May 2026

Solving the brittleness problem: Redefining airmanship in the age of increasing complexity

J.F.W. (Frederik) Mohrmann[1]

[1] Delft University of Technology, Delft, the Netherlands

ABSTRACT: The increasing performance of our aviation system has in-part been made possible through growing complexity of systems and operations, together with a drive toward standardization and compliance as safety foundations. Despite high levels of safety performance and reliability, commercial aviation accidents in the past two decades reveal a systemic, brittle crew response when faced with opacity and ambiguity resulting from this complexity. Building on research into effective flight crew behaviours and strategies in complex fourth generation commercial transport aircraft, this paper proposes a new concept for the human pilot operator – Airmanship 2.0 – to address the brittleness of the existing socio-technical flight deck system and further improve aviation safety. This concept proposes a compound pilot role to systematically harmonize compliant and adaptive behaviours and mitigate the current bifurcation of and related strain on the pilot role. This paper contrasts the proposed Airmanship 2.0 concept against notable accidents in the past decades.

KEYWORDS: Resilience, Aviation, Safety, Sensemaking, Complexity

1 Introduction

Per 2019 the worldwide commercial air transport system transported approximately 4.5 billion passengers with an accident rate of only 1.13 per million flights (IATA, 2020). Compared to the year 2001, 2019 featured a 235% increase in number of sectors flown, while also achieving 60% reduction in the number of accidents per one million flights (IATA, 2002; IATA, 2020). However, as the aviation industry's safety performance is approaching a 10^{-7} accident rate, it is showing signs of stabilizing into a non-plus ultra-safe system of the existing system design (Stoop, 2017; Airbus, 2025).

Safety studies into accidents and incidents from the past two decades (IATA, 2014; Man4Gen Consortium, 2015; Mohrmann et al., 2015) reveal specific flight deck dynamics in modern aircraft as a key limiting factor of system safety performance. This is the result of increasing cognitive distance between modern crews' and aircraft systems and operations, suffering from both the growing complexity of modern aircraft and operations, as well as a decrease in the requisite crew capacity to respond adaptively to such complexity. This cognitive distance is not a wholly novel phenomenon: Mosier et al. (1998) and Manzey et al. (2012) described similar performance-limiting mechanisms such as automation bias and automation complacency. However, in recent years these mechanisms are becoming the major barriers to improving safety, and may be exacerbated by future increase of complex and potentially even more opaque systems such as AI-teaming (Eurocontrol, 2020; EASA, 2023; Lorrig & Daw, 2024; Kirwan 2025; Hörhager, 2025) as well as the advent of Reduced Crew Operations (RCO) (Malick & Gollnick, 2016; Sprengart et al., 2018; Schmidt & Stanton, 2020), or their combination (Minaskan et al., 2022).

Research into this control gap has been able to reproduce flight crews' difficulty to manage opaque and ambiguous situations in the flight deck (Sarter & Woods, 1994; Saurin & Carim, 2012; IATA, 2014; Man4Gen Consortium, 2015; Mohrmann et al., 2015; Stoop & Van Kleef, 2015; Wolter et al., 2015; Strauch, 2017; Kharoufah et al., 2018; Proctor & Van Zandt, 2018; Kelly & Efthymiou, 2019; Prinzel et al., 2024; Gago et al., 2025), as well as revealing flight crew responses that *are* effective in managing complexity and related opacity (Rankin et al., 2016; Field et at., 2017; Mohrmann et al., 2017; FutureSky P6 Consortium, 2017; Field et al., 2018; Banks et al., 2020; Landman et al., 2020; Hancock et al., 2022; Sarter, 2024). Behavioural studies such as Man4Gen as well as industry initiatives such as Competency Based Training and Assessment (CBTA) and Evidence Based Training (EBT) (ICAO, 2013; EASA, 2020; IATA, 2024) acknowledge the need to adjust the pilot role toward the inclusion of more effective responses to further improve flight safety.

The dynamic nature of such responses, however, lies in stark contrast to the long-held (and still robust) performance standards centred on compliance, error-reduction and reproducibility which account for the strong safety performance in recent decades. The contrast between and need for both compliant and adaptive control paradigms strains the pilot role. This paper explores this dichotomy of compliant versus adaptive control and sets forth a new, compound concept for the pilot role, *Airmanship 2.0*, as an effective, efficient and responsible assimilation of both control modes. The concept is explored in the context of seven commercial aviation accidents which occurred in the past two decades.

2 The current flight deck control paradigm

The proposal of a new framework for the pilot model is incumbent on several concurrent developments in aircraft systems, the pilot profession and the interaction between them.

2.1 Evolution of aircraft systems

The past decades have seen significant evolutions in flight deck systems across four generations of flight decks (Abbott, 2017; Airbus, 2025). To meet increasing operational, environmental and economic demands, flight deck systems feature high levels of integration and level of autonomy (Boy, 2020; Clark & Wilson, 2024). Key examples of such systems include flight envelope protection and active upset recovery technology, GPS-based precision navigation, automatic centre-of-gravity fuel-balancing systems, automated engine startup and the advent of paperless flight decks.

The growing number of systems also drives more system integration such as EICAS/ECAM[1] fault management systems which directly invoke Quick Reference Handbook (QRH) checklists to be executed. Similarly, digital checklists can detect the state of several aircraft system configurations and monitor checklist completion (e.g., the “No Blue” callout during Airbus approaches) (Landry, 2009). The Flight Management System (FMS) can receive a flight plan from an operator’s dispatch team via a wireless

---

[1] EICAS: Engine Indication and Crew Alerting System; ECAM: Electronic Centralized Aircraft Monitor

connection at the gate or in the air via CPDLC[2], which automatically does performance calculations for weight and balance, take-off performance, flight navigation and auto-flight system configurations, all with minimal crew interactions (mostly confirmations). Bowles (2018) and Feary (2018) provide an overview of this evolution in automation, including an interesting perspective of *Automation Readiness Levels*, for further reading.

A growth in system autonomy is also seen in flight envelope protection systems (Landry, 2009), Primary Flight Display (PFD) declutter modes (SAIA, 2016), automated ground-based support systems (Ho et al., 2017) and future free-flight operations with aircraft autonomously interacting to manage traffic flows (Enea & Porretta, 2012). The developments and promises of Artificial Intelligence may further expand the autonomy of systems (Kulida & Lebedev, 2020; Lorrig & Daw, 2024), also echoed in the AI roadmaps provided by EASA and the FAA (EASA, 2023; FAA, 2024).

2.2 Evolution of the pilot profession

Concurrently, the pilot role has shifted. Boeing predicts 804,000 new pilots (Boeing, 2019) required in the next two decades, while average pilot salaries have dropped to less than 20,000 USD in the past decade (Crouch, 2020) together with reduced job security (Fraher, 2019). This trend is very present in the US where the first 1,500 flight hours requires hour-building in at lower-pay regional operators (Caraway, 2020). In contrast, the financial burden of flight training has increased with higher fuel prices and state- or airline-funded training programs being replaced with loans and even pay-to-fly schemes (Valenta, 2018; Maxwell & Grant, 2021).

In addition to this, pilot training efforts have been under pressure to answer for pilot shortages by increasing throughput, as well as remaining commercially viable in a competitive, low-margin market

---

[2] CPDLC: Controller Pilot Data Link Communications

(Soo, 2018; Adanov et al., 2020). On-aircraft familiarisation has made way for more simulator and line training such as Zero Flight Time Training (ZFTT), ab-initio training has made way for Multi Pilot Licence (MPL) training (Wikander & Dahlström, 2016), reducing (frozen) Air Transport Pilot Licence (ATPL) small aircraft flying time from 200 to 120 hours and condition cadets specifically towards the right seat of a regional commercial aircraft such as an A320 or 737.

Lastly, the (r)evolution in airline networks is negatively affecting pilot fitness. Twelve- to fifteen-hour flight routes and 24-hour airport operations have extended work into non-circadian times (Sallinen et al., 2020; Shah, 2024). This is acknowledged by efforts which put considerable effort into recognizing and mitigating fatigue risks for example through the implementation of Fatigue Risk Management Systems (FRMS) (ICAO, 2016; UK CAA, 2019). The combination of financial stress, circadian stress and other demands such as cultural diversity in the flight deck, forced relocations and the strong cyclic nature of the industry puts increasing strain on the physical, mental and emotional capacity of the human pilot (Cullen et al., 2021).

2.3 Evolution of pilot-system interactions

The above developments in technology, training and operations describe a shift in flight deck collaboration (Landry, 2009; Abbot, 2017; Boy, 2020; Clark & Wilson, 2024). The ten levels of automation listed in Table 1 serve as a useful reference to describe this shift (adapted from Parasuraman et al., 2000). “Implementation” tasks were the first to be offloaded to automation around the 1950’s with the advent of autopilot systems (around levels 3-4), “Generate” tasks shifted to automation since the 1960’s with the introduction of warning systems (around levels 5-6), “Select” shifted to automation since the 1990’s with the introduction of EICAS and ECAM systems (around levels 7-8), leaving “Monitor” for the pilot in the most recent years (currently around levels 8-9).

Table 1. Levels of Automation (Parasuraman et al., 2000)

Information-Processing Functions
H: Human, C: Computer

	Levels of Automation	**Monitor**	**Generate**	**Select**	**Implement**
1.	*Manual control*	H	H	H	H
2.	*Action support*	H/C	H	H	H/C
3.	*Batch processing*	H/C	H	H	C
4.	*Shared control*	H/C	H/C	H	H/C
5.	*Decision support*	H/C	H/C	H	C
6.	*Blended decision making*	H/C	H/C	H/C	C
7.	*Rigid system*	H/C	C	H	C
8.	*Automated decision making*	H/C	H/C	C	C
9.	*Supervisory control*	H/C	C	C	C
10.	*Full automation*	C	C	C	C

This evolution can be described as a shift from a pilot-centric model (where the automation supports the human) to an automation-centric model (where the human monitors the automation), shepherded through key human-machine design philosophies such as human-centred design (Kirwan, 2025), cognitive ergonomics (Harris, 2017) and human-systems integration (Boy, 2020). More recently, research and policy making in human-AI teaming (HAT) also anticipate a similar transition toward greater autonomy and a reduced role for the human operator (EASA, 2023; Lorrig & Daw, 2024; Kirwan, 2025). An example of this can be found in the EU HAIKU project which defines an evolution through six levels of HAT from enhancing information for human operators through to full AI autonomy (Kirwan, 2025).

These developments in autonomy and system integration cognitively isolated the flight crew from control loops and, combined with the highly reliable nature of systems, induce monitoring errors and automation bias (Mosier et al., 1998; Mumaw et al., 2001; Manzey et al., 2012; Bradshaw et al., 2013;

Man4Gen Consortium, 2012; Schutte, 2015; Man4Gen Consortium, 2015; Sarter, 2024). Ironically, the mitigation of such new human error modes furthers the shift toward automation in the flight deck. This process of automation deference was already described by Bainbridge (1983). Klein (2011) aptly describes this process as the *Self-Reinforcing Complexity Loop*. Kirwan (2025) also highlights the risk of future AI systems exacerbating the already-present opacity of automated systems and suggests that existing human factors design methodologies may not be sufficient for such future teaming modalities.

2.4 The brittleness problem

The strong interactions between the above developments in systems, the pilot role and operations require appreciating their collective performance as a Joint *Cognitive System* (JCS) (Hollnagel & Woods, 2005; Woods, 2019). The combination of rigid (deterministic) automated systems, flight crews strongly conditioned toward procedural compliance and rationed information-sharing between systems and pilots results in a JCS that is performing very strongly within the domain of deterministic (known) operations.

However, as the accident cases in Section 5 will illustrate, beyond the realm of deterministic operations, when faced with opacity, this JCS struggles to respond in an adaptive way and make sense of “unscripted” situations. The strongly conditioned rigidity of both systems and pilots in the JCS sacrifices cognitive flexibility, dynamic responses and sensemaking (Manzey et al., 2012; Man4Gen Consortium, 2012; Mohrmann et al., 2015; Mohrmann et al., 2017; Sarter, 2024). As such, the JCS can better be described as *brittle*: featuring high performance through reproducibility within the deterministic operational domain, yet catastrophic failure (to adapt) beyond that domain (Woods, 2015; Hancock et al., 2022).

Figure 1 organizes several factors that drive this brittle characteristic. These factors have been reproduced in several research activities such as the EU FP7 project MAN4GEN (Man4Gen Consortium, 2015), a CAA UK investigation into pilot fatigued performance (UK CAA, 2019), an EU-contracted global pilot fatigue data study (Sallinen et al., 2020), an EASA-contracted startle and surprise management study (Field et al., 2018), research into performance-based training such as Evidence Based Training (EBT) (ICAO, 2013) and Horizon 2020's "FutureSky Safety: Human Performance Envelope" project (FutureSky Safety P6 Consortium, 2016).

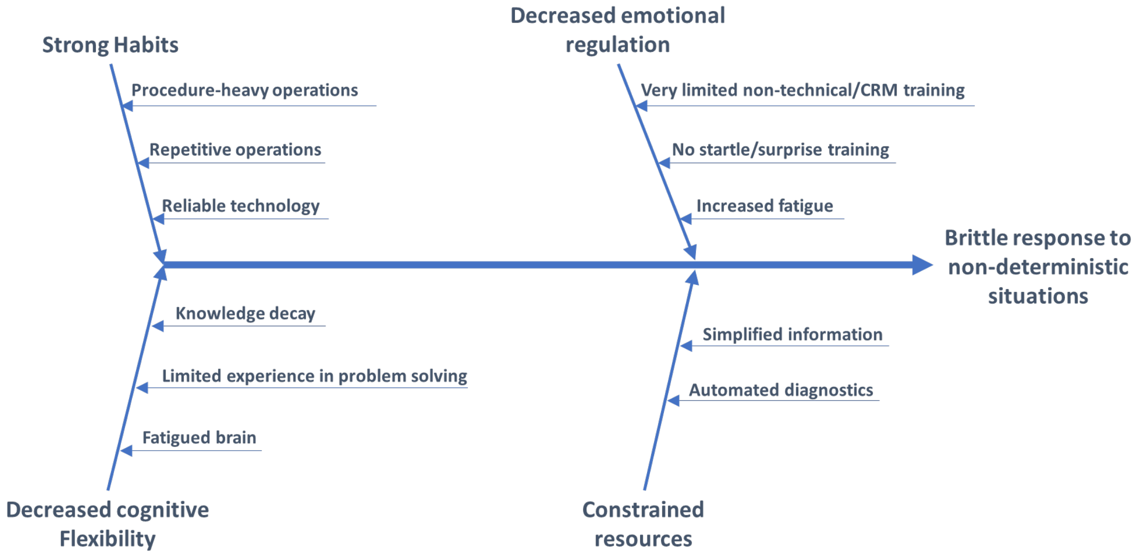

Figure 1. Contributory factors for brittleness in modern flight deck JCS response to complexity

Although this brittle JCS currently still maintains a high level of safety performance, problems arise when underlying design assumptions no longer hold true. Constant (1973) and Vincenti (1990) describe this as a *presumptive anomaly*:

> *"[A presumptive anomaly] occurs when scientific insight or assumptions derived from science indicate either that under some future conditions the conventional system will fail (or function badly) or that a radically different paradigm will do a much better job or will do something entirely novel."* (Constant, 1973 – pg. 555)

Research into pilot resilience, cognitive control and modern human-machine teaming showcase two types of anomalies that this JCS is presumed to be subjected to.

The first (Type 1) presumptive anomaly is that the JCS will not maintain sufficient flight safety due to progressive, compliance-driven *erosion* of flight crew adaptiveness. Flight crew knowledge, awareness and flexibility degrade more than presumed due to prolonged exposure to high system reliability, operational consistency and procedural emphasis. This results in relatively high *experienced* opacity (even with known, but rare, system states). In analogy: the JCS has become more brittle than assumed.

The second (Type 2) presumptive anomaly is that this deterministic JCS will not maintain sufficient flight safety in modern and future aircraft systems featuring greater *true* opacity (due to increased complexity, integration and autonomy). This is because the JCS lacks the ability to *systematically* respond in a dynamic and unscripted way to non-deterministic situations. In analogy: future aircraft may require even greater JCS "ductility"[3] than the current (brittle) JCS is designed for.

Figure 2 depicts both presumptive anomalies in the complexity space. Figure 7 in Section 5 will illustrate six notable accidents (actual anomalies) within this complexity space as early evidence of the presumed anomalies. By recognizing and appreciating these two anomaly mechanisms can a new—less brittle—JCS be considered as a new control paradigm.

---

[3] Ductility in the material sense: a physical property that measures a material's ability to undergo plastic deformation under stress without fracturing. I.e., changing shape in reaction to the specific stress applied.

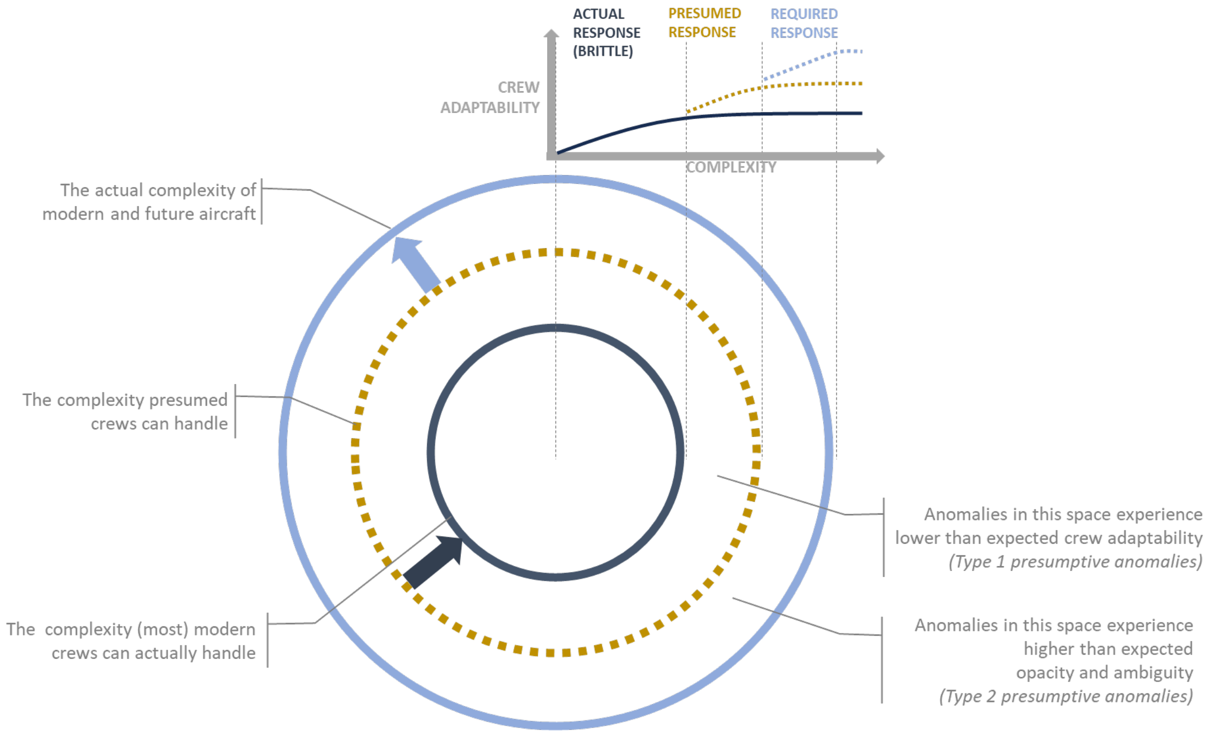

Figure 2. Visualisation of JCS brittleness Type 1 and Type 2 presumptive anomalies

3 Solving for the brittleness problem

The two presumptive anomaly types described in subsection 2.4 share a common mitigation by means of reinforcing flight crews' ability to adapt their responses in opaque and ambiguous situations. This raises the question how to achieve this without disrupting the existing, compliance-based safety paradigm. The following subsections take a closer look at deterministic and non-deterministic control paradigms to identify how they may be assimilated.

3.1 The dichotomy of pilot control paradigms

In a general sense for all paradigms, pilots[4] have control when their *understanding*, *decisions* and *actions* reflect and resonate well with the specific nature of the situation they are experiencing. This is

---

[4] The term pilot refers to the entire flight crew in a multi-crew JCS.

the basic concept of Neisser's Perceptual Control Cycle (PCM) (Neisser, 1976). In deterministic situations, past occurrences have resulted in checklists and procedures to streamline the process of recognizing a situation and selecting appropriate actions (e.g., memory items during a birdstrike on departure). The evolution toward such prescribed and trained recognition-primed decision making has been very effective for most (if not all) unambiguous operational situations (O'Hare, 1992; Li et al., 2014). Pilots proficient in such proceduralized problem-solving trust their training, apply procedures diligently and are sensitive to recognizing and correcting for errors. Tony Kern's airmanship model (Kern, 2010), shown in Figure 3, provides an apt model of the pilot role in the deterministic control paradigm.

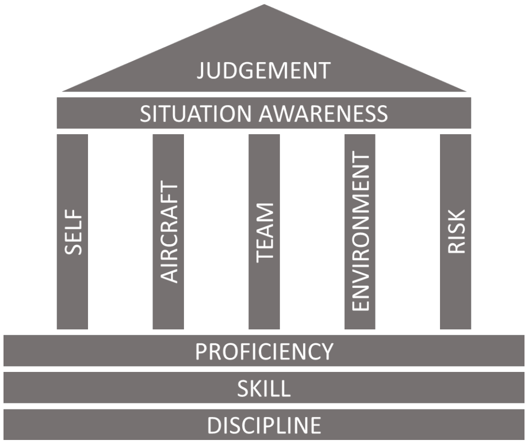

Figure 3. Tony Kern's Airmanship Model (Kern, 2010)

The apex of Kern's model is proper awareness and judgement, supported by five pillars of knowledge, which are in turn supported by a foundation of proficiency, skill and discipline. This model has a good fit for deterministic operations, where an upkeep of proficiency and knowledge supports an effective recognize-and-act way of working.

In non-deterministic situations, crews undergo the same basic process of understanding a situation, deciding on a course of action and executing accordingly, but without predetermined solutions. This goes beyond existing resolution frameworks such as T-DODAR, FORDEC and DECIDE[5] (Banks et al., 2020), as these still emphasize "*recognize* the failure" and "performance *appropriate* actions" in situations which (by definition) lack known references and standards. Rather than a recognize-act response, the crew must cognitively engage in *sensemaking* (Hollnagel & Woods, 2005; Rankin et al., 2016).

Sensemaking is a generative process in which a person (or crew) assimilates raw information about the aircraft, flight path and situation into hypothesized scenarios that explain symptoms and observations. This is similar to the concept of "fluid intelligence" (Ziegler et al., 2012; Kent, 2017, Landman et al., 2017) described as the ability to arrange (new) variables into a coherent mental model. In contrast, "crystalline intelligence" is a stockpile of knowledge, insights and beliefs gained over a pilot's lifetime (like Kern's knowledge pillars). Where sensemaking relies strongly on fluid intelligence, crystalline intelligence (i.e., experience and knowledge) contributes to the sensemaking process through heuristic-based problem solving (Kahneman, 2011). Behaviours echoing sensemaking have recently also been operationalized through the ICAO pilot competency framework (ICAO, 2013; EASA, 2020; IATA, 2023) in competencies such as "*Problem Solving and Decision Making*" and to a lesser extent "*Situation Awareness and Management of Information*".

Effective sensemaking requires higher cognitive functioning, which often decays quickly under stress (Field et al., 2018; Landman, 2020). For this reason, effective sensemaking requires a pilot to be aware of his/her own cognitive capacity and be able to re-engage it after a setback. Prime examples of this are startle and surprise recovery training (Field et al., 2018; Landman, 2020), increased emotional intelligence for collaboration (Matthews et al., 2003) and active fatigue prevention (Sallinen, 2020). In

---

[5] T-DODAR: Time, Diagnose, Options, Decide, Assign, and Review
FORDEC: Facts, Options, Risks and benefits, Decide, Execute, Check
DESIDE: Detect, Estimate, Set safety objectives, Identify, Do, Evaluate

the ICAO pilot competency framework, "*Workload Management*" describes similar behaviours such as "*OB 8.9 Manages and recovers from interruptions, distractions, variations and failures effectively while performing tasks*" (ICAO, 2013). Table 2 illustrates the control dichotomy by contrasting the deterministic (prescriptive, compliant) paradigm versus the non-deterministic (sensemaking, learning) paradigm.

Table 2. Comparison of deterministic and non-deterministic control paradigms

	Deterministic control paradigm	**Non-deterministic control paradigm**
Safety by	Reproducibility	Adaptability
Most effective in	Transparent situations	Opaque situations
Promoted behaviour	Respect procedures Remain within all limits Focus on execution	Challenge assumptions Focus on understanding Generate options
Undesired behaviour	Non-compliance Question procedures Lack of punctuality	Making assumptions No-cross checking Lack of system interest/knowledge

3.2 The need for a singular pilot role

Both control paradigms have strong but different contributions to safety, although neither will suffice on their own. *Reproducibility* has been effective in reducing workload as increasingly demanding aircraft and flight operations strained pilot cognitive capacity. An overreliance on this paradigm, however, has resulted in the brittleness problem presented. *Adaptability* (sensemaking) provides a pathway in opaque and ambiguous situations. In turn, an overreliance on this paradigm would result in unreasonable workload as it is much more cognitively demanding than a procedural response.

Figure 4 illustrates how the ratio of these paradigms has approximately changed over time (blue), contrasted with the industry safety performance (green). The asymptotic limit seen in the safety performance (Maligne, 2011; Airbus, 2025) drives the re-appraisal of these control paradigms, particularly considering the two brittleness-related anomalies. Not doing so may even result in a decline in safety performance (purple) with more Type 2 anomalies. To ensure the safety performance improves beyond the current level (orange, upward trajectory), it is imperative that the new pilot role leverages both workload-reducing and adaptivity-enhancing mechanisms.

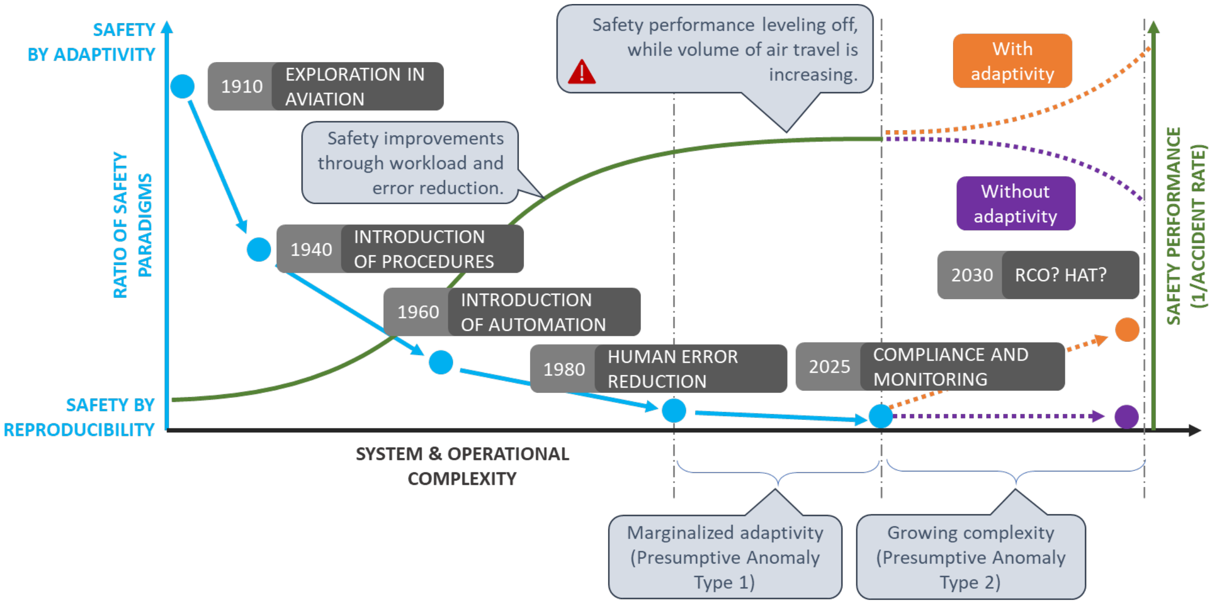

Figure 4. Conceptual visualisation of control paradigm ratio (blue) and safety performance (green) over time with future scenarios (orange and purple)

There are three other arguments supporting a singular role rather than split paradigms. The first is the maintenance of recency and competence in both compliance and adaptiveness. As these paradigms are diametrically opposed, consistent use of one will result in the decay of the other (e.g., presumptive anomaly Type 1). A pilot role that consistently and repeatedly engages both compliance and adaptability ensures that pilots maintain proficiency in responding to both deterministic and non-deterministic situations. While the ICAO pilot competency framework (ICAO, 2013) appreciates behaviours from both paradigms, adaptive behaviours are a clear minority, most notably without any *systemic integration* to mitigate their decay. This “role-conflict” was already appreciated by Degani and Weiner (1994), describing it as an undesired “switching” between two operator modes.

The second argument is for removing standards duplicity for pilot behaviour. The current strong emphasis on compliance presents accountability risks for pilots adapting and deviating from procedures, even when it may have improved the situation (Huijbrechts & Van Paassen, 2023). Unless it is established as part of the accepted behavioural framework, the requisite adaptability behaviours discussed in Section 3.2 will be subject to a barrier as they are (often in hindsight) mutably labelled as a *merit* (e.g., in Quantas Flight 32[6] and the DHL A300 missile strike in Baghdad (Maligne, 2004)) or an *error* (e.g., in Air Asia Flight 8501[6] and Asiana 214[6]).

The third argument for providing a single, unambiguous pilot role is its function as a design principle for efficient (re)design of the larger JCS. This allows for all human actors, systems and interactions to support the new pilot role rather than maintaining functionalities strongly engineered toward only one of the control paradigms. The DLR Risk Information System (RIS) developed in Man4Gen (Buch et al., 2017) or Reduced Crew Operations (RCO)-supporting automation as proposed by Ligda et al. (2015) are examples of such recent design efforts enabling a more dynamic pilot role.

---

[6] These cases are not examples of desired behaviour but rather illustrate *poorly scaffolded* adaptive behaviour. These cases will be discussed further in Section 5.

4 Airmanship 2.0: A compound pilot role

The Airmanship 2.0 concept proposes an adaptation to Tony Kern's airmanship model by systematically *nesting* compliant behaviour *within* a resilient attitude. This mitigates Type 2 presumptive anomalies (i.e., more frequent opaque operations) by structuring a *fail-safe* approach to the blend of transparent and opaque situations. In other words:

> *It is safer for flight crew*[7] *to assume opaque operations and, upon realizing that a situation is transparent, transition to a prescribed action, rather than assuming that an opaque situation is simpler than it is.*

This nesting is achieved through three extensions of Kern's model, as illustrated in Figure 6. The first is establishing a foundation of belief and *acknowledgement* of non-deterministic realities in modern flight operations. By voiding the illusion of a deterministic operation, pilots feature a wider safety aperture, which connects well with the principles of proactive, continuous Threat and Error Management (TEM) (Merrit & Klinect, 2006).

A second key adaptation is introducing the pilot's *cognitive self-regulation* to maximize cognitive facilities of oneself and others. This ability is essential to maintain high mental working capacity for sensemaking and engaging Kern's pillars of knowledge during opaque situations, and as such serves as a foundation for them in the model. It is placed above discipline, skill and proficiency as these capacities support cognitive regulation habits (e.g., meditation, breathing exercises, fatigue management). The increased propensity for startle and surprise that comes with highly reliable and complex systems

---

[7] This same principle may also apply to other human actors facing potentially opaque situations in aviation and other domains.

(Martin et al., 2012; Field et al., 2018; Vlaskamp et al., 2025) calls for pilot cognitive recovery as an indispensable and explicit part of the new pilot role.

The third model adaptation requires *adaptability of mental models* as an essential foundation for situation awareness in opaque situations. This does not imply that pilots should doubt every mental model they have learned but that they recognize that these are only models with underlying assumptions, and that reality may prove to work differently. This attitude mitigates confirmation bias when seeking to re-establish situation awareness and invites a learning attitude. Such learning can be operationalized through practical sensemaking strategies as proposed in the Man4Gen project (Field et al., 2017; Mohrmann et al., 2017). Figure 5 illustrates these three adaptations to Kern's airmanship model.

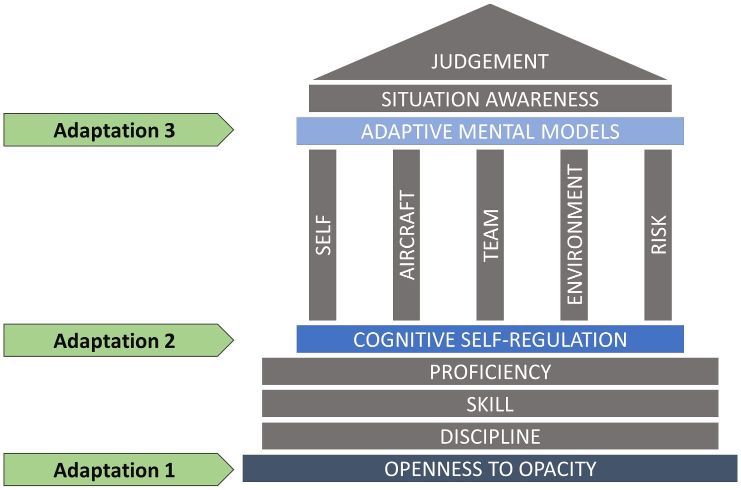

Figure 5. The Airmanship 2.0 adaptations of Kern's airmanship model

With this nested approach the Airmanship 2.0 concept aims to increase JCS adaptiveness while retaining established airmanship components. In keeping with the analogy to material brittleness and ductility: This nested approach is similar to compound materials featuring complementary performance through multiple substrates (e.g., rebar in concrete, or carbon fibre reinforced polymers), this concept facilitates the co-existence of contributions of both existing compliance (providing "strength") and new sensemaking behaviours (providing “ductility”). This can potentially result in three practical synergies between them. The first potential synergy is that the appreciation of opacity and mental model flexibility may invite aircrew to recognize critical (often implicit) assumptions underlying *diagnostic* procedures and actions, to improve their relative priority according to situational applicability.

The second potential synergy is that targeted deterministic responses (i.e., Standard Operating Procedures (SOP’s), checklists, memory items) may improve the efficiency of the sensemaking process by reducing mental workload through manageable work items, shared situation/action awareness and reducing the number of slips and lapses. It is critical that such execution has an underlying *sensemaking objective*, and feeds back into the crew’s situation awareness.

The third potential synergy may be that improved sensemaking of the current situation or system characteristics can help to identify which standardized *recovery* procedures and actions may be most appropriate, and which to avoid. Effective interventions may in turn continue to contribute to the (continuous) sensemaking process to reduce opacity. Figure 6 visualizes these three potential synergies between efficient (procedural) and sensemaking (adaptive) responses.

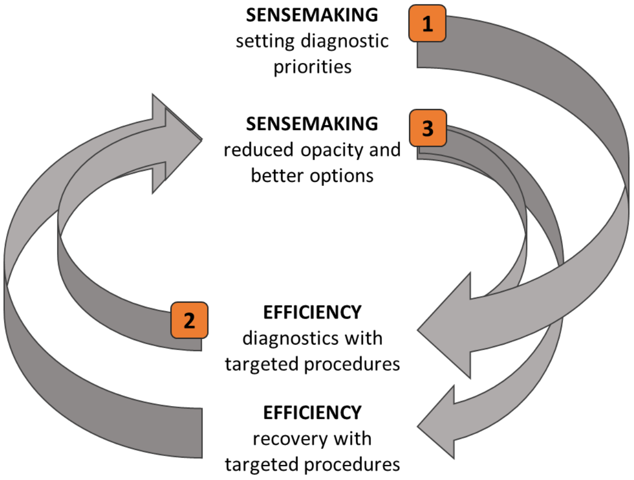

Figure 6. Airmanship 2.0 synergies between sensemaking and procedural efficiency

As a final note, it is worthwhile to consider the Airmanship 2.0 concept at the JCS level, particularly how systems can better support this new pilot role. For example: to maximize pilot cognitive capacity for sensemaking, high workload tasks including calculations and manual flight may be delegated to automation (ACROSS Consortium, 2016; Stanton et al., 2016). Aircraft systems could also proactively support mental model adaptation through varying the level of system details and providing trend data (Cahill et al., 2017). In the future, situation-sensitive systems (e.g., AI-based) may also support pilots in option generation and risk assessment (Kulida & Lebedev, 2020). Table 3 illustrates possible differences between the existing JCS and the Airmanship 2.0 JCS.

Table 3. High-level contrasting of the existing flight deck JCS with the proposed Airmanship 2.0 JCS

	Existing JCS	**Airmanship 2.0 JCS**
Safety by	Reproducibility	Adaptability & Reproducibility
Most effective in	Transparent situations	Opaque & transparent situations
Pilot role	Monitor automation Manual flight Normal procedures Non-normal procedures Problem solving & decision making Optimize flight	Continuously build and improve understanding of the situation Identify and execute diagnostic queries Re-evaluate mission priorities (Re-)assess downstream options
Automation role	Autoflight Warnings Status information (on demand) Autonomous system reconfiguration	Autoflight Accelerate restructuring of mental models Support option-generation Transparent system reconfiguration

A review of the JCS will invariably invite a review of aircrew training programs, where existing competency-oriented, scenario-based training such as CBTA and EBT may serve a solid departure point for training preparing pilots for such an Airmanship 2.0 flight deck JCS. Section 5 discusses the Airmanship 2.0 concept by studying seven accident cases from the past two decades from the perspective of Airmanship 2.0.

5 Seven anomalies of opacity

Figure 8 reproduces Figure 3 with the placement of seven aviation accidents characteristic of Type 1 or Type 2 brittleness anomalies. There is one exception (Quantas Flight 32) which illustrates a so-called “black swan event” (Taleb, 2015), where the event likelihood is orders of magnitude smaller than current industry accident likelihoods.

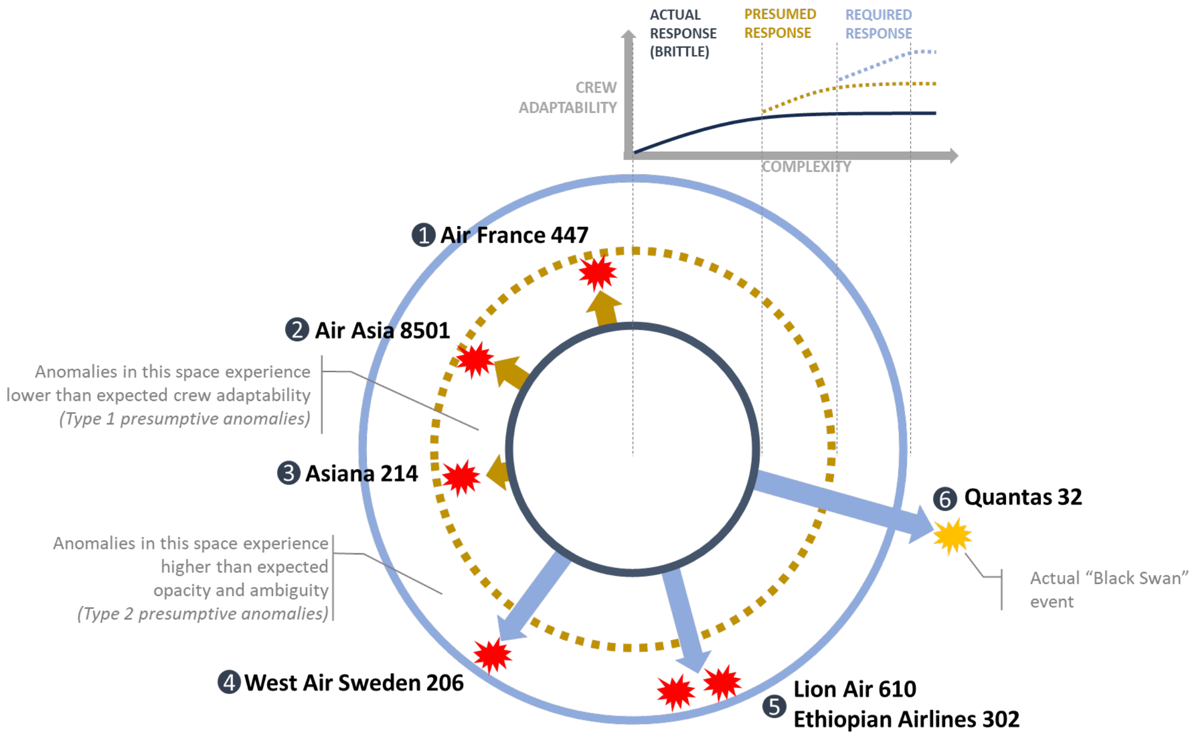

Figure 7. Illustration of seven accidents (anomalies) in the complexity space of Figure 2

Each case is first briefly summarized and then features a reflection on how the Airmanship 2.0 concept could have contributed to a less brittle response. The exception is Quantas 32 which illustrates a strong application of Airmanship 2.0 principles *avant la lettre*. The purpose of these cases is to revisit the underlying opacity and ambiguity flight crews experienced and how Airmanship 2.0 could have applied to such actual cases. The focus is not on the outcome and related crew actions and associated hindsight bias.

5.1 Air France 447

5.1.1 Synopsis (BEA, 2012)

An Airbus A330 experienced icing on the pitot tubes, resulting in inaccurate air data. The autopilot disconnected promptly, the aircraft control law switched from normal to alternate law, followed by the autothrust system disengaging. The pilots were unable to maintain control of the aircraft which was flying at an altitude where overspeed and stall speeds approach each other, combined with a removal of airspeed information. The resulting stall at night in low visibility was not detected, the aircraft crashed into the ocean.

5.1.2 Airmanship 2.0 reflection

Investigation into the crew functioning indicated a possible startle reaction (freeze) that may have caused inappropriate control inputs, possibly in reaction to greater *true* opacity (Type 2 anomaly). However, one year prior, the Airbus Operator Conference specifically addressed 1) the issue of a loss of airspeed indication, 2) the startle reaction that it could cause, providing guidance to operators for loss of airspeed indication in both operations bulletins and training guides (BEA, 2012). As such, this accident may have instead featured greater *experienced* opacity without higher true opacity (Type 1 anomaly). Despite this, the removal of a primary flight metric (airspeed) introduced ambiguity about the aircraft's attitude, speed margins (with stall speed and overspeed limits in close proximity at such cruise altitudes) and autoflight capabilities hours aways from the nearest airfield. The disappearance of such a primary indicator may well induce surprise and possibly even startle, requiring pilots' cognitive and emotional recovery before engaging in sensemaking and taking further action. Such self-restoration may be counter-intuitive when warning systems and disengaging autoflight systems are indeed "panicking" without airspeed indications, but a recovery allows crews to re-engage in bigger picture sensemaking (energy management and energy margins) and prevent over-focus on (narrow-scoped) system warnings. Several BEA report recommendations (BEA, 2012) align with Airmanship 2.0, notably training of surprise management skills and improving the design of flight deck information and warning systems to reduce the opacity of a situation and provide crews with more, clearer cues to diagnose the situation, even for such extremely rare, *but known*, situations.

5.2 Air Asia 8501

5.2.1 Synopsis (KNKT, 2015)

An Airbus A320 experienced a malfunction of the Rudder Traveler Limiter Unit (RTLU), later attributed to a tiny crack in a soldered connection. This specific aircraft experienced a similar RTLU malfunction 23 times in the previous year. Airbus procedures require resetting the RTLU. During this flight, the system presented an RTLU failure four times. After the fourth time taking prescribe ECAM reset actions, the captain elected to reset the two Flight Augmentation Computers (FAC) using the circuit breaker, placing the aircraft in alternate law and manual control. Shortly after, a small, sustained rudder deflection resulted in an undetected roll, after which the aircraft entered a stall and crashed.

5.2.2 Airmanship 2.0 reflection

Similar to the Air France 447 accident, the chain of events begins with a small component failure (solder crack) causing a great deal of opacity in the flight deck, showcasing the highly leveraged nature of complex systems. With ECAM actions not effective in resolving the recurring RTLU fault, the crew was presented with both opacity ("what is going on in the system?") and ambiguity ("reset to recover rudder or leave it and continue with limited rudder control?"), particularly with the deteriorating weather in the area. While the report notes contributory factors such as insufficient UPRT, poor crew coordination (e.g., no pilot flying/monitoring role divisions) and poor Threat and Error Management (TEM) briefings, it was also suspected that the captain's non-standard FAC-reset was based on a previous maintenance action to reset the RTLU with a circuit breaker reset and confirming that this could be done every time the failure was presented on the ECAM (KNKT, 2015). While this may have been permitted on the ground, Airbus did not list this reset as permitted during flight, stating that such a reset must be carefully considered. While seeking other resolutions when prescribed actions are ineffective can be regarded as resilient or adaptive behaviour when prescribed actions are ineffective, it must be sufficiently scaffolded in a sensemaking process that considers threats and contingencies. An Airmanship 2.0 approach would activate a sensemaking process at the very first ECAM warning indicating an RTLU failure, where prescribed ECAM actions would be part of the sensemaking process. By seeing actions as a way to make sense of the situation, rather than only a resolution activity, may have elicited a more thorough appreciation of the threats and risks of a non-standard circuit breaker reset before engaging in it. This may have reduced the crew's strong *switching* from compliance to "creativity" without validating their adapted mental models before acting. Furthermore, an improved appraisal of the situation by evaluating flight continuation options (e.g., could the flight proceed to the destination or diversion airfield with an unresolved RTLU fault?) may have identified other means to maintain safe flight rather than focusing on RTLU resets. A JCS improvement similar to the Man4Gen RIS (Buch et al., 2017), discussed in subsection 3.2, could help identify the *actual* controllability risks.

5.3 Asiana 214

5.3.1 Synopsis (NTSB, 2014)

A Boeing 777 on final approach descended below the glideslope and impacted the sea wall just before the runway threshold, resulting in three fatalities. During the descent air traffic control requested the crew to maintain a higher speed until closer to the airport. As a result the aircraft few above the glideslope. The crew made successive changes to the aircraft's vertical speed autoflight modes to return to the glideslope, resulting in the use of Flight Level Change (FLCH) mode which is not part of Asiana's approach SOPs. This mode, in combination with a thrust lever manual override to idle thrust, disengaged the autothrust system, putting it into an autothrust "hold" mode without low-speed protection modes. As the aircraft descended toward the glideslope from above, the crew let it recapture the glidepath at around 500ft AMSL, after which the aircraft slowly bled speed as the autothrust was still in a hold mode, resulting in a low energy state which was recognized too late for a successful go-around.

5.3.2 Airmanship 2.0 reflection

The sequence of events and automation responses was clearly traceable without any indication of a system failure. However, the fact that the crew's non-standard use of autoflight and autothrust modes resulted in a situation they themselves could not manage, despite the aircraft being fully serviceable in good weather conditions, illustrates the brittleness that may arise through current crew conditioning (Type 1 anomaly). The investigation also underscores the complexity of current autoflight and autothrust systems in relation to the level of details flight crews are trained on. The foundation of the Airmanship 2.0 *openness to opacity* applies well to this situation, where a general appreciation that the aircraft may feature more complexity than the flight crew's own familiarity with it. Such an approach to the aircraft would be able to elicit a stronger threat evaluation during the approach which could in turn drive effective mitigation such as enhanced autoflight mode monitoring, preparing for a missed approach decision or requesting a more familiar approach profile from air traffic control. This case highlights the pro-active, attitudinal aspect of Airmanship 2.0, rather than its sensemaking aspects highlighted by other cases.

5.4 West Air Sweden 206

5.4.1 Synopsis (SAIA, 2016)

During a cruise flight at night a Canadair CRJ200 with the autoflight system engaged, the left-seat Primary Flight Display (PFD) showed a sudden, very rapid increase in pitch attitude due to a failure of one of the attitude reference systems. This automatically disengaged the autopilot (which is coupled to the left-seat instruments), requiring the crew to take immediate manual control. The PFD initially provided a very brief indication of a “*pitch angle disagreement”* between the left and right PFD. However, as the indicated pitch angle rapidly increased over a certain limit, a PFD *declutter mode* activated and removed this pitch disagreement indication, also adding a “push down” indication on the PFD. The captain obliged immediately, resulting in the aircraft entering an uncontrolled dive and impacting the ground less than two minutes later.

5.4.2 Airmanship 2.0 reflection

This case showcases how system integration complexities can result in undesired, emergent system behaviour. The pitch-induced disconnection, pitch-disagreement indication and declutter functions were all each designed from a safety perspective. However, their interaction likely presented the crew with a particularly opaque situation, exacerbated by a call for immediate action (i.e., a “push down” indication). While similar to the AF447 case in terms of information degradation and reduced autoflight capabilities, this Type 2 case featured greater *true* opacity, showcasing the need for crews to be prepared for (safety) systems to act in unpredictable ways (openness to opacity). However, the similarly with AF447 also lies in the need for flight crew to be aware of the need for personal recovery in light of acute and unexpected developments. Such a capacity would have dampened crew actions to allow for a brief window of sensemaking (e.g., for visual and vestibular sensing) prior to taking action. From a JCS perspective, more graceful system degradations (Woods, 2015) could also cater for a safe transition from automated control to human control and sensemaking.

5.5 Lion Air 610 & Ethiopian Airlines 302

5.5.1 Synopsis (KNKT, 2019; EAIB, 2022)

Two Boeing 737MAX aircraft exhibited very similar failures of its Manoeuvring Characteristics Augmentation System (MCAS). Both investigations suspected malfunctions of the single Angle of Attack (AOA) sensor inputting to the MCAS which subsequently trimmed the aircraft nose down. Crew attempts to pull the nose up resulted in multiple, cumulative MCAS nose-down inputs, until the aircraft pitch was unrecoverable. In both cases the crew was simultaneously confronted with airspeed- and altitude-disagree messages, master caution warnings and a stick shaker response, possibly inducing a startle response. Both crews were unable to correctly execute the non-normal checklist to disengage the pitch trim system. Neither aircraft was equipped with an optional *angle-of-attack disagree* warning light. A few days before the Lion Air Flight 610 accident, Lion Air flight 43 operated the same aircraft as Lion Air 610 and also experiencing the same MCAS pitch down and alerts (KNKT, 2019). This crew was able to operate the pitch trim disconnect switches in time. This recovery took minutes with significant pitch control challenges, rather than several seconds as was assumed in the 737MAX certification program (Senate Committee, 2020). However, as Flight 43 was only an incident it unfortunately was not subject to a full investigative report as the following accident flight.

5.5.2 Airmanship 2.0 reflection

Like the West Air Sweden 206 case, the MCAS system likely contributed to system opacity well beyond the competence of today's aircrew (Type 2 anomaly). In particular, the repeated and cumulative nose-down pitch trimming resulted in a near-reversal of longitudinal control: pulling to raise the nose only led a stronger force pushing it down. The assumption that crews would recover from an MCAS failure by identifying it as a stabilizer runaway doesn't hold as such a runaway provides a loud, continuous trim wheel activation (strong cue), while the MCAS pitch changes were much more subtle and progressive. Lastly, the lack of pilot training, system knowledge and crew notifications likely shrouded a critical control malfunction (true opacity). Both accident cases suffered further from uncoordinated crew responses, likely startle induced and possibly even fear-potentiated (Field et al., 2018) due to progressive nose down trimming. The Lion Air crew did not communicate about the control difficulties they were struggling with (possibly a freeze response), and the Ethiopian Airlines crew repeatedly attempted to engage the autoflight system rather than responding to the warnings and alerts they were experiencing (possibly a fixed mental model). The Airmanship 2.0 concept may have improved outcomes through prioritizing (some) cognitive recovery and establishing some action feedback as part of a sensemaking process, rather than repeating the same actions (e.g., trying to reactivate the autoflight system) and hoping for a different outcome. A more salient report about the preceding Lion Air Flight 43 incident may have provided a more insight into the crew’s experiences, mental models and resolution of MCAS opacity. From a JCS perspective, greater transparency in systems goes a long way in helping flight crews differentiate between multiple situations with similar symptoms. Since these accidents, the AOA disagree warning indication has become standard rather than optional (Boeing, 2019), which is a small step in this direction.

5.6 Quantas 32

5.6.1 Synopsis (ATSB, 2013)

An Airbus A380 experienced an uncontained engine failure in cruise flight. A separated fan blade damaged multiple hydraulic, electric and structural systems in the left wing, causing malfunctions in flight-, engine- and fuel control systems with the ECAM displaying 53 faults. In comparison, pilot training scenarios deemed challenging by proficient crews may feature only two or three ECAM faults. By coincidence, this flight featured three additional experienced crew members performing trainer qualification checks. After two and a half hours in the vicinity of Singapore diagnosing the state of the aircraft, the crew managed to land successfully with one engine still uncontrollable near max thrust, reduced hydraulic braking, failed antiskid and only thirty minutes until a lateral fuel imbalance would have resulted in complete loss of control.

5.6.2 Airmanship 2.0 reflection

This case qualifies as a true “black swan” (Taleb, 2015) event as it was considerably less probable than current certification standards for catastrophic events (10^{-9}) as per EASA AMC CS25.1309-7 (EASA, 2023). According to Captain De Crespigny the likelihood was nearer to 10^{-14} (De Crespigny, 2012). Unfortunately, the ATSB report is very thin on human performance, merely stating the crew worked as a team and followed SOPs (ATSB, 2013). De Crespigny’s own account (De Crespigny, 2012) features much more detail about the crew’s cognitive processes and experiences. Despite the considerable opacity and ambiguity experienced by the crew, the crew’s ability to diagnose system states, create an *adapted* mental model of the aircraft and identify a novel recovery plan is a strong example of Airmanship 2.0. Their use of ECAM and other procedures as *part* of their sensemaking process drove efficient and effective adaptation of their mental model of the aircraft’s capabilities. Ultimately, reducing the opacity by clearly establishing *what functionality they still had* (rather than focusing on what failed) most likely gave them a much more actionable sense of the situation. This is a great example of an adaptive mental model, which is reconstructed bottom up, rather than maintaining mental models of systems and the aircraft which are not accurate in that context. A final key element is effective time management to facilitate sensemaking. In line with recommendations and results from Mohrmann et al. (2015) and Mohrmann (2026), managing short term planning to create time (e.g., by entering a holding pattern) provides both temporal and cognitive space for sensemaking, reconfiguring and mitigating downstream options and threats. In effect, time management is a precursor to most of the Airmanship 2.0 principles: allowing time to recovery cognitively, allowing time to think/reflect and allowing time to perform actions and reflect on their impact. Other accidents such as Lion Air 610 and West Air Sweden 206 also provide similar contexts where building in time may have resulted in better outcomes.

It must be mentioned that the above accidents may have been (partially) mitigated with other non-aircrew countermeasures such as improved safety management (e.g., preceding instances of MCAS and frozen pitot tubes) or automation interaction (e.g., crew notifications). However, such hindsight does not generate sufficient technical foresight to void the requirement for resilient flight crews. Opacity and ambiguity are already a reality and will increasingly be part and parcel of modern flight operations. A reappraisal of the pilot role is essential to mitigate what cannot always be anticipated (without an accident to point it out).

6 Discussion

This paper revisits basic premises of human-automation teaming in complex, safety-critical systems to reduce systemic brittleness. It brings to the surface critical underlying assumptions and non-sequiturs in introducing adaptive pilot responses in a currently compliance-centred paradigm. The Airmanship 2.0 concept is proposed, which leverages the particularly human capacity of sensemaking to adapt in the face of opacity and ambiguity. It systematically retains existing compliance safety performance by nesting existing workload-reducing behaviours (e.g., procedures) within the process of sensemaking. By defining a singular, compound role, the concept aims to (1) prevent pilot skill decay in both compliance and adaptivity, and (2) avoid straining pilots in role-bifurcation. Successful introduction of this concept will likely require a more succinct analysis and appreciation of the various JCS interactions and automation roles potentially supporting a pilot model based on Airmanship 2.0, as well as the potential impact on flight operations and aircrew training.

The Airmanship 2.0 concept is relevant for today's Type 1 and Type 2 brittleness anomalies, but even more so for two significant developments on the horizon. The first is Reduced Crew Operations (RCO) exploring a single-pilot JCS (Sprengart et al., 2018; Schmid & Stanton, 2020) as well as concepts with ground-based pilots (Niedermeier et al., 2023). These concepts explore shifting more tasks to automated systems (Shively et al., 2016; Ho et al., 2017), further propagating Klein's self-reinforcing complexity loop and potentially increasing the risk of Type 1 brittleness anomalies. Harris (2023) also calls for the reappraisal of the pilot role for RCO concepts, shifting from executive activities to strategic, decision-making activities. RCO potentially also increases the likelihood of Type 2 anomalies as the effective cognitive capacity for sensemaking in the flight deck has reduced by (at least) half. Furthermore, RCO potentially affects the sensemaking response by removing valuable human interaction during 1) cognitive recovery (i.e., startle management crew cross-checking as described by Field et al. (2018)), 2) the mutual challenging of personal biases and 3) mitigation of (cognitive) blind spots (Sarter et al., 2023). This is not to say that RCO should not be explored, but it does underscore the need to explicitly design required crew behaviours (including adaptive responses) into a new (Airmanship 2.0) JCS, rather than inheriting the modus operandi from existing multi-crew flight decks.

The second development on the horizon is the introduction of AI-based systems in aviation (EASA, 2023; FAA, 2024). The potential for significant autonomy of automated systems (Kirwan, 2025) brings with it considerable risk of opacity and complexity and increased risk of Type 2 anomalies. This in part is driving robust and paced integration as proposed in EASA and FAA AI roadmaps. Research into human-AI teaming underscores limitations of AI to fully replicate the human ability of causal mental modelling (Sarter et al., 2023) through cognitive processes such as critical thinking, assumption testing, and intuition. As AI agents enter the flight deck JCS, a reappraisal of the pilot role toward such cognitive capacities is well called for. This may be particularly relevant in combination with RCO concepts (Minaskan et al., 2022). In contrast, the Airmanship 2.0 concept may also serve as an operator model for other safety-critical systems such as Air Traffic Management (ATM), which faces the integration of unmanned aviation and autonomous non-piloted systems together with increased levels of automation, possibly requiring a reappraisal of the human role in traffic management (Malakis et al., 2023; Zou, 2025).

As a final reflection: explicating the brittleness problem, its presumptive anomalies, a revised pilot role and future developments around RCO and AI raises the question: how the industry can identify in time when Airmanship 2.0 also reaches its limits? When would we need Airmanship *3.0*? This would be the hallmark of *industry-level resilience*. Kirwan (2025) provides a clear histological overview showing how human factors methods have evolved as the flight deck's JCS evolved. Traditional models such as Human Factors Analysis and Classification System (HFACS) don't provide sufficient resolution to address the cognitive and behavioural nuances that drive today's JCS performance. More fundamentally, the research field of "human" factors may be too narrowly scoped for these nuances, possibly requiring an evolution toward a systems perspective at the JCS level as suggested by Woods (2019).

One attempt to proactively defend against future blind spots is the System Theoretic Accident Model and Processes (STAMP) approach, and in particular its applications to design with Systems-Theoretic Process Analysis (STPA), and to accident investigation with Causal Analysis Using System Theory (CAST) (Leveson, 2016). By mapping entire systems "as they are", the STAMP approach attempts to uncover interactions, behaviours and sensitivities not visible to frameworks scoped around the human. In an industry with such a high level of safety, reactive incident- and accident-led learning may not provide sufficient evidence to understand modern flight deck JCS criticalities and sensitivities in time. Rather, the expansion of scenario-based training through CBTA (ICAO, 2013; IATA 2024) may provide the requisite immersion and operational relevance to better understand these cognitive nuances in the JCS. As a first step in this direction, the IATA white paper on EBT already proposed that the pilot competency framework replace the existing flight crew countermeasure safety taxonomies (IATA, 2024). As the near future heralds radical operational concepts for tomorrow's flight decks, the industry must be more vigilant than ever on the continued validity of design assumptions inherited from the past.

7 References

Abbott, K. H. (2017). Human factors engineering and flight deck design. *Digital avionics handbook*, 302-328.

ACROSS consortium. (2016). Final Report Summary – ACROSS (Advanced Flight deck for Reduction Of StreSs and workload).

Adanov, L., Efthymiou, M., & Macintyre, A. (2020). An exploratory study of pilot training and recruitment in Europe. *International Journal of Aviation Science and Technology*, *1*(02), 44-51.

Airbus. (2025). A Statistical Analysis of Commercial Aviation Accidents 1958 - 2024. Blagnac, France.

Australian Transportation Safety Bureau (ATSB). (2013). In-flight uncontained engine failure Airbus A380-842, VH-OQA. https://www.atsb.gov.au/publications/investigation_reports/2010/aair/ao-2010-089

Bainbridge, L. (1983). Ironies of Automation. *Automatica*, 19, 775-779.

Banks, V. A., Plant, K. L., & Stanton, N. A. (2020). Leaps and shunts: designing pilot decision aids on the flight deck using Rasmussen's ladder. *Contemporary ergonomics and human factors*.

Boeing, (2019). Pilot & Technician Outlook 2019-2038.

Boeing. (2019). Boeing statement on AOA disagree alert [Press release]. https://boeing.mediaroom.com/news-releases-statements?item=130431

Bowles, G. (2018). Automation Readiness Levels. *Personal Communication*.

Boy, G. A. (2020). Aerospace human systems integration: Evolution over the last 40 years. *A framework of human systems engineering: Applications and case studies*, 113-128.

Bradshaw, J. M., Hoffman, R. R., Woods, D. D., & Johnson, M. (2013). The seven deadly myths of" autonomous systems". *IEEE Intelligent Systems*, *28*(3), 54-61.

Buch, J. P., Niedermeier, D., & Stepniczka, I. (2017). Managing the Unexpected-Human-in-the-Loop Simulation as Effective Tool for the Assessment of the Risk Information System in an Operationally Relevant Context. In *AIAA Modelling and Simulation Technologies Conference* (p. 4155).

Bureau d'Enquêtes et d'Analyses (BEA). (2012). Final Report On the accident on 1st June 2009 to the Airbus A330-203 registered F-GZC operated by Air France flight AF 447 Rio de Janeiro – Paris. https://bea.aero/fileadmin/documents/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf

Cahill, J., Callari, T. C., Fortmann, F., Suck, S., Javaux, D., Hasselberg, A., ... & van Doorn, B. A. (2017, June). Adaptive automation and the third pilot: managing teamwork and workload in an airline flight deck. In *International Symposium on Human Mental Workload: Models and Applications* (pp. 161-173). Cham: Springer International Publishing.

Caraway, C. L. (2020). A looming pilot shortage: It is time to revisit regulations. *International Journal of Aviation, Aeronautics, and Aerospace, 7(2), 3.*

Clark, A. N., & Wilson, A. L. (2024). From Levers to Glass: The Evolution of Aircraft Flight decks and the Future of Aviation Innovation.

Constant II, E. W. (1973). A model for technological change applied to the turbojet revolution. *Technology and Culture, 14(4), 553-572.*

Crouch, V. (2020). Analysis of the airline pilot shortage. *Scientia et Humanitas, 10, 93-106.*

Cullen, P., Cahill, J., & Gaynor, K. (2021). A qualitative study exploring well-being and the potential impact of work-related stress among commercial airline pilots*. Aviation Psychology and Applied Human Factors.*

De Crespigny, R. (2012). *QF32*. Macmillan.

Degani, A., & Wiener, E. L. (1994). *On the design of flight-deck procedures* (No. A-94095).

Enea, G., & Porretta, M. (2012, September). A comparison of 4D-trajectory operations envisioned for Nextgen and SESAR, some preliminary findings. In *28th Congress of the International Council of the aeronautical sciences* (Vol. 5, pp. 4152-4165). Edinburgh, UK: Optimage Ltd.

Eurocontrol. (2020). The FLY AI Report: Demystifying and accelerating AI in Aviation/ATM.

European Union Aviation Safety Agency (EASA). (2020). Easy access rules for flight crew licencing (Part-FCL) (June 2020 version).

European Union Aviation Safety Agency (EASA). (2023). Artificial intelligence roadmap 2.0.: Human-centric approach to AI in aviation. Version 2.0. www.easa.europa.eu/ai

European Union Aviation Safety Agency (EASA). (2023). Certification Standards Part 25.

Ethiopia Aircraft Accident Investigation Bureau (EAIB). (2022). Investigation report on accident to the B737-MAX8 reg. ET-AVJ operated by Ethiopian Airlines. https://avherald.com/files/ethiopian_b38m_et-avj_190310_final_report_20221226.pdf

Feary, M. (2018). A first look at the evolution of flight crew requirements for emerging market aircraft (No. ARC-E-DAA-TN57004).

Federal Aviation Administration. (2024, July 25). Roadmap for artificial intelligence safety assurance (Version I). U.S. Department of Transportation.

Field, J., Rankin, A., Mohrmann, J.F.W., Boland, E., & Woltjer, R. (2017). Flexible procedures to deal with complex unexpected events in the flight deck. *In Resilience Engineering Association Symposium Liége, Belgum, 26-29 June, 2017.* Resilience Engineering Association.

Field, J.N., Boland, E.J., Van Rooij, J.M., Mohrmann, J.F.W., Smeltink, J.W., 2018. Startle Effect Management. (Report Nr. NLR-CR-2018-242). European Union Aviation Safety Agency.

Fraher, A. L. (2019). The vulnerability of quasi-professional experts: A study of the changing character of US airline pilots' work. *Economic and Industrial Democracy*, 40(4), 867-889.

FutureSky Safety P6 Consortium (2016). D6.4 Recommendations recovery measures and HMI implementation.

Gago, C. P., Hansman, R. J., Edmondson, M. K., & Mosqueda, M. A. (2025, September). A Study of Pilot Response to System Failure in Transport Category Aircraft from 2000 to 2024. In *2025 AIAA DATC/IEEE 44th Digital Avionics Systems Conference (DASC)* (pp. 1-8). IEEE.

Hancock, P. A., Cruit, J., Kochan, J. A., Kaplan, A. D., Diaz, Y., & Pruchnicki, S. (2022). Pilots' responses to unexpected events: Conceptual, theoretical, methodological, and analytical issues. *The International Journal of Aerospace Psychology*, *32*(4), 254-282.

Harris, D. (2017). *Human factors for civil flight deck design*. Routledge.

Harris, D. (2023). Single-pilot airline operations: Designing the aircraft may be the easy part. *The aeronautical journal*, *127*(1313), 1171-1191.

Ho, N., Johnson, W., Panesar, K., Wakeland, K., Sadler, G., Wilson, N., ... & Brandt, S. (2017, September). Application of human-autonomy teaming to an advanced ground station for reduced crew operations. In *2017 IEEE/AIAA 36th digital avionics systems conference* (DASC) (pp. 1-4). IEEE.

Hollnagel, E., & Woods, D. D. (2005). Joint cognitive systems: Foundations of cognitive systems engineering. CRC Press.

Hörhager, P. (2025). Aircraft Automation and Human Performance: An Analysis of Flight Deck Failures (Doctoral dissertation, Technische Universität Wien).

Huijbrechts, E. J., & Van Paassen, M. M. (2023). Flight deck procedures for a new generation of pilots.

International Air Transport Association (IATA) (2002). Safety Report 2001.

International Air Transport Association (IATA) (2014). Safety Report 2014.

International Air Transport Association (IATA) (2020). Safety Report 2019.

International Air Transport Association (IATA) (2024). Evidence-Based Training Implementation Guide Edition 2.

International Air Transport Association (IATA) (2024). Competency-Based Training and Assessment (CBTA) Expansion within the Aviation System.

International Civil Aviation Organisation (ICAO). (2013). Doc 9995 Manual of Evidence-based Training, First Edition.

International Civil Aviation Organisation (ICAO). (2016). Doc 9966: Manual for the Oversight of Fatigue Management Approaches, Second Edition.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kelly, D., & Efthymiou, M. (2019). An analysis of human factors in fifty controlled flight into terrain aviation accidents from 2007 to 2017. *Journal of Safety Research*, *69*, 155–165.

Kent, P. (2017). Fluid intelligence: A brief history. *Applied Neuropsychology: Child, 6(3), 193-203.*

Kern, T. (2010). Foundations of Professional Airmanship and Flight Discipline. *Convergent Performance*. Colorado Springs, Colorado.

Kharoufah, H., Murray, J., Baxter, G., & Wild, G. (2018). A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016. *Progress in Aerospace Sciences*, *99*, 1–13.

Kirwan, B. (2025). Human factors requirements for human-ai teaming in aviation. *Future Transportation*, *5*(2), 42.

Klein, G. A. (2011). *Streetlights and shadows: Searching for the keys to adaptive decision making*. MIT Press.

Komite Nasional Keselamatan Transportasi (KNKT). (2015). Aircraft Accident Investigation Report PT. Indonesia Air Asia Airbus A320-216; PK-AXC. http://www.aaiu.ie/sites/default/files/FRA/KNKT%20Indonesia%20Final%20Report%20PK-AXC%20Airbus%20A320-216%20Air%20Asia%20PT%20Indonesia%202015-12-01.pdf

Komite Nasional Keselamatan Transportasi (KNKT). (2019). Aircraft Accident Investigation Report PT. Lion Mentari Airlines Boeing 737-8 (MAX); PK-LQP. https://knkt.go.id/Repo/Files/Laporan/Penerbangan/2018/KNKT.18.10.33.04-Final-Report.pdf

Kulida, E., & Lebedev, V. (2020, September). About the use of artificial intelligence methods in aviation. In *2020 13th International Conference" Management of large-scale system development"(MLSD) (pp. 1-5). IEEE.*

Landry, S.J. (2009). Flight Deck Automation. In: *Nof, S. (eds) Springer Handbook of Automation.* Springer Handbooks. Springer, Berlin, Heidelberg.

Landman, A., Groen, E. L., van Paassen, M. M. (René), 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), 1161–1172.

Landman, A., van Middelaar, S. H., Groen, E. L., Van Paassen, M. M., Bronkhorst, 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.

Leveson, N. G. (2016). *Engineering a safer world: Systems thinking applied to safety* (p. 560). The MIT Press.

Li, W. C., Li, L. W., Harris, D., & Hsu, Y. L. (2014). The application of aeronautical decision-making support systems for improving pilots' performance in flight operations.

Ligda, S. V., Fischer, U., Mosier, K., Matessa, M., Battiste, V., & Johnson, W. W. (2015, July). Effectiveness of advanced collaboration tools on crew communication in reduced crew operations. In *International conference on engineering psychology and cognitive ergonomics* (pp. 416-427). Cham: Springer International Publishing.

Lorrig, P., & Daw, Z. (2024, September). Advances and Challenges Towards Enabling Human-AI-Teaming Applications for Flight Deck Operations. In *2024 AIAA DATC/IEEE 43rd Digital Avionics Systems Conference (DASC)* (pp. 1-10). IEEE.

Malakis, S., Baumgartner, M., Berzina, N., Laursen, T., Smoker, A., Poti, A., ... & Kontogiannis, T. (2023, July). A framework for supporting adaptive human-AI teaming in air traffic control. In *International Conference on Human-Computer Interaction* (pp. 320-330). Cham: Springer Nature Switzerland.

Malinge, Y. (2004). A300B4 loss of all hydraulics, Baghdad. International Society of Air Safety Investigators (ISASI) Annual Seminar.

Malinge, Y. (2011). The views from an aircraft manufacturer: What have we learned.

Malik, A., & Gollnick, V. (2016). Impact of reduced crew operations on airlines-Operational challenges and cost benefits. In *16th AIAA aviation technology, integration, and operations conference* (p. 3303).

Man4Gen Consortium. (2012). D1.1 literature review. Technical report, Man4Gen Consortium. Retrieved from www.man4gen.eu.

Man4Gen Consortium (2015). 6.7 - Final report of research methods and results.

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human performance consequences of automated decision aids: The impact of degree of automation and system experience. *Journal of Cognitive Engineering and Decision Makin*g, 1555343411433844.

Martin, W., Murray, P., & Bates, P. (2012). The effects of startle on pilots during critical events: A case study analysis. In *EAAP 30*. EAAP.

Matthews, G., Emo, A. K., Funke, G., Zeidner, M., & Roberts, R. D. (2003, October). Emotional intelligence: Implications for human factors. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 47, No. 9, pp. 1053-1057).* Sage CA: Los Angeles, CA: SAGE Publications.

Maxwell, G. A., & Grant, K. (2021). Commercial airline pilots' declining professional standing and increasing precarious employment. *The International Journal of Human Resource Management*, *32*(7), 1486-1508.

Merritt, A. C., & Klinect, J. R. (2006). *Defensive flying for pilots: an introduction to threat and error management*. University of Texas Human Factors Research Project, The LOSA Collaborative.

Minaskan, N., Alban-Dromoy, C., Pagani, A., Andre, J. M., & Stricker, D. (2022, June). Human intelligent machine teaming in single pilot operation: a case study. In *International Conference on Human-Computer Interaction* (pp. 348-360). Cham: Springer International Publishing.

Mohrmann, J.F.W., Lemmers, A., & Stoop, J. (2015). Investigating flight crew recovery capabilities regarding system failures in highly automated fourth generation aircraft. *Aviation Psychology and Applied Human Factors.*

Mohrmann, J.F.W., Field, J.N. & Stoop, J. (2017). Resilient decision making in the flight deck: Does it work?. *Proceedings from Resilience Engineering Association Symposium Liége, Belgum, 26-29 june, 2017.* Resilience Engineering Association.

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. (1998). Automation bias: Decision making and performance in high-tech flight decks. *The International journal of aviation psychology*, 8(1), 47-63.

Mumaw, R. J., Sarter, N., & Wickens, C. D. (2001, May). Analysis of pilots' monitoring and performance on an automated flight deck. In *11th International Symposium on Aviation Psychology, Columbus, OH* (Vol. 8).

National Transportation Safety Board (NTSB). (2014). Descent Below Visual Glidepath and Impact with Seawall, Asiana Airlines Flight 214, Boeing 777-200ER, HL7742, San Francisco, California, July 6, 2013. Aircraft Accident Report NTSB/AAR-14/01. Washington, DC. https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR1401.pdf

Neisser, U. (1976). *Cognition and Reality*. W. H. Freemand.

Niedermeier, D., Papenfuß, A., & Wies, M. (2023). Simulator study on a spatially separated airline crew in a complex decision-making scenario. In *AIAA AVIATION 2023 Forum* (p. 3474).

O'Hare, D. (1992). The 'artful' decision maker: A framework model for aeronautical decision making. *The international journal of aviation psychology*, *2*(3), 175-191.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans, 30(3), 286-297.*

Prinzel, L. , Krois, P., Ellis, K., Vincent, M., Stephens, C., Oza, N., Chancey, E., Davies, M., Mah, R., Ackerson, J., Infeld, S., Kiggins, D., & Matthews, B. (2024). *The Adaptable and Resilient Safety System: The Human Factor in Future In-Time Aviation Safety Management Systems*.

Proctor, R. W., & van Zandt, T. (2018). *Human Factors in Simple and Complex Systems, Third Edition*.

Rankin, A., Woltjer, R., & Field, J. (2016). Sensemaking following surprise in the flight deck—a re-framing problem. *Cognition, Technology & Work*, *18*(4), 623-642.

Sallinen, M., van Dijk, H., Aeschbach, D., Maij, A., & Åkerstedt, T. (2020). A large-scale European Union study of aircrew fatigue during long night and disruptive duties. *Aerospace medicine and human performance, 91(8), 628-635.*

Sarter, N. B., & Woods, D. D. (1994). Pilot Interaction With Cockpit Automation 11: An Experimental Study of Pilots' Model and Awareness of the Flight Management System. In *THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY* (Vol. 4, Issue 1).

Sarter, N., Panesar, K., & Bhardwaj, A. (2023). The Contributions of Human Operators to Safety and Risk Mitigation: Implications for Crew Complements and Automation/Autonomy Levels in Commercial Transport Operations (No. DOT/FAA/TC-22/40). United States. Department of Transportation. Federal Aviation Administration. William J. Hughes Technical Center.

Sarter, N. (2024). Information Management on the Flight Deck of Highly Automated Aircraft.

Saurin, T. A., & Carim Junior, G. C. (2012). A framework for identifying and analyzing sources of resilience and brittleness: A case study of two air taxi carriers. *International Journal of Industrial Ergonomics*, *42*(3), 312–324.

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

Schutte, P. C. (2015, July). How to make the most of your human: Design considerations for single pilot operations. In *International Conference on Engineering Psychology and Cognitive Ergonomics* (pp. 480-491). Cham: Springer International Publishing.

Senate Committee (2020). US Senate Committee Investigation Report – Aviation Safety Oversight.

Shah, D., (2024). Timesaving: Why Airlines Operate Overnight Flights. *Simply Flying.* https://simpleflying.com/time-saving-airlines-overnight-flights/

Shively, R.J., Brandt, S. L., Lachter, J., Matessa, M., Sadler, G., & Battiste, H. (2016, June). Application of human-autonomy teaming (HAT) patterns to reduced crew operations (RCO). In *International conference on engineering psychology and cognitive ergonomics* (pp. 244-255). Cham: Springer International Publishing.

Soo, K. (2018). The learning and development process of pilots during initial airline training.

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

Stanton, N. A., Harris, D., & Starr, A. (2016). The future flight deck: Modelling dual, single and distributed crewing options. *Applied Ergonomics*, *53*, 331-342.

Stoop, J. A., & van Kleef, E. A. (2015). Reliable or Resilient: Recovery from the Unanticipated. In *International Journal of Performability Engineering* (Vol. 11, Issue 2).

Stoop, J. (2017). How did aviation become so safe, and beyond? *In Proceedings of the 53rd ESReDA Seminar, 14 – 15 November 2017: European Commission Joint Research Centre, Ispra, Italy*

Strauch, B. (2017). *Investigating Human Error*. CRC Press.

Swedish Accident Investigation Authority (SAIA). (2016). Final report RL2016:11e Accident in Oajevágge, Norrbotten County, Sweden on 8 January 2016 involving the aeroplane SE-DUX. https://shk.se/download/18.2d6f089b18faca29dc81cffa/1700491384504/RL-2016_11e.pdf

Taleb, N. (2015). The black swan: The impact of the highly improbable. Victoria, 250, 595-7955.

UK Civil Aviation Authority (UK CAA). (2019). Pilot Fatigue Measurement Research Report.

Valenta, V. (2018). Effects of airline industry growth on pilot training. *MAD-Magazine of Aviation Development, 6(4), 52-56.*

Vincenti, W., (1990). *What Engineers Know and How They Know It. Analytical Studies from Aeronautical History*. The John Hopkins University Press.

Vlaskamp, D., Landman, H. M., van Rooij, J., & Blundell, J. (2025). Recovery from startle and surprise.

Wikander, R., & Dahlström, N. (2016). The multi crew pilot licence-revolution, evolution or not even a solution. *A review and analysis of the emergence, current situation and future of the multi-crew pilot licence (MPL). Lund University*.

Woltjer, R., Field, J., & Rankin, A. (2015). Adapting to the unexpected in the flight deck. In *Proceedings of the 6th resilience engineering association symposium. Lisbon, Portugal: REA*.

Woods, D. D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. *Reliability engineering & system safety, 141, 5-9.*

Woods, D. D. (2019). Essentials of resilience, revisited. *Handbook on resilience of socio-technical systems*, 52-65.

Ziegler, M., Danay, E., Heene, M., Asendorpf, J., & Bühner, M. (2012). Openness, fluid intelligence, and crystallized intelligence: Toward an integrative model*. Journal of Research in Personality, 46(2), 173-183.*

Zou, Y. (2025). Transparent Path Planning For Uncrewed Air Traffic Management.