communications, preparation, execution and cooperation, that contributed to the degree of success of landing in Sioux city.

5.1 Comparison of similar adverse events with different outcomes

The preceding section discussed 14 accidents where the outcome of the event was better than might have been expected, thanks to positive contributions to safety and risk mitigation made by the flight crew. In this section, five accidents will be compared, which involve a very similar type of adverse event (dual engine failure/flame-out), but resulted in very different outcomes. Possible reasons, such as positive and negative flight crew actions/behaviors and contextual factors, will be discussed.

The first three accidents in Table 3 below (highlighted in green), TACA Airlines Flight 110 (B 737-300), Air Transat Flight 236 (Airbus A330) and SAS Flight 751 (MD 81), were described in detail in section 3.1. These three accidents resulted in a significantly better outcome than the two accidents described below, Tuninter Flight 1153 (ATR 72) and TransAsia Airways 235 (ATR 72):

Tuninter Flight 1153 was a scheduled passenger flight from Bari, Italy, to Dierba, Tunisia. The accident happened on August 6, 2005, when the aircraft, an ATR 72, lost both engines due to fuel exhaustion, which was the result of the installation of inappropriate fuel quantity indicators that were designed for the smaller ATR 42. The flight crew did not detect the fuel exhaustion because the incorrectly installed ATR 42 gauge indicated an adequate amount of fuel in the tanks; even after all usable fuel had been consumed. The aircraft's right engine failed at 23,000 feet, and 100 seconds later, the left engine failed, at 21,900 feet. After the engine failures, the Captain requested an emergency landing in Palermo, Sicily. The crew tried repeatedly but unsuccessfully to restart the engines as they navigated to Palermo. The aircraft glided for 16 minutes but was unable to reach the runway. It was forced to ditch into the sea, 23 nm northeast of the Palermo International Airport, at a speed of 145 miles per hour. The aircraft broke into three sections upon impact (Figure 21). Sixteen of the 39 people on board lost their lives.

Figure 21. Tuninter flight 1153 crash

TransAsia Airways Flight 235 was a domestic scheduled passenger flight from Taipei to Kinmen, Taiwan. The accident happened on February 4, 2015. Shortly after take-off, a fault in the autofeather unit of the #2 engine caused the automatic take-off power control system to autofeather that engine. The flight crew misdiagnosed the problem and reported an engine flameout. They then shut down the still-functioning #1 engine. The aircraft climbed to a maximum height of 1,510 ft, then descended. The aircraft, an ATR 72-600, banked sharply left, clipped a taxi travelling west on the Huandong Viaduct and then the viaduct itself with its left wing before it crashed into the Keelung River (Figure 22). Of the 53 passengers and five crew on board, only 15 people survived.

Figure 22. TransAsia Airways flight 235 crash

Table 3 provides information on all five flights, including:

- the altitude at which the dual engine failure/flame-out occurred
- the number of pilots involved in handling the emergency
- whether any airports were within range when the adverse event took place
- crew communication
- ATC communication
- total flight hours for flight crew (Captain | FO)
- whether or not pilots experienced panic and confusion
- the number of fatalities/injuries

Table 3 highlights that the three accidents with a more positive outcome share aspects that distinguish them from the less successful cases. Most notably, crew and ATC communication were handled well in all three cases (but not on Tuninter Flight 1153 and TransAsia Flight 235). This confirms our observation of excellent communication and coordination in the 14 accidents that were analyzed in Section 3. In contrast, in the case of Tuninter Flight 1153, the air traffic controller was not fully proficient in English; as a result, requests had to be repeated and time was lost. The pilots on TransAsia 235 delayed their mayday call until about 1 minute before the crash. This did not necessarily affect the outcome of the event but, in combination with poor flight crew communication and coordination in this case, suggests that the pilots were struggling.

Table 3. Comparison of five aircraft accidents involving a dual engine failure/flame-out in different contexts
(green = relatively positive outcome; red = catastrophic outcome)

Flight	**Failure**	**Altitude**	**# Of pilots**	**Airport(s) in range?**	**Crew comm.**	**ATC comm.**	**Total flight hours**	**Panic and confusion**	**Fatalities/Injuries**
TACA Airlines Flight 110 (B 737-300)	Dual engine flame-out (water ingestion when flying through thunderstorms)	16,500ft, during descent	3	No	Good	Good	13,410 | 12,000	No (based on interview)	1 minor injury out of 45 on board
Air Transat Flight 236 (A330)	Dual engine flame-out (fuel leak/exhaustion)	39,000 ft and 33,000ft	2 (one experienced glider pilot)	Yes	Good	Good	16,800 | 4,800	No	16 minor, 2 serious injuries out of 306 on board
SAS 751 (MD 81)	Dual engine failure (ice ingestion) after takeoff	~3000 ft	3	No	Good	Good	8,020 | 3,015	No	8 serious, 84 minor injuries out of 129 on board
Tuninter Flight 1153 (ATR 72)	Dual engine flameout due to fuel exhaustion (wrong gauges installed)	23,000 ft and 22,000ft	2	Yes	Good	Poor (Requests had to be repeated because ATC could not understand English well)	7,182 | 2,431	Yes. One pilot started praying after engines failed.	16 fatalities out of 39 on board
TransAsia Airways 235 (ATR 72)	Engine #2 failure 37 seconds after takeoff; engine #1 shut down mistakenly	~1500 ft	2, plus one observer	No	Insufficient (e.g., taking action without cross-checking)	Delayed (Mayday call was made ~1min before crash)	4,914 | 6,922	Confusion	43 fatalities out of 58 on board

Among the three accidents with a relatively successful ending, SAS flight 751 had a worse outcome than the other two cases, which may be explained, in part, by the fact that this event occurred at a much lower altitude. The other mishap that occurred shortly after takeoff, with much more catastrophic consequences—TransAsia 235—, and also the second accident at high altitude that resulted in a large number of fatalities, Tuninter flight 1153, suggest another important factor, namely that the pilots involved failed to stay calm in the face of danger. They showed signs of confusion or panic which may have been related to experiencing startle and surprise. The following sections will define both terms, discuss their performance effects, and describe mitigation strategies through training and design.

6 Startle and surprise

The Navy SEALs use the maxim "calm is contagious" to encourage officers to remain composed in the middle of chaos, uncertainty, and adverse conditions. A calm and composed demeanor instills confidence in other members of the team and fosters an environment that invites a focused diagnosis of the problem at hand. Aviation accident investigators have attributed the opposite experience—startle and surprise—as likely contributors to the negative outcome of recent aircraft accidents such as Air France Flight 447 and Colgan Air Flight 3407. In the case of AF 447, for example, physiological and psychological effects from startle following the sudden disconnect of the autopilot due to unreliable airspeed data may have resulted in the pilot making rapid and high-amplitude roll control inputs, as well as a nose-up input that increased the pitch attitude up to 11 degrees in 10 seconds. This ultimately resulted in a stall from which the crew was unable to recover.

Startle has been defined as "a sudden exposure to intense stimulation that generates an involuntary physiological reflex, similar to a flight/fight reaction, with an emotional response component" (Blumenthal, et al., 2005; Bradley, Cuthbert, & Lang, 1993; Koch, 1999). While this definition focuses on startle as a response to high-intensity external stimuli, other definitions consider startle to be the result of a violation of a pilot's expectations (Federal Aviation Administration, 2017). The latter definition overlaps with that of surprise which is "a cognitive-emotional response to something unexpected, which results from a mismatch between one's mental expectations and perceptions of one's environment" (Rivera, Talone, Boesser, Jentsch, & Yeh, 2014). Contrary to startle, a surprise reaction can be triggered by the presence or absence of an expected signal/event; it does not require the stimulus to be of high intensity. In this section, we discuss how startle and surprise events affect human performance and some ways in which these effects can be mitigated.

Startle in response to a sudden and intense stimulus triggers involuntary physiological and dexterous impairments, such as the tightening of muscles, increased blood pressure and heart rate, and brief disorientation (May & Rice, 1971). Disruption of motor response from a startle reflex generally lasts for a very brief time only (between ~0.1 – 0.3s) but can, in some circumstances, last up to 1.5 seconds (Sternbach, 1960; Thackray, 1965).

Startle also triggers cognitive impairments and affects human information processing in several ways, with effects lasting for up to 30, and sometimes 60, seconds following the event. Most notably, perceptual and attentional narrowing (also known as attentional tunneling or cognitive tunneling) are experienced where attentional resources are directed (almost) exclusively towards the startling stimulus, which, operationally, implies a tendency to reduce the number of environmental cues that are sampled for tasks and procedures peripheral to the problem source (Hilscher, Breiter, & Kochan, 2005; Lynn, 1966). Another effect of startle events is that they can lead to poor judgment of the passage of time and thus potentially to not applying appropriate control inputs at the right time or for the required duration (McKenney, 2010). Martin et al. (2016), for example, found in a simulator study that startle resulted in a significant delay in executing go-around decisions, regardless of pilot experience (measured by total flight hours) or age. In a scenario with a hand-flown ILS approach that required a missed approach on reaching the decision altitude, their data show that pilots were, on average, about five seconds slower in starting the missed approach following startle. While it is not clear from this research what differentiates those pilots who perform well from those who struggle to recover from startle events, other research outside of the aviation domain suggests that the magnitude of a startle reflex is positively correlated with people’s inherent physiological reactivity to startle, stress levels, and state anxiety levels (Poli & Angrilli, 2015; Thackray, 1988).

Surprising events result in effects that are similar to those induced by startle. Both startle and surprise result in elevated heart rate and blood pressure, and a narrowed attentional focus on the source of the surprise event. Unlike startle, however, surprise also often leads to confusion, impairment of working memory, and an inability to remember operating procedures. The duration of the surprise response is typically longer than that of the startle reflex. Like startle, the narrowing of attention following a surprise event increases the operator’s ability to focus, which can help discover relevant cues and evaluate the situation, at the cost of an increased likelihood of overlooking or forgetting other cues and procedures that may be relevant to current operation but not the specific problem encountered.

6.1 Mitigating the effects of startle and surprise

Domain expertise (the level of declarative, procedural, and structural knowledge a person has about a subject) and, even more so, judgement skills (a person's decision-making skills, metacognitive skills, cognitive flexibility, and adaptive expertise(Kochan, 2005)) affect people's ability to handle unexpected events. Both procedures and design interventions have been developed to help operators develop these abilities and cope with startle and surprise.

6.1.1 Mitigation through procedures

Research suggests that the use of mnemonics is particularly effective in helping pilots make better decisions and improve their performance following surprising and startling situations (Landman, et al., 2020). One example of this strategy is the anomalous event management framework IHTAR developed by Nutter & Anthony (2012). IHTAR comprises three "waypoints" that lead to a successful solution of an anomalous event or situation: 1) IHTAR (Upon noticing a problem, the Captain announces "I have the aircraft and radios. You have got everything else."), 2) HITSI (after stabilizing the aircraft, the Captain summarizes his/her view of the situation "Here is the way I see it"), and 3) WAYFI (the Captain allows time for the First Officer to diagnose the problem and inquires, "What are you finding?"). Step 1 ensures that the first priority is to maintain the aircraft in a safe operating state and disambiguates role assignments through specific communication. Step 2 establishes a common understanding of the problem and seeks different perspectives on the issue. Lastly, step 3 continues the communication process by seeking an independent assessment by the other pilot.

More recently, (Landman, et al., 2020) developed and tested the calm down, observe, outline, lead (COOL) strategy that encourages pilots to take a moment to relax, observe the situation, analyze the problem, and select a course of action. They found that the use of the COOL mnemonic improved decision-making performance in simulated off-nominal events such as mass shift (loose cargo shifting towards the tail of the aircraft) and flap asymmetry (left flap remains up when selecting flaps 25). However, the authors note that some pilots found the procedure to be too elaborate to execute in actual startle situations and suggested that it should be simplified by letting only the PM perform the *Observe* step in a two-pilot crew. Similar techniques include TPA which focuses on managing **t**ime, **p**ower, and **a**ttitude, respectively (Gillen, 2016); BAD (Breathe-Analyze-Decide) (Martin W. , 2017), and URP (Unload-Roll-Power) by Field et al. (2018).

6.1.2 Mitigation through display design

In addition to procedures, effective display design can aid diagnosis and information processing following an anomalous event. This may include careful filtering and placement of information for the pilot to prevent data overload and assist noticing and location critical data, proper assignment of information to/across various modalities (vision, audition, touch) to support simultaneous processing, and making pilots aware of their attentional narrowing through display prompts (Moacdieh & Sarter, 2017; Prinet, Mize, & Sarter, 2016). In particular, command displays with integrated status information have been shown to reduce the frequency of inaccurate intuitive responses, improve handling of an unexpected event, and reduce time to initial recovery from upset conditions (McGuirl & Sarter, 2006; Sarter & Schroeder, 2001; Wickens, Small, Andre, Bagnall, & Brenaman, 2008) . Since command displays eliminate the need to recall and activate relevant rules and commands, they increase the attentional resources available for practitioners to operate at the knowledge-based level in case of novel unexpected events.

7 Contributions of the second pilot

This report has so far discussed contributions of pilots to safety and risk mitigation irrespective of crew complement. In this section, we focus specifically on contributions made by the currently required second pilot on the flight deck. Identifying these contributions is important in light of two concepts for reduced crew sizes that are being proposed: (1) Reduced Crew Operations (RCO) and (2) Single Pilot Operations (SPO) (Bailey, Kramer, Kennedy, Stephens, & Etherington, 2017). RCO refers to the case where two human pilots are on-board the aircraft but during the cruise phase of flight, only one pilot is actively engaged in flying the airplane. The other pilot is resting or napping. SPO, in contrast, assumes that there is only one pilot on board who serves as the Captain and PIC, making all decisions and performing actions pertaining to command of the flight. If needed, a ground operator may provide support during high-workload conditions, such as approach and landing operations.

Operational experience and a series of empirical studies of RCO and SPO suggest that there are various reasons and circumstances for preserving the presence of a second onboard pilot (ALPA, 2019; Bailey, Kramer, Kennedy, Stephens, & Etherington, 2017; Vu, Lachter, Battiste, & Strybel, 2018). These include workload sharing/balancing, communication, complementarity, emergency handling, scanning/monitoring, as well as the impairment and/or incapacitation of a pilot. The following sections will discuss each of these motives in some detail.

7.1 Workload sharing/balancing

In RCO and SPO, the absence of a second pilot on the flight deck will likely increase the workload for the PIC who needs to fly the aircraft while performing other important duties, such as running checklists. Empirical studies have shown, for example, that in both types of operations, it takes the PIC longer to start the correct checklist following an adverse event, compared to two-pilot crew operations (Etherington, Kramer, Bailey, Kennedy, & Stephens, 2016; Etherington, Collins, Kramer, Bailey, & Kennedy, 2017). In addition, the need for the PIC to handle ATC communication during an emergency was shown to significantly disrupt checklist flow in SPO.

A pilot-in-the-loop high-fidelity motion simulation study to quantify pilot contributions to flight safety during normal flight and in response to aircraft system failures (Bailey, Kramer, Kennedy, Stephens, & Etherington, 2017) showed that, in RCO, the post-rest debriefing is critical to bring the resting pilot up to speed quickly when they rejoin the flight during an ongoing emergency. Failure to mention important events or pieces of information can lead to wrong assumptions/actions. For example, Bailey et al. also noted in one particular case, the failure of the PF to mention which checklist items had already been completed resulted in the rejoining pilot to erroneously assume completion of remaining items and thus led to a breakdown in workload sharing.

Bailey et al. (2017) also revealed that, in nominal conditions, the workload ratings for Captains and First Officers differed (Figure 23). Captains reported significantly lower workload in RCO conditions than in the traditional two-pilot crew configuration and in SPO, the latter may have been an artifact of the study design, however. For First Officers, workload in RCO was also significantly lower than in the traditional two-crew case but workload in SPO was significantly higher than in RCO.

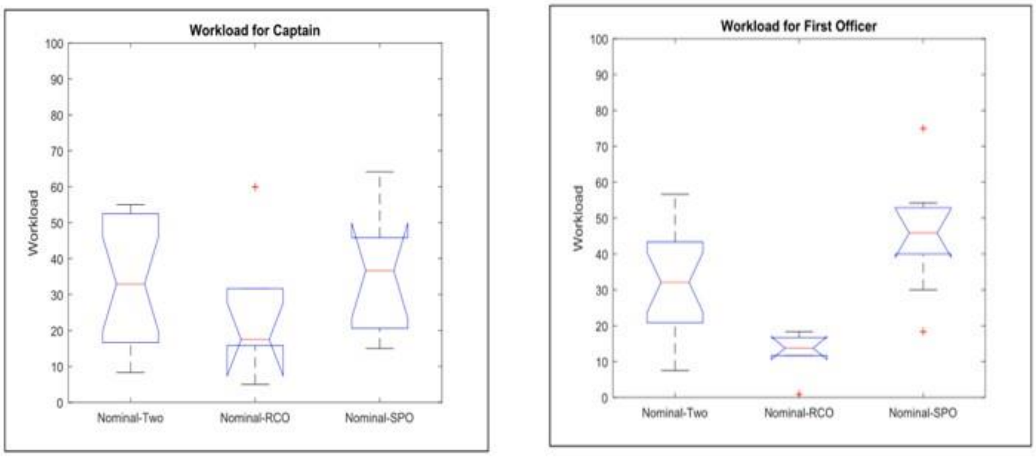


Fig 1. CP TLX Workload Rating by Crew Configurations – Nominal Flight

Fig. 2. FO TLX Workload Rating by Crew Configurations – Nominal

Figure 23. Workload ratings for the Captain and FO for nominal, RCO, and SPO operations

During non-normal events, the workload ratings for the Captain and First Officer were very similar (Figure 24, (Bailey, Kramer, Kennedy, Stephens, & Etherington, 2017)). There was a statistically significant increase in workload for the SPO condition, compared to the nominal two-crew and RCO configurations (which did not statistically differ from each other).

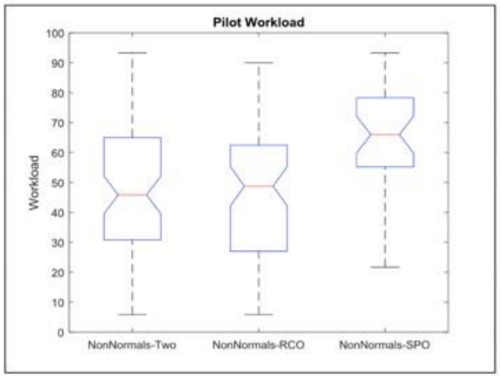


Figure 3: Pilot Workload for Non-Normal Events

Figure 24. Pilot workload during non-normal events for nominal, RCO, and SPO operations

7.2 Communication

Communication and coordination are critical on the flight deck. Two pilots sitting next to each other can coordinate their actions via voice communication and by using non-verbal cues. This was important, for example, in the case of Aloha Airlines Flight 243 where, due to the ambient noise on the flight deck, the pilots could not hear each other and used hand signals to communicate during the initial part of the descent. Non-verbal cues, such as body language, posture, and silence, can serve as cues indicating to the second pilot the cognitive and emotional state of their colleague. The co-presence of two pilots also avoids delays that can occur in SPO if a single pilot needs to communicate with a ground-based operator who may be responsible for multiple aircraft.

7.3 Complementarity

A second pilot on the flight deck can contribute to safety by complementing the skill set and experience of their colleague. For example, if one pilot lacks knowledge or experience with a particular task, procedure, or equipment, the second pilot may be able to adapt their own monitoring strategy to compensate for the deficiency:

> *"…from a CRM standpoint, the Captain told me he needed me to help him by backing him up because he was not as current in the airplane as I was at that time" (Young, 2020, p. 143).*

Even though, in this example, one pilot was relatively unfamiliar with the aircraft, the flight crew as a whole was able to compensate by changing how closely one pilot monitors the actions of the other.

7.4 Emergency handling

Handling an emergency can overwhelm a single pilot because a multitude of faults can occur in very short order and require performance of numerous checklists while continuing to fly the airplane and communicate with ATC, crew and passengers. A division of labor between two pilots is essential so that one pilot can focus on safe flight path management while the other pilot is assigned all non-flying responsibilities.

7.5 Scanning/monitoring

A recent review of the literature on visual scanning on modern flight decks by Sarter & Thomas (Sarter & Thomas, 2022) revealed how a second pilot can assist and complement their colleague across flight phases and in off-nominal conditions. Specifically, empirical studies using eye tracking have shown that:

- On takeoff, the PF and PM complement one another.

  Specifically, on takeoff in VFR conditions, the PF spends more time looking out the window, while the PM allocates more attention to the airspeed indicator and the ECAM.

- On approach, PF and PM show similar scanning strategies.

  Both pilots spend the majority of their time looking out the window, followed by fixating the PFD. The PM distributes their attention more widely, between a larger number of interfaces.

- During go-arounds, PF and PM complement each other.

  The PF primarily monitors the PFD (mostly attitude indicator), while the PM's attention is spread more broadly on interfaces related to configuration management, MCP/FCU, airspeed, altitude, FMAs and vertical speed.

Studies have also shown that, in the absence of a PM, i.e., during single-pilot operations, scanning patterns become more dispersed as the single pilot needs to spend more time on secondary instruments (e.g., flaps, landing gear, electronic centralized aircraft monitor) that traditionally tend to be monitored by the PM (Faulhaber, Friedrich, & Kapol, 2020). This has been shown to result in degradation in overall monitoring performance. For example, in an off-nominal scenario involving unreliable airspeed indications, Etherington et al. (Etherington, Collins, Kramer, Bailey, & Kennedy, 2017) found that SPO and RCO crews were less likely (33% and 56%, respectively, compared to 67% of the two-pilot crews) to detect an IAS disagree warning light before other, more salient, warnings were issued, such as an overspeed clacker.

7.6 Impairment and/or incapacitation of a pilot

The onboard presence of a second pilot is critical also in case of pilot impairment or incapacitation. Although pilot incapacitation is a rare event (e.g., 1 out of 34,000 flights (Australian Transportation Safety Bureau, 2016); and an annual rate of 40 cases (Evans & Radcliffe, 2012)), its consequences can be catastrophic in SPO or RCO. A distinction has been made between obvious and subtle incapacitation (IFALPA, 2013). Obvious incapacitation tends to be sudden, prolonged, and usually immediately apparent to the remaining flight crew member(s). This type of incapacitation would likely have immediate consequences that could be detected by pilots not on the flight deck or by ground-based operators. Subtle incapacitation, on the other hand, tends to be partial in nature, transient (seconds or minutes) and can be a significant operational hazard because it is difficult for other crew members (and even more so remote operators) to detect. The affected flight crew members may look well and be conscious but they may be unaware of their condition.

Several solutions for detecting and handling pilot impairment and incapacitation have been proposed (Liu, Gardi, Ramasamy, Lim, & Sabatini, 2016), such as methods for determining whether the pilot is alert for making the strategic and tactical decisions necessary to safely navigate the aircraft to an acceptable destination. For example, a procedural deviation or lack of appropriate communication can serve as a first indication of incapacitation (IFALPA, 2013). Still, considerable work is needed to determine the effectiveness of these methods and to develop effective and efficient procedures for handling cases of pilot incapacitation.

This section has highlighted contributions made by the second pilot on the flight deck. It is important to note that the presence of a second pilot seems necessary but may not be sufficient to improve safety. It will be important to improve training to ensure that the second pilot understands and is proficient on their specific role/responsibilities in a given circumstance, including how to scan instruments and monitor flight parameters as PM across flight phases and in off-nominal conditions.

In this report, we have identified and described the important contributions to safety and risk mitigation that are made by human pilots. The question often raised in discussions between practitioners and human factors researchers on the one hand, and engineers and technologists on the other is whether advanced technologies can make the same or even greater contributions to safety and thus replace the human operator. The following section will discuss this issue and present edge and corner cases that will likely require human involvement in the future, even in the presence of automation and autonomous systems.

8 Edge and corner cases that require human involvement even in the presence of advanced technologies

In this section, we describe, in generic terms, some edge and corner cases that will likely require people to stay involved in flight operations in the future, even in the presence of advanced technologies. The literature differentiates edge and corner cases. Edge cases are events that occur only at the extreme (maximum or minimum) end of an operating parameter (such as an airplane flying at its maximum airspeed). Corner cases are situations that are encountered when multiple variables or conditions are simultaneously at extreme levels but each parameter itself is within its specified range. In other words, an edge case involves pushing one variable to a minimum or maximum whereas a corner case involves doing the same with multiple variables at the same time (thus putting an aircraft at a "corner" of its multidimensional flight envelope).

The identified edge and corner cases in this section highlight two critical aspects of maintaining or recovering the system to safe operation following a failure or adverse event. First, they involve one or more human operators who play a vital—but not necessarily sufficient—role in the recognition and resolution of the problem at hand. Second, they provide opportunities for the integration of advanced technologies for early detection—and prevention—of accidents and for enhancing joint human-automation performance at the rule-based level when system safety is significantly compromised.

To identify the edge and corner cases, we revisited, grouped, and compared the accidents described in Section 3 according to the types of failures involved. In general, accidents requiring human intervention involved (1) manifestations of latent failures of hardware and software, (2) so-called “N+1” problems, i.e., problems that were considered impossible or highly unlikely by engineers and evaluators and for which therefore no procedures or checklists were developed, and (3) alarm floods due to coupling and complexity. These cases require the real-time adaptation of existing, or the development of entirely new solutions to an adverse event by human operators. The edge cases, along with example accidents highlighting human contributions, are elaborated in the following sections.

8.1 In-flight manifestation of a latent failure resulting in significant structural damage

Latent failures are a serious threat to system safety because they often lie dormant in a system for a long time until just the right set of conditions and circumstances come together to trigger a completely unexpected and often severe adverse event. In several of the analyzed accidents, the latent failure took the form of a manufacturing- or maintenance-related deficiency in an airplane component that went unnoticed until many flight hours had accrued on the affected aircraft. Edge cases required pilots to quickly examine the nature and extent of the damage. Pilots determine, based on this information as well as their knowledge of the system and current conditions, how that damage changes the constraints and performance of the aircraft/system, and develop a course of action to mitigate any further damage and keep the aircraft/system operating until it can be safely landed or shut down.

The first step in handling these events—assessing the type and severity of structural damage—is critical for making correct decisions and developing a reasonable plan of action. It often involves visual inspection and thus requires human involvement, such as a pilot going back to the cabin to visually inspect the damage sustained by an engine and/or wing. Efforts are under way to develop and utilize ultrasound and robotic technologies for aircraft body/skin and engine inspection on the ground (Jovančević, Larnier, Orteu, & Sentenac, 2015). These technologies have the potential to reduce the risk for inspectors, reduce inspection time, and lead to more reliable detection of flaws and early signs of damage, thus limiting the risk of undetected latent failures leading to catastrophic in-flight events. However, to our knowledge, no technology exists to support in-flight visual inspection of aircraft structures and components.

Developing and implementing a course of action in response to sudden structural damage in flight also benefits from pilot involvement. One specific example is Qantas Flight 32 where a sudden uncontained engine failure occurred due to non-conformance with design specifications during manufacturing of an oil pipe deep within one of the four engines. The pilots realized that roll control and ailerons were compromised and maintaining appropriate yaw and heading were critical to flight safety. They configured the outboard engines 1 and 4 to provide symmetric thrust and used inboard engine for speed control, which aided them in successfully flying and landing the airplane. Similarly, Aloha Airlines 243 suffered a sudden explosive decompression and multiple-site fatigue cracking due to a deficient inspection and maintenance program. The pilots observed that the airplane was less controllable with higher flaps settings and speeds below 170 knots and decided to use flaps 5 and IAS 170 knots for approach and landing to maintain the structural integrity of the airframe while minimizing the time to land at the nearest airport.

8.2 Unexpected aircraft behavior due to latent failures in software

The high and increasing degree of complexity of software that controls safety-critical systems on board modern airplanes creates challenges for verification of code that can lead to latent failures. Such latent failures, in turn, can have catastrophic consequences unless humans intervene quickly and effectively. For example, Qantas Flight 72 suffered two uncommanded rapid pitch-downs due to spikes in data from an inertial reference unit. The final accident report stated that the mishap "occurred due to the combination of a design limitation in the flight control primary computer (FCPC) software of the Airbus A330/A340, and a failure mode affecting one of the aircraft’s three air data inertial reference units (ADIRUs). The design limitation meant that, in a very rare and specific situation, multiple spikes in angle of attack (AOA) data from one of the ADIRUs could result in the FCPCs commanding the aircraft to pitch down." The aircraft was fitted with three ADIRUs to provide redundancy. The FCPCs used the independent AOA values from the three ADIRUs to check for consistency. When all three AOA values were valid and consistent, the FCPCs used the average value of AOA 1 and AOA 2 for their computations. However, if either AOA 1 or AOA 2 significantly deviated from each other or from AOA 3, the FCPCs used a memorized value for 1.2 seconds. The underlying algorithm had not been designed to manage a case where there were multiple spikes in either AOA value 1.2 seconds apart. The pilots on Flight 72 took several actions that resulted in a successful landing. They quickly recovered their altitude after each pitch-down event. The Captain used the First Officer's PFD for flying the airplane after deducing that the information on his own PFD was unreliable.

He also took manual control of the aircraft to be able to respond more quickly to any potential subsequent pitch-downs and conducted a series of wide left turns to maintain the aircraft's speed below 330kts while losing altitude for landing.

8.3 Dealing with the unexpected – 'n+1' problems

It is widely acknowledged that safety in high-risk domains relies on people's ability to generate new solutions to solve unforeseen problems for which no rules or procedures exist. This is often referred to as the 'n+1' problem, meaning that there is a high likelihood that there will be one or more failure modes that were not predicted or even rejected and that were not planned for by engineers and designers. These failure modes require human creativity and ingenuity under time pressure and resource constraints (e.g. Bourgeois-Bougrine (2020)). United Airlines Flight 232 is an excellent example of this challenge. The aircraft experienced a catastrophic failure of its tail-mounted engine. The separation and discharge of fan-rotor assembly parts from the engine led to the loss of all three hydraulic systems that powered the airplane's flight controls. This type of failure had been deemed by Douglas Aircraft Company, the FAA, and UAL so remote that no procedure had been developed for handling the situation. The flight crew, augmented by a UAL check airman who happened to be a passenger on the flight, was left to its own devices. They experienced severe difficulties controlling the airplane but gained limited control of pitch and roll by using asymmetric engine power.

8.4 Excessive number of alerts due to coupling and complexity

Due to the tight coupling of hardware and software components in complex systems, the onset of a failure often results in damage to and failure of related components. This can lead to an excessive number of alerts and warnings – an alarm flood (Wan & Sarter, 2022). Some of the alarms may simply signify expected effects at a distance of the original failure and not necessarily require any action. Successful resolution of such cases requires that the pilots remain calm, carefully filter out irrelevant tasks and warnings based on their mental model of the aircraft and their understanding of the situation, and prioritize tasks based on their urgency relative to the safe operation of the aircraft. Qantas Flight 32 is an example of this challenge.

The latent oil pipe failure in the case of Qantas 32 was exacerbated by damage to flight control surfaces and a cascade of nearly 60 ECAM messages. Pilots had to actively suppress irrelevant alerts to focus on aircraft subsystems that were operational and critical to performing a successful landing. For example, the crew realized that some ECAM messages could be ignored, such as several FUEL messages indicating that the aircraft was going outside its lateral imbalance limits. This was an obvious result of the fact that the aircraft was leaking fuel from various points on the left wing. ECAM procedures instructed the crew to open cross-feed valves in order to transfer fuel from the undamaged heavier right wing into the damaged lighter wing. The crew ignored these messages and did not perform the checklists. It still took the crew 50 minutes to complete all relevant ECAM procedures and checklists, and this was possible only because of the presence of five pilots on the flight deck who shared in the overall workload.

9 Conclusions and research needs

This project examined the role of pilots in risk mitigation, both during routine day-to-day operations and in the context of mishaps. Unique human capabilities and the margin of safety provided by a second crewmember were considered. Edge and corner cases were identified that will likely require the presence of human pilots for the foreseeable future. Possible ways to improve system performance further through technology and training of specific skills were explored. The effort consisted of an extensive literature review, the analysis of select accidents where human actions ‘saved the day’ and a comparison of similar adverse events with different outcomes, and an abstraction from specific cases to identify generic edge and corner cases that will likely require human involvement for some time to come.

The main conclusion from this work is that human pilots still play a pivotal role in maintaining safety in aviation operations. They contribute through highly proficient performance at the skill-based and rule-based level and by coping, on a daily basis, with extensive operational variability. Pilots must also recognize when and how to deviate from standard operating procedures and checklists to account for contextual variations in adverse events and engage in knowledge-based performance in the face of novel unexpected events that require them to develop solutions in real-time through trial-and-error. Latent failures of hardware and software, unanticipated failure modes as well as coupling and complexity leading to alarm floods were identified as some important edge and corner cases calling for pilot involvement in operations. In addition, in response to suggested in reductions in crew complements, the important role and responsibilities of a second pilot on the flight deck were highlighted. In fact, our accident analysis suggests that a better-than-expected outcome of a mishap can often be attributed, in part, to the presence of an even larger number of pilots working together.

Finally, as part of this effort, we also identified important knowledge gaps in safety research and management as well as system development. Additional empirical data is needed on resilient behavior in daily commercial transport operations as well as single-pilot and reduced- crew operations, improved pilot training, and the development of more capable and transparent flight deck technologies. Each of these areas will be discussed in more detail below.

9.1.1.1 Pilot contributions

Investigate resilience in routine operations:

- Develop markers of resilient behavior and methods to measure, understand and document how pilots use expertise and past experiences to support anticipating, monitoring, and responding to disturbances in everyday operations (Holbrook, Prinzell III, Stewart, Smith, & Matthews, 2019; Null, et al., 2019).
- Analyze qualitative data on resilient performance from databases like ASRS to study the types of actions and behaviors that enable human operators to deal with system variability and disturbances (Holbrook, Prinzell III, Stewart, Kiggins, & Kiggins, 2020).
- Connect disparate data sources to develop a robust and complete picture of resilient pilot behaviors (Holbrook, Prinzell III, Stewart, Smith, & Matthews, 2019; Kiernan, Cross, & Scharf, 2020).
- Investigate whether and under what circumstances variability in pilots' behavior should be minimized or accepted/promoted (Weber & Dekker, 2017).

Improve training:

More effective use of and collaboration with automated systems necessitates substantial changes to pilot training. In a longitudinal study, for example, Soo et al. (2021) found that current training practices and classroom-style learning enable only a basic understanding of automated systems. Efforts to elevate the role of flight deck automated systems to a virtual copilot or a teammate will require changes to training practices that support the formation of a better mental model of the automated systems. In particular, more time is needed for exploring the system and observing its behavior in a range of circumstances:

> *"Being exposed to it more has been helpful. In the sim[ulator] it was a lot about manual handling of the aeroplane, that was really what a lot of it was. There was a little bit on the automation [which] you just kind of picked up as you went. Unless you're really exposed to all these random little mode changes . . . [sometimes] you're having to just disengage the autopilot completely and just hand fly it" (Soo, Mavin, & Kikkawa, 2021, p. 723)*

The variability of training scenarios needs to increase to capture what pilots may experience in real-world situations (Clewley & Nixon, 2019). This will also reduce the predictability of scenarios and allow for the experience of startle/surprise which, in turn, supports the development of resilient behavior that includes anticipation, monitoring, responding, and learning skills (Landman et al., 2017).

9.1.1.2 Reduced crew and single pilot operations

Conduct additional research on effects of RCO/SPO on safety and performance:

- The co-location vs. remote crew paradigm needs to be investigated more closely where 1) the ground station operator is handling multiple aircraft and thus faces considerable attentional demands and 2) the crew faces off-nominal events and failures (Gore & Wolter, 2014).
- Determine what factors affect how long it takes RCO performance to stabilize and match that of two-crew complements.

Improve testing scenarios:

A major shortcoming of current approaches to alternate crew configurations where the pilot can call for dedicated ground support is that it assumes the PF will be able to detect the problem reliably in the first place. For situations that develop slowly over time and do not produce a discrete alert to the operator, the PF may not notice or diagnose the problem in time to call for dedicated ground support.

- Test monitoring, problem identification, and failure troubleshooting performance in the context of RCO and SPO with and without the presence of a remote operator and human-autonomy tools. A majority of the research on distributed crew configurations has been conducted with simple off-nominal events such as diversion due to bad weather, wheel-well fire, etc. Where play-based tools are provided, for example, how will ground station operators deal with situations in which no pre-defined play exists to handle the situation? How will operators handle situations where multiple aircraft experience severe problems and need assistance at the same time?
- Develop scenarios to test automated systems in the operational context in which it will be used.

9.1.1.3 *Develop more capable and transparent flight deck technologies*

- Improve reliability of natural-language processing and voice-based communication for integration into RCO and SPO. Even current industry-leading voice recognition tools are not accurate enough to be integrated reliably into a cockpit (Arthur III, Shelton, Prinzel III, & Bailey, 2016; Cummings, Stimpson, & Clamann, 2016).

- Investigate ways to improve the ability of automated systems to observe CRM etiquette and practices. Automated systems must provide clear indication of an aircraft's present status and its expected future state. They must also balance false alarms without withholding too much information. Additionally, automated systems need to "be aware of goals, seek clarification, offer timely information regarding limitations, and unambiguously communicate intended changes." (Geiselman, Johnson , & Buck, 2013).
- Investigate ways to make automated systems and enable them to make judgement calls. Pilots do not simply want "more automated systems" but rather systems that can take on additional responsibility, communicate using natural language, and comprehend and execute higher level cognitive functions (Bailey, Kramer, Kennedy, Stephens, & Etherington, 2017; Fennell, Pruchnicki, McKenney, Reidemar, & Comstock, 2009; Geiselman, Johnson , & Buck, 2013). According to some pilots, lack of judgement on the part of an autonomous copilot precludes its ability to be a teammate (Tokadlı, Dorneich, & Matessa, 2021).
- Develop and test ways to help pilots diagnose and understand affected systems/subsystems/functionalities in cases of equipment malfunction or failure.
- Develop tools and interfaces that support higher-level cognitive functions such as 3D graphical flight planning tools, and the ability to visualize/modify the desired flight plan relative to the terrain, other aircraft, and time (Harris, 2007).
- Investigate how automating checklist items affects pilot performance (Reitsma, van Paassen, Borst, & Mulder, 2021).

10 References

(Director) Sinyi, M., & (Writer) Kazazian , A. (2012, April 13). Impossible Landing. *(Season 11, Episode 13) [TV Series Episode] In S. Lloyd and R. Parker (Executive Producers), Air Crash Investigation. Cineflix Productions*.

Abbott, K. (2018). An overview of flight deck human factors in the U.S. Aviation Regulatory system. In *Handbook of Human Factors in Air Transportation Systems* (pp. 137-145). CRC Press. doi:https://doi.org/10.1201//9781315116549-7

ALPA. (2019). *The dangers of single pilot operations [Whitepaper]*.

American Airlines Department of Flight Safety. (2021). *Charting a new approach: What goes well and why at American Airlines.* Retrieved from https://skybrary.aero/bookshelf/books/AA%20LIT%20White%20Paper%20II%20-%20SEP%202021.pdf

Arthur III, J. J., Shelton, K. J., Prinzel III, L. J., & Bailey, R. E. (2016). *Performance evaluation of speech recognition systems as a next-generation pilot-vehicle interface technology*.

Australian Transportation Safety Bureau. (2016). *Pilot incapacitation occurrences 2010–2014.* (AR-2015-096).

Bailey, R. E., Kramer, L. J., Kennedy, K. D., Stephens, C. L., & Etherington, T. J. (2017). An assessment of reduced crew and single pilot operations in commercial transport aircraft operations. *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, (pp. 1-15). doi:https://doi.org/10/gf4stv

Billman, D., Mumaw, R., & Feary, M. (2020). A model of monitoring as sensemaking: Application to flight path management and pilot training. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64(1)*, pp. 244-248. doi:https://doi.org/10/giph2z

Bilmoria, K. D., Johnson, W. W., & Schutte, P. C. (2014). Conceptual framework for single pilot operations. *Proceedings of the International Conference on Human-Computer Interaction in Aerospace*, (pp. 4:1-4:8). doi:https://doi.org/10/ggxpxs

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42(1)*, 1-15. doi:https://doi.org/10.1111/j.1469-8986.2005.00271.x

Bourgeois-Bougrine, S. (2020). What does creativity mean in safety-critical environments? *Frontiers in Psychology, 11*. Retrieved from https://www.frontiersin.org/article/10.3389/fpsyg.2020.565884

Boy, G. A. (2014). Requirements for single pilot operations in commercial aviation: A first high-level cognitive function analysis. *Complex Systems Design and Management Conference*, (pp. 227-234).

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1993). Pictures as prepulse: Attention and emotion in startle modification. *Psychophysiology, 30*(5), 541-545. doi:https://doi.org/10.1111/j.1469-8986.1993.tb02079.x

Brandt, S. L., Lachter, J., Battiste, V., & Johnson, W. (2015). Pilot situation awareness and its implications for single pilot operations: Analysis of human-in-the-loop study. *Procedia Manufacturing, 3*, 3017-3024. doi:https://doi.org/10/gf4sht

Carrico, M., Matessa, M., & Stover, K. (2018). Reduced crew operations. *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, (pp. 1-8). doi:https://doi.org/10/gjpm4p

Carroll, M., Wilt, D., Sanchez, P., & Carstens, D. (2019). Examining pilot decision making in information-rich cockpits. *20th International Symposium on Aviation Psychology*, (pp. 73-78). Retrieved from https://corescholar.libraries.wright.edu/isap_2019/13

Certification procedures for products and articles, 14 CFR Part 21 (2022). Retrieved from https://www.govinfo.gov/content/pkg/CFR-2022-title14-vol1/pdf/CFR-2022-title14-vol1-part21.pdf

Chauvin, C., Lardjane, S., Morel, G., Costermann, J.-P., & Langard, B. (2013). Human and organisational factors in maritime accidents: Analysis of collisions at sea using the HFACS. *Accident Analysis & Prevention, 59*, 26-37.

Christian, B. (2020). *The alignment problem: Machine learning and human values* (1st ed.). W.W. Norton & Company.

Clewley, R., & Nixon, J. (2019). Understanding pilot response to flight safety events using categorisation theory. *Theoretical Issues in Ergonomics Science, 20*(5), 572-589. doi:https://doi.org/10/gjpm4k

Cover, M., Reichlen, C., Matessa, M., & Schnell, T. (2018). Analysis of airline pilots subjective feedback to human autonomy teaming in a reduced crew environment. In S. Yamamoto, & H. Mori (Eds.), *Human Interface and the Management of Information. Information in*

Applications and Services (pp. 359-368). Springer International Publishing. doi:https://doi.org/10/gjgpn4

Cummings, M. L., Stimpson, A., & Clamann, M. (2016). Functional requirements for onboard intelligent automation in single pilot operations. *AIAA*, 1-7. doi:https://doi.org/10/gjgpsq

Dao, A. Q., Koltai, K., Cals, S. D., Brandt, S. L., Lachter, J., Matessa, M., . . . Johnson, W. W. (2015). Evaluation of a recommender system for single pilot operations. *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences. 3,* pp. 3070-3077. AHFE 2015. doi:https://doi.org/10/gjgpr3

Demir, M., Ligda, S., Cooke, N., Seeds, M., Harris, M., & Niemczyk, M. (2019). Controller-pilot communication as an index of human performance in the National Airspace system. *20th International Symposium on Aviation Psychology*, (pp. 217-222). Retrieved from https://corescholar.libraries.wright.edu/isap_2019/37

Department for Transport. (2010). *British Airways flight 38.* Accident Report EW/C2008/01/0. Retrieved from https://www.faa.gov/sites/faa.gov/files/2022-11/BritishAirways38_Accident_Report.pdf

Dickson, B. (2020). *Why deep learning won't give us level 5 self-driving cars.* Retrieved from https://bdtechtalks.com/2020/07/29/self-driving-tesla-car-deep-learning

Etherington, T. J., Kramer, L. J., Bailey, R. E., Kennedy, K. D., & Stephens , C. L. (2016). Quantifying pilot contribution to flight safety for normal and non-normal airline operations. *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, (pp. 1-14). doi:https://doi.org/10/gjpm4f

Etherington, T., Collins, R., Kramer, L., Bailey, R., & Kennedy, K. (2017). Quantifying pilot contribution to flight safety during an in-flight airspeed failure. *19th International Symposium on Aviation Psychology*, (pp. 95-100).

Evans, S., & Radcliffe, S. A. (2012). The annual incapacitation rate of commercial pilots. *Aviation, Space, and Environmental Medicine*, *83*(1), 42-49. doi:https://doi.org/10.3357/ASEM.3134.2012

Faulhaber, A. K. (2019). From crewed to single pilot operations: Pilot performance and workload management. *20th International Symposium on Aviation Psychology*, (p. 7).

Faulhaber, A. K., Friedrich, M., & Kapol, T. (2020). Absence of pilot monitoring affects scanning behavior of pilot flying: Implications for the design of single-pilot cockpits. *Human Factors*. doi:https://doi.org/10/gg4wp3

Federal Aviation Administration. (2014). The safety continuum - A doctrine for application.

Federal Aviation Administration. (2017). Upset prevention and recovery training, AC-120-111. Retrieved from https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_120-111_CHG_1_FAA_Editorial_Update.pdf

Fennell, K., Pruchnicki, S., McKenney, D., Reidemar, H., & Comstock, K. (2009). Aviation automation design and implementation - The need for human factors consideration. *2009 International Symposium on Aviation Psychology*, (pp. 299-304). Retrieved from https://corescholar.libraries.wright.edu/isap_2009/65

Field, J. N., Boland, E. J., Van Rooij, J. M., Mohrmann, J. W., & Smeltink, J. W. (2018). *Startle effect management.* European Aviation Safety Agency. Retrieved from https://www.easa.europa.eu/downloads/67174/en

Flight Safety Foundation. (2021). *Learning from all operations [White paper]*. Retrieved from https://flightsafety.org/wp-content/uploads/2021/07/Learning-from-All-Operations-FINAL.pdf

Geiselman, E. E., Johnson , C. M., & Buck, D. R. (2013). Flight deck automation: Invaluable collaborator or insidious enabler? *Ergonomics in Design: The Quarterly of Human Factors, 21*(3), 22-26. doi:https://doi.org/10/gfkwdw

General operating and flight rules, 14 CFR Part 91 (2022). Retrieved from https://www.ecfr.gov/current/title-14/chapter-I/subchapter-F/part-91?toc=1

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum , J. B. (2021). A counterfactual simulation model of causal judgements for physical events. *Psychological Review, 128*(5), 936-975. doi:https://doi.org/10.1037/rev0000281

Gillen, M. (2016). A study evaluating if targeted training for startle effect can improve pilot reactions in handling unexpected situations in a flight simulator. *Theses and Dissertations, 345*. Retrieved from https://commons.und.edu/theses/345

Goh, G., Cammarata, N., Voss, C., Carter , S., Petrov, M., Schubert, L., . . . Olaf, C. (2021). *Multimodal neurons in artificial neural networks.* Retrieved from https://distill.pub/2021/multimodal-neurons/

Gore, B. F., & Wolter, C. (2014). A task analytic process to define future concepts in aviation. In V. G. Duffy (Ed.), *Digital human modeling. Applications in health, safety, ergonomics, and risk management* (Vol. 8529, pp. 236-246). Springer International Publishing. doi:https://doi.org/10.1007/978-3-319-07725-3_23

Harris, D. (2007). A human-centred design agenda for the development of single crew operated commercial aircraft. *Aircraft Engineering and Aerospace Technology, 79*(5), 518-526. doi:https://doi.org/10.1108/00022660710780650

Haynes, A. C. (1991). *United 232: Coping with the "one-in-a-billion" loss of all flight controls.* Flight Safety Foundation. Retrieved from https://flightsafety.org/ap/ap_jun91.pdf

Heaven, D. (2019). Why deep-learning AIs are so easy to fool. *Nature, 574*, pp. 163-166. doi:https://doi.org/10.1038/d41586-019-03013-5

Hilscher, M. B., Breiter, E. G., & Kochan, J. A. (2005). From-the-couch-to-the-cockpit: Psychological consideration during high-performance flight training. *Proceedings of the Third Annual Meeting of the Society for Human Performance in Extreme Environments.* Retrieved from https://www.semanticscholar.org/paper/From-the-Couch-to-the-Cockpit-%3A-Psychological-Hilscher-Breiter/1ade20c275a5a387e325f3d9b1d6b84a6b87befc

Holbrook, J., Prinzell III, L. J., Stewart, M. J., Kiggins, D., & Kiggins, D. (2020). How do pilots and controllers manage routine contingencies during RNAV arrivals? In P. M. Arezes, & R. L. Boring (Eds.), *Advances in safety management and human performance* (pp. 331-338). Springer International Publishing. doi:https://doi.org/10/gjgps5

Holbrook, J., Prinzell III, L. J., Stewart, M. J., Smith, B. E., & Matthews, B. L. (2019). Resilience and safety for in-time monitoring, prediction, and mitigation of emergent risks in commercial aviation. *20th International Symposium on Aviation Psychology*, (pp. 109-114). Retrieved from https://corescholar.libraries.wright.edu/isap_2019/19

Hollnagel, E. (2014). *Safety-I and safety-II: The past and future of safety management* (1st ed.). Ashgate.

Hollnagel, E. (2015). *RAG- Resilience Analysis Grid.* Introduction to the Resilience Analysis Grid (RAG). Retrieved from https://erikhollnagel.com/onewebmedia/RAG%20Outline%20V2.pdf

Hollnagel, E. (2016). *Resilience engineering perspectives, volume 2, preparation and restoration.* (C. P. Nemeth, Ed.) CRC Press. doi:https://doi.org/10.1201/9781315244389

IFALPA. (2013). *In-flight incapacitation - flight crew training.* Retrieved from https://www.ifalpa.org/media/2133/14pos06-in-flight-incapacitation-flight-crew-training.pdf

Johnson, W., Lachter, J., Feary, M., Vernol, B., & Mogford, R. (2012). Task allocation for single pilot operations: A role for the ground. *HCI Aero 2012 - International Conference on Human-Computer Interaction in Aerospace.* Brussels. Retrieved from https://ntrs.nasa.gov/search.jsp?R=20120016028

Jovančević, I., Larnier, S., Orteu, J. J., & Sentenac, T. (2015). Automated exterior inspection of an aircraft with a pan-tilt-zoom camera mounted on a mobile robot. *Journal of Electronic Imaging, 24*(6), 061110. doi:https://doi.org/10.1117/1.JEI.24.6.061110

Kennedy, M. (2019, August). *Al Haynes, pilot from miraculous 1989 crash landing, has died.* Retrieved from NPR: https://www.npr.org/2019/08/26/754458583/al-haynes-pilot-from-miraculous-1989-crash-landing-has-died

Kiernan, K., Cross, D., & Scharf, M. (2020). Developing a taxonomy for success in commercial pilot behaviors. *Collegiate Aviation Review International, 38*(1), 16. doi:https://doi.org/10/gjgpsv

Kinney, L., & O'Hare, D. (2020). Responding to an unexpected in-flight event: Physiological arousal, information processing, and performance. *Human Factors, 62*(5), 737-750. doi:https://doi.org/10/ghhmhj

Klein, G. (2008). Naturalistic decision making. *Human Factors, 50*(3). doi:https://doi.org/10/bn54qh

Klein, G. (2017). *Sources of power: How people make decisions.* The MIT Press. Retrieved from http://direct.mit.edu/books/book/3647/Sources-of-PowerHow-People-Make-Decisions

Klein, G., & Wolf, S. (1998). The role of leverage points in option generation. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 28*(1), 157-160. doi:https://doi.org/10.1109/5326.661098

Koch, M. (1999). The neurobiology of startle. *Progress in Neurobiology, 59*(2), 107-128. doi:https://doi.org/10.1016/S0301-0082(98)00098-7

Kochan, J. A. (2005). *The role of domain expertise and judgment in dealing with unexpected events.* [Doctoral dissertation, University of Central Florida]. Retrieved from http://www.proquest.com/docview/305363659/abstract/F8E8613D74AF4C64PQ/1

Koltz, M., Roberts, Z. A., Sweet, J., Battiste, H., Cunningham, J., Battiste, V., . . . Strybel, T. Z. (2015). An investigation of the harbor pilot concept for single pilot operations. *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences. 3,* pp. 2937-2944. AHFE 2015. doi:https://doi.org/10/gjgpnz

Lachter, J., Battiste, V., Matessa, M., Dao, Q. V., Kotesky, R., & Johnson, W. W. (2014). Toward single pilot operatoins: The impact of the loss of non-verbal communication on the flight deck. *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI- Aero '14*, (pp. 1-8). doi:https://doi.org/10/gf4stt

Lachter, J., Brandt, S. L., Battiste, V., Matessa, M., & Johnson, W. W. (2017). Enhanced ground support: Lessons from work on reduced crew operations. *Cognition, Technology, & Work,* 279-288. doi:https://doi.org/10/gf69p8

Lachter, J., Brandt, S., Battiste, V., Ligda, S. V., Matessa, M., & Johnson, W. W. (2014). Toward single pilot operations: Developing a ground station. *Proceedings on Human-Computer Interaction in Aerospace*, (pp. 1-8). doi:https://doi.org/10/gf4stz

Landman, A., Groen, E. L., van Paassen, M., Bronkhorst, A. W., & Mulder, M. (2017). Dealing with unexpected events on the flight deck: A conceptual model of startle and surprise. *Human Factors, 59*(8), 11661-1172. doi:https://doi.org/10/gd4wgk

Landman, A., van Middelaar, S. H., Groen, E. L., van Paassen, M. M., Bronkhurst, A. W., & Mulder, M. (2020). The effectiveness of a mnemonic-type startle and surprise management procedure for pilots. *The International Journal of Aerospace Psychology, 30*(3-4), 104-118. doi:https://doi.org/10.1080/24721840.2020.1763798

Ligda, S. V., Fischer, U., Mosier, K., Matessa, M., Battiste, V., & Johnson, W. W. (n.d.). Effectiveness of advanced collaboration tools on crew communication in reduced crew operations. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics* (Vol. 9174, pp. 416-427). Springer International Publishing. doi:https://doi.org/10.1007/978-3-319-20373-7_40

Lim, Y., Gardi, A., Ramasamy, S., & Sabatini, R. (2017). A virtual pilot assistant system for single pilot operations of commercial transport aircraft. *17th Australian International Aerospace Congress (AIAC17).*

Liu, J., Gardi, A., Ramasamy, S., Lim, Y., & Sabatini, R. (2016). Cognitive pilot aircraft interface for single pilot operations. *Knowlege-Based Systems, 112,* 37-53. doi:https://doi.org/10/f89czc

Lynn, R. (1966). *Attention, arousal, and the orientation overreaction: International series of monographs in experimental psychology*. Elsevier.

Lyons, J. B., Sycara, K., Lewis, M., & Capiola, A. (2021). Human-autonomy teaming: Definitions, debates, and directions. *Frontiers in Psychology, 12:589585*. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2021.589585

Madan, K., Ke, N. R., Goyal, A., Schölkopf, B., & Bengio, Y. (2021). Fast and slow learning of recurrent independent mechanisms. *arXiv: 2105.08710*. doi:https://doi.org/10.48550/arXiv.2105.08710

Martensson, L. (1995). The aircraft crash at Gottrora: Experiences of the cockpit crew. *The International Journal of Aviation Psychology, 5*(3), 305-326. doi:https://doi.org/10.1207/s15327108ijap0503_5

Martin, W. (2017). *Developing startle and suprise training interventions for airline training programs.* [PowerPoint Slides]. Retrieved from https://pacdeff.com/wp-content/uploads/2017/08/PACDEFF-FC-Forum-Presentation-on-Startle.pdf

Martin, w. L., Murray, P. S., Bates, P. R., & Lee, P. Y. (2016). A flight simulator study of impairment effects of startle on pilots during unexpected critical events. *Aviation Psychology and Applied Human Factors, 6*(1), 24-32. doi:https://doi.org/10.1027/2192-0923/a000092

Matessa, M., Strybel, T., Vu, K., Battiste, V., & Schnell, T. (2017). Concept of operations for RCO/SPO. Retrieved from https://core.ac.uk/download/pdf/84913320.pdf

May, D. N., & Rice, C. G. (1971). Effects of startle due to pistol shots on control precision performance. *Journal of Sound and Vibration, 15*(2), 197-202. doi:https://doi.org/10.1016/0022-460X(71)90534-7

McGuirl, J. M., & Sarter, N. B. (2006). Supporting trust calibration and the effective use of decision aids by presenting dynamic system confidence information. *Human Factors, 48*(4), 656-665. doi:https://doi.org/10/dbc243

McKenney, D. G. (2010). Best practices relating to stall and stick pusher training. *Flight Safety Foundation 63rd Annual International Air Safety Seminar.* Milan.

McNeese, M., Salas, E., & Endsley, M. R. (Eds.). (2001). *New trends in cooperative activities: Understanding system dynamics in complex environment.* Human Factors and Ergonomics Society.

Metz, C. (2016, March). One genius' lonely crusade to teach a computer common sense. *Wired*. Retrieved from https://www.wired.com/2016/03/doug-lenat-artificial-intelligence-common-sense-engine

Moacdieh, N. M., & Sarter, N. B. (2017). The effects of data density, display organization, and stress on search performance: An eye tracking study of clutter. *IEEE Transactions on Human-Machine Systems, 47*(6), 886-895. doi:https://doi.org/10/gck328

National Transportation Safety Board. (1989). *Aloha Airlines flight 243.* Accident Report NTSB/AAR-89/03. Retrieved from https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR8903.pdf

National Transportation Safety Board. (1990). *United Airlines Flight 232, McDonnel-Douglas DC-1010, Sioux Gateway Airport, Sioux City, Iowa, July 19, 1989.* Aircraft Accident Report NTSB/AAR-90/06. Retrieved from https://www.ntsb.gov/investigations/accidentreports/reports/aar-90-06.pdf

National Transportation Safety Board. (1991). *TACA International Airlines flight 110 - In flight with the weather.* FTW88IA109; p 5.

National Transportation Safety Board. (2009). *In-flight left engine fire American Airlines flight 1400.* Accident Report NTSB/AAR-09/03. Retrieved from https://www.ntsb.gov/investigations/accidentreports/reports/aar0903.pdf

National Transportation Safety Board. (2020). *United Airlines flight 1175.* Aviation Investigation Final Report DCA18IA092. Retrieved from https://data.ntsb.gov/Docket?ProjectID=96738

NTSB. (1989). *Safety Recommendation July 21, 1989.* A-89-53 through 69 . Retrieved from https://www.ntsb.gov/safety/safety-recs/recletters/A89_53_69.pdf

Null, C. H., Adduru, V., Amman, O. C., Cardoza, C. T., Stewart, M. J., Avrekh, I., . . . Smith, B. E. (2019). *Human performance contributions to safety in commercial aviation.* NASA. Retrieved from https://ntrs.nasa.gov/api/citations/20190001429/downloads/20190001429.pdf

Nutter, C., & Anthony, T. (2012). *The IHTAR model - A three step process for manageing anomalous events and maintaining aircraft control.* Flight Safety Foundation. Retrieved from https://flightsafety.org/asw-article/the-ihtar-model

Orlady, L. M., & Orlady, H. W. (2016). *Human factors in multi-crew flight operation.* Retrieved from https://www.taylorfrancis.com/books/9781315092881

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect (Illustrated edition).* Basic Books.

Pérez, A. A., & Behrend, F. (2019). A holographic checklist assistant for the single pilot. *2019 Integrated Communications, Navigation, and Surveillence Conference (ICNS),* (pp. 1-10). doi:https://doi.org/10/ghfjhk

Poli, E., & Angrilli, A. (2015). Greater general startle reflex is associated with greater anxiety levels: A correlational study on 111 young women. *Frontiers in Behavioral Neuroscience, 9*. Retrieved from https://www.frontiersin.org/article/10.3389/fnbeh.2015.00010

Prinet, J. C., Mize, A. C., & Sarter, N. (2016). Triggering and detecting attentional narrowing in controlled environmnets. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60 (1)*, pp. 298-302. doi:https://doi.org/10/ggxp3r

Pruchnicki, S., Key, K., & Rao, A. H. (2019). *Problem solving/decision making and procedures for unexpected events: A literature review.* [DOT/FAA/AM-20/01; pg 43]. Office of Aerospace Medicine. Retrieved from https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2020s/media/202001.pdf

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13*(3), 257-266. doi:https://doi.org/10/gfkwdc

Reason, J. T. (2008). *The human contribution: Unsafe acts, accidents, and heroic recoveries.* Ashgate.

Reitsma, J., van Paassen, M. M., Borst, C., & Mulder, M. (2021, January). Quantifying automatable checklist items on a commercial flightdeck. *AIAA Scitech 2021 Forum*. doi:https://doi.org/10/gjpm4q

Rivera, J., Talone, A. B., Boesser, C. T., Jentsch, F., & Yeh, M. (2014). Starte and suprise on the flight deck: Similarities, differences, and prevalence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58(1),* pp. 1047-1051. doi:https://doi.org/10.1177/1541931214581219

Roitblat, H. L. (2020). *Algortihms are not enough: Creating general artificial intelligence.* The MIT Press.

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control.* Penguin Books.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9*(3), 185-211. doi:https://doi.org/10.2190/DUGG-P24E-52WK-6CDG

Sarter, N. B., & Schroeder, B. (2001). Supporting decision making and action selection under time pressure and uncertainty: The case of in-flight icing. *Human Factors, 43*(4), 573-583. doi:https://doi.org/10/dss5s8

Sarter, N. B., & Thomas, R. (2022). *Flight crew: Visual scanning techniques on transport category aircraft.*

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1*(1), 45-57. doi:https://doi.org/10/bp6p75

Schmid, D., & Korn, B. (2017). A tripartite concept of a remote-copilot center for commercial single-pilot operations. *AIAA Information Systems - AIAA Infotech@Aerospace*. doi:https://doi.org/10/gjgpnf

Schmid, D., & Stanton, N. A. (2019). The training of operators in single pilot operations: An initial system theoretic consideration. *20th International Symposium on Aviation Psychology (ISAP 2019).* Dayton, OH. Retrieved from https://elib.dlr.de/122762

Schmid, D., & Stanton, N. A. (2020). Progressing toward airliners' reduced-crew operatoins: A systematic literature review. *The International Journal of Aerospace Psychology, 30*(1-2), 1-24. doi:https://doi.org/10/ggwkj2

Schmid, D., & Stanton, N. A. (Dec. 2020). Systems thinking in aerospace: The contributions to the design of future airliners' single pilot operations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 64,* pp. 188-192. doi:https://doi.org/10/gjpm4r

Soo, K. K., Mavin, T. J., & Kikkawa, Y. (2021). Mastering automation: New airline pilots' perspective. *International journal of human-computer interaction, 37*(7), 717-727. doi:https://doi.org/10/gjph48

Sprengart, S., Neis, S. M., & Schiefele, J. (2018). Role of the human operator in future commercial reduced crew operations. *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, (pp. 1-10). doi:https://doi.org/10/gf8q4n

Sternbach, R. A. (1960). Some relationships of among various "dimensions" of autonomic activity. *Psychosomatic Medicine, 22*(6), 430-434.

Stewart, M., Matthews, B., Janakiraman, V., & Avrekh, I. (2018). Variables influencing RNAV STAR adherence. *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*, (pp. 1-10). doi:https://doi.org/10/gjpm4m

Thackray, R. I. (1965). Correlates of reaction time to startle. *Human Factors, 7*(1), 74-80. doi:https://doi.org/10.1177/001872086500700109

Thackray, R. I. (1988). *Performance recovery following startle: A laboratory approach to the study of behavioral response to sudden aircraft emergencies [DOT/FAA/AM-88/4]*. Federal Aviation Authority. Retrieved from https://rosap.ntl.bts.gov/view/dot/21299

Thoroman, B., Goode, N., Salmon, P., & Wooley, M. (2019). What went right? An analysis of the protective factors in aviation near misses. *Ergonomics, 62*(2), 192-203. doi:https://doi.org/10/gjpm4c

Toews, R. (2021, June 1). What artifical intelligence still can't do. *Forbes*. Retrieved from https://www.forbes.com/sites/robtoews/2021/06/01/what-artificial-intelligence-still-cant-do

Tokadlı, G., Dorneich, M. C., & Matessa, M. (2021). Evaluation of playbook delegation approach in human-autonomy teaming for single pilot operations. *International Journal of Human-Computer Interaction*, 1-14. doi:https://doi.org/10/gjgpr5

Ververs, P. M. (1997). *Understanding a pilot's tasks.*

Vu, K. P., Lachter, J., Battiste, V., & Strybel, T. Z. (2018). Single pilot operations in domestic commercial aviation. *Human Factors, 60*(6), 755-762. doi:https://doi.org/10/ggxp2k

Wan, Y., & Sarter, N. (2022). Attention limitations in the detection and identification of alarms in close temporal proximity. *Human Factors*. doi:https://doi.org/10.1177/00187208211063991

Weber, D. E., & Dekker, S. W. (2017). Assessing the sharp end: Reflections on pilot performance assessment in the light of safety differently. *Theoretical Issues in Ergonomics Science, 18*(1), 59-78. doi:https://doi.org/10/gjgpn6

Wickens, C. D., Small, R. L., Andre, T., Bagnall, T., & Brenaman, C. (2008). Multisensory enhancement of command displays for unusual attitude recovery. *The International Journal of Aviation Psychology, 18*(3), 255-267. doi:https://doi.org/10.1080/10508410802073491

Wood, M. (2021, May). *Self-driving cars might never be able to drive themselves*. Retrieved from Marketplace Tech: https://www.marketplace.org/shows/marketplace-tech/self-driving-cars-might-never-drive-themselves

Woods, D. D., & Allspaw, J. (2019). Revealing the critical role of human performance in software. *Queue, 17*(6), 28-40. doi:https://doi.org/10.1145.3380774.3380776

Young, J. M. (2020). *The lived-experience of inflight automation failures: A qualitative descriptive phenomenological study.* [Doctoral Dissertation, Grand Canyon University]. Retrieved from http://search.proquest.com/docview/2471484216/abstract/3C1B1DBF1D6B4F5CPQ/1