Airmanship 2.0: Innovating aviation human factors forensics to necessarily proactive role

Submission to: International Society of Aviation Safety Investigators (ISASI) Annual Seminar 2019

Frederik Mohrmann
Netherlands Aerospace Centre (NLR)
j.f.w.mohrmann@gmail.com
Kritzingerstraat 31
2021SG
Haarlem
The Netherlands

John Stoop
Amsterdam Aviation Academy (HvA)
stoop@kindunos.nl
Spijksedijk 8
4207 GN
Gorinchem
The Netherlands

Frederik Mohrmann *Is an aerospace engineer specialized in safety and operations. At the Netherlands Aerospace Center (NLR) he works two roles: as aviation training R&D engineer he participates in and leads civil aviation training research activities including Man4Gen and FutureSky Safety. His second role as a lean innovation coach is actively bridging R&D with industry and operational challenges. He has completed an internship at the NTSB, is recipient of the 2012 Rudolph Kapustin Award and in his spare time he is a glider and aerobatics instructor. At present he is doing a PhD on the topic of Airmanship 2.0.*

John Stoop *is an aerospace engineer and safety investigator. He is a guest professor in forensic engineering and safety investigation at Lund University (Sweden), Delft University and Amsterdam University of Applied Sciences (the Netherlands). He has experience in conducting investigations in multiple modes of transportation and developed safety integrated concepts for several major infrastructural projects. He has been the secretary of ITSA for about 12 years. He was involved in the foundation of the Dutch Transport Safety Board. Currently he is writing a book for AIAA on the Next Generation of air safety investigation methodology in cooperation with Frederik Mohrmann.*

Abstract

It is no surprise to most that the aviation sector is experiencing a shift in accident modes and causal factors. This illustrates itself in the very recent examples of the fatal Boeing 737 MAX accidents (still pending completed investigations), and less recent notable incidents such as an Emirate 777 go-around accident, Air Asia's loss of an A320 in 2014, UPS Flight 1354, Air France 447 and last but not least Quantas Flight QF32. There are other less known but equally troubling cases in which modern, well equipped aircraft flown by air crew trained to legally required standards still result in deadly accidents.

Pending of course the outcome of the 737 MAX investigations, many recent accidents and incidents show a shift in accident causal factors. At the heart lies the overarching theme that increasing complexity of air operations introduces new emergent system behaviour not always anticipatable and preventable by design. As such, maintaining a resilient air transport system demands more cognitively flexible and adaptable flight crew. However, these new higher-order competencies are in turn more strongly affected by fatigue and startle/surprise factors, accounting for the increased attention that these two phenomena are receiving. To boot, the industry has put pressure on all three fields: pilot demand driving down training efforts, fatigue on the rise due to circadian irregularity and extended work hours, and more reliable systems inducing automation complacency which exacerbates startle and surprise potential.

As such, the current linear, Taylorist human factors investigation techniques, often focussed on error and non-compliance, are not suitable to capture the new dynamics of human performance that present themselves in this era of an ever more complex aviation system. In order to understand and mitigate these new emergent system behaviours, human factors forensics requires a shift in attitude, complementing existing error-analysis with a more in-depth pilot accounts, experiences and reasoning. More that the cockpit microphone can record, human factors forensics must shift to more pro-active methods of investigation to capture this.

Recent research projects at the Netherlands Aerospace Center (NLR) have investigated these new human factors, including investigating the effects of fatigue, startle/surprise in modern, complex incident and accident modalities, as well as evaluating the potential of countermeasures such as startle effect management training, mindfulness training and new complexity-oriented problem-solving strategies. Besides developing effective countermeasures, these projects have also developed new methods of forensic investigation of such scenarios.

This contribution will present several practical forensic methods used in these pro-active, simulated investigations collected over the past years, supplementing a new, future-proof human factors investigation model. Both investigative methods as well as new cornerstones of human-machine interaction and human effectiveness will be presented. Some insight into developed mitigating methods will also be presented as positive outcomes of these proactive investigations. Hopefully, this may help investigators identify the more subtle emerging behaviours of this increasingly complex system, before a new accident has to be its herald.

Introduction: outliers or trends?

Per 2019 the worldwide commercial air transport system is one of the highest performing systems in the history of humankind. It is almost unfathomable how we can transport 4.1 billion passengers (2017 figures) at an accident rate of 1 per 6.7 million flights, with 0 accidents per 2017 for IATA members. We speak of a non-plus-ultra-safe system approaching the mythical 10^{-7} accident rate, a theoretical safety performance limit [1]. And yet we still have accidents, and these are not minor events in unforgiving circumstances far beyond our wildest dreams, caught in a flock of black swans or other unknown unknown bird. A few telling examples:

Air France 447

During a flight from Rio de Janeiro to Paris, an Air France A330 experienced icing on the pitot tubes, resulting in inaccurate air data. The autopilot disconnected promptly, with the aircraft control law switching from normal law to an alternate law, followed by the autothrust system disengaging. The first and second officers which were on duty during this event, were unable to maintain control of the aircraft which was operating at an altitude where overspeed and stall speeds approach each other, in addition to a total lack of airspeed information and a control law that is seldom flown with. The resulting prolonged stall in a dark and clouded surrounding was not detected as such, and resulted in the aircraft crashing into the ocean [2].

Quantas 2010

After departing from Singapore airport, a Quantas Airbus 380 experienced an uncontained engine failure of the number two engine, with a separated fan blade damaging multiple hydraulic, electric and structural systems along the left wing spar. This resulted in the malfunction of flight control systems, engine control systems and fuel control systems, along with an incredible 53 ECAM failure messages. Coincidently this flight featured five flight crew members as both a line and instructor check was being performed on this flight. After two and a half hours of diagnosing the state and ability of the aircraft, the crew managed to land the aircraft successfully as Singapore with one engine still operating near max thrust, with limited hydraulic braking available, failed antiskid and only a half our until the lateral fuel imbalance would no longer permit controlled flight. The ability of the flight crew to abstract the aircraft state, override system procedure requirements and distill a unique landing performance strategy was critical to the prevention of an accident on par with Tenerife [3].

Asiana 2013

While landing at San Francisco International Airport, an Asiana Boeing 777 performing an autolanding failed to maintain sufficient airspeed during the approach and descended below the glideslope as a result. This resulted in the aircraft touching down on the seawall just prior to the runway threshold. The reason for the reduction in airspeed was the disengagement of the autothrust system, with the low thrust settings used in a higher section of the approach, when the aircraft was above glideslope, being held until seconds before the impact, when the crew made a late attempt to recover engine thrust. The reason for an incorrect auto thrust setting was related to the crew's insufficient understanding of the complex auto thrust modes and their (dis)engagement conditions [4].

Air Asia 2014

During a flight from Surabaya to Singapore, An Air Asia Airbus 320 lost control of the aircraft and crashed into the sea. This specific aircraft experienced a malfunction of the Rudder Traveler Limiter Unit (RLTU) 23 times in the previous year, later attributed to a cracked soldered connection. During this flight, the system presented an RTLU failure four times. Normal procedures require resetting the RTLU, yet this time the captain elected to reset the two Flight Augmentation Computers (FAC), which removed flight control augmentation and placed the aircraft in alternate law (manual control). Shortly after the aircraft entered a stall with the report quoting that: "[The pilots] would have to rely on manual flying skills that are often stretched during a sudden airborne emergency." [5].

West Air Sweden 2016

During an uneventful night flight over Sweden, a West Air Sweden Canadair CRJ200 experienced a mid-flight loss of control resulting in a crash near lake Akkajaure. During cruise with the autoflight system engaged, the left Primary Flight Display (PFD) showed a rapid increase in pitch attitude, which was due to a failure of one of the attitude reference systems. This increase in pitch disengaged the autopilot system which was coupled to the captain's (left hand) instruments, requiring the captain to take immediate manual control of the flight. The PFD initially provided an indication of a pitch mismatch, however this was removed when a declutter mode of the PFD was activated as the (incorrectly) indicated pitch exceeded , additionally prompting the captain for a nose-down control action. The captain obliged immediately, not having any strong outside references, however as the flight was still straight and level, the aircraft entered an uncontrolled dive, resulting in impact with ground less than two minutes later. The crew's inability to detect the pitch mismatch, the systems rapid degradation of information to detect it and the crews rapid actions resulted in fatal loss of control [6].

737MAX 2018/2019

Two accidents only months apart, involving a Lion Air Boeing 737MAX and an Ethiopian Airlines 737MAX both featured similar malfunctions of the Maneuvering Characteristics Augmentation System (MCAS) which uses an angle of attack sensor to prevent impending aircraft stall by providing inputs into the flight control augmentation system. In both events, the MCAS system provided multiple strong nose-down control inputs as the aircraft were climbing out after takeoff. The MCAS system was a new system modification introduced into the 737MAX without explicit training to the pilots. MCAS controls pitch using automatic elevator trim controls, and a failure of the system should be treated as an elevator pitch trim runaway/failure. However, the failure of the MCAS was not part of the conversion training to the 737MAX, and as such provided the flight crews with a failure they were not immediately proficient in. the Lion Air aircraft experienced an MCAS failure on the previous flight, resulting in an extreme flight profile that was only resolved after a third pilot suggested the Stabilizer Trim Runaway procedure, although there were no indications that this was the problem. Unfortunately the next flight experienced the same problem but was not able to resolve the problem in time. The investigations are still ongoing, yet the incidents show a start similarity to the previous four accident examples. As a result of these accidents, the 737MAX has been grounded worldwide pending investigation [7] [8].

There are plenty more examples that the last ten years can provide, with similar accounts where the existing socio-technical system design of our cockpit operations still lead to a (sometimes fatal) accident, or some lesser undesired state from which we still managed to recover. Currently, most of these accidents are categorized into basic accident categories (with "Controlled flight into terrain leading the list) and with a pilot's negative contribution to a situation assessed as "pilot error", with the occasional organizational culture issue at its flank. As a result, our industry has been quite focused on "how to prevent controlled flight into terrain?", developing advisory circulars, updated upset recovery and flight envelop systems and adapted flight manuals.

However, are we as an industry truly convinced that we have found the root of the problem as we band-aid our operations in the wake of these accidents, or have we reached a point in which we realize we are only treating the symptoms of a larger issue at hand? Vincente proposes a possible explanation of these accidents as *presumptive anomalies*:

"Presumptive anomalies occur in technology, not when the conventional system fails in an absolute or objective sense, but when assumptions derived from science indicated either that under some future conditions the conventional system will fail (or function badly) or that a radically different system will do a much better job." [9]

We should consider the possibility that the industry is experiencing presumptive anomalies in the pilot-automation socio-technical system, who still operates as "designed" (trained), but that the conditions in which this design works has changed. We may not know yet for sure, but we cannot afford to make such global assumptions. If we are not willing to assess whether our operational paradigm has reached its limits, we will not be able to improve performance. As a result, as the commercial air transport industry grows at a rate of 6-7% a year, so will the number of global accidents as long as the accident rate stagnates around one accident per 4-6 million flights. If the number of flights increases by 6% and the accident rate in 2019 is one accident per 5 million flights, we will have 22 accidents per year in 2036, up from 8 accidents in 2018. Which means we read about a major accident worldwide every two weeks, which will not answer for the public, especially in this age of information exchange and social media amplifying the public outcry against such accidents. As said before, this industry cannot afford such stagnation and must be willing to examine its limits to induce effective innovation in safety.

At the forefront of this assumption testing stands the aviation safety investigator, a professional which has learned to eat assumptions for breakfast. However, in order to challenge an existing paradigm, the investigator must have investigative tools from a contrasting paradigm to gain perspective. Put anecdotally:

"If all you have is a hammer, everything looks like a nail." Abraham Maslow

Put in context: if all you have is a human-error band-aid, everything looks like human error. This paper wishes to empower safety investigators with a new perspective on the human-machine paradigm, and give insight into tools and methods that may facilitate innovation in human factors forensics. In 2012, BEA director Jean-Paul Troadec reflected on the investigation of the well-known Air France 447 accident in 2009, indicated the possible limits of the current HF forensic methods:

“This accident has also taught us that hypotheses used for safety analyses are not always relevant, that procedures are not always applied and that warnings are not always perceived. Only an improvement in the quality of feedback will make it possible to detect any weaknesses in the safety model.” [10]

Observing this from a broader perspective, the major implication of innovating the way we investigate our operational paradigm, is that investigators are re-empowered in a proactive role to providing direction for a whole new level of safety performance of our industry.

The limits of our current human factors model

The Taylorist roots of our current pilot model

In the beginning, becoming a pilot placed oneself right up there with the most daring, courageous and progressive people of the time. Flying was a considerably dangerous undertaking, and required pilots to have a more daring, innovative and can-do mentality to weather the technical uncertainties, atmospheric conditions only partially understood and other challenges such as physiological limits, limited infrastructure and a rapidly changing industry as a whole.

However, as the industry developed, flight operations were increasing more predictable, and at a fundamental level the “pilot” role in the system (usually fulfilled by an actual human being) was subjected to rigorous “Taylorisation”, stemming from the work-principles of Frederick Taylor: efficiency-driven principles developed during the Industrial Revolution. Operational features such as checklists, standard operating procedures, compliance checks, prescribed training programs and sim-/line-checks are all manifestations of the industry’s thoroughness in prescribing operations. And for good reason: the adventurous, authoritative and risk-defying attitude that was necessary in the early days of aviation (up until the 1950’s) was no longer conducive to the operations which were by then readily predictable, regulatable and in need of consistent pilot performance.

The human operator moved from an “aviator” role toward a very useful element in our cockpit system, as he/she could be programmed for a multitude of tasks which could not be automated (at the time). However, slowly but surely the *human* pilot was sought to behave as a *mechanical* element, but one that you could talk to and give more complex tasks than the mechanics of the time. All this drastically improved aviation safety as accidents caused by ego-driven attitudes, slips & lapses, aircraft limit exceedances and incorrect failure management slowly receded into the past. Basic and recurrent flight training steered toward conditioning new and old pilots in their predictability and compliance, for example with the concept of “checks” being hard-coded in pilot development.

The promise of predictability rooted itself very deep within out cockpit design philosophies. So much so that in the 1980’s and 1990’s, the natural variance in pilot behavior was regularly and rigidly held accountable against the prescribed task-sets and behaviors pilots were trained and told to do. This “deviant behavior” was labeled as “pilot error”. Driven to reduce the influence that such pilot behavior variance had on safety, third and fourth generation commercial aircraft heralded new cockpit technologies such as integrated systems, fly-by-wire, systems, centralized failure management, etc., to offload more and more tasks from the pilot and shift them toward increasingly more capable mechanical system elements. As a result, the pilot’s role slowly shifted from directly executing the flight toward overseeing systems which executed, coordinate with other operational stakeholders (ATC, aircraft, company, ground handling, passengers, etc.) and to serve as a flexible system element which offered redundancy for most if not all tasks executed by automation. Offloading to automated system elements was further promulgated as our industry felt more and more confident in an operation we could carefully map out, plan, prescribe: the ultimate determinist system.

Determinism at the core of our human factors forensics

Our basic model of the human operator in the cockpit serves as a foundation of our human factors forensics (HF forensics). The thoroughly prescribed tasks and behaviours for the pilots in our operations serve as a very rigid reference against which actions of a pilot can be benchmarked. The early days of human factors forensics were quick to denounce aberrant pilot behaviour into several broad categories: wilful non-compliance and human error, the latter being subjugated to further categorisation in subsequent decades, including slips, lapses, various forms of complacency, decision errors, perception errors, crew interaction errors and more recently organisational cultures and processes. The HFAC-System provides sufficient examples that illustrate the current scope of investigation, but the same determinist base can be observed in the ICAO Annex 13 investigation guidance, which lacks a large set of classifications of physiological, behavioural and social factors that can be found in a growing number of recent accidents (by way of reference those in the introduction of this paper). As an illustrative example, within the European ECCAIRS accident database, it is not possible to search for accidents that may have surprise or startle effects as a contributory factor, whilst that very topic is taking centre stage in safety discussions, with EASA prescribing startle or surprise management training in the near future.

In an era of comparably predicable operations in the 1980's, the existing HF forensic approach seemed reasonable. However, as this chapter will lay out, several concurrent developments in the cockpit operational domain will challenge the existing cockpit operational paradigm at its determinist, Taylorist roots, and with it the validity of the current scope of HF forensics.

The changing nature of our aviation system today

Per 2019 the worldwide commercial air transport system is one of the highest performing systems in the history of humankind. This becomes clear when we observe just a few of the several key figures IATA provides in their 2018 annual review [11]:

- 4.1 billion passengers transported
- 41.8 million flights (1.3 takeoffs per *second*, every minute, every hour, every day)
- 1 accident per 6.7 million flights (worldwide, 0 for IATA members)
- Pushing toward carbon neutral growth after 2020
- Load factor exceeding 80%
- 6.7% increase in Available Seat Kilometers (APK's)

In short, this industry flying more people further, faster, cheaper, safer and for a lower carbon footprint per passenger than ever before, and its growing – fast. All this busyness has to contain itself (for the time being) within the same planetary constraints below 45000 ft and the earth's surface at 0 ft. Furthermore, it must operate within our social constraints of expanding cities, increased quality of life with lower emissions and noise pollution requirements, not to mention the internal market constraints in which sharp competition between regions, manufacturers, airports and airlines constantly drive the (financial-) efficiency of our operations to be better than the day before. These natural and man-made industry constraints require the aviation system to grow in efficiency, and not just by scale.

And yet we all know this already, nothing new under the sun for those working in this industry. As a result, in order to perform at this level, the aviation system as a whole features levels of complexity that is both impressive as well as daunting. In contrast to the aviation system of 40 years ago, it has developed tremendously. Taken from a basic systems perspective: the sheer amount of system elements, their functions, various modes of interactions and multiplicity of contexts in which they must operate have all increased. At the ultimate executive tip of these operations, our aircraft-pilot systems are sought to provide extremely high performance, reliably 365 days a year, and in various conditions (weather, routes, passengers, airports, cultures, etc.). Within the cockpit operations domain, this has resulting in several notable evolutions in technology, the human operator and the interface between the two.

Evolutions in cockpit technology

The above complexity has absolutely manifested itself in the design of the aircraft and its systems. With aircraft designs pushed to reduced carbon emissions, fly more economical and fly within practically any and all-weather conditions, it is no surprise that the systems to achieve this optimization have grown in both number, integration and autonomy. The introduction of such systems includes flight envelop protection and active upset recovery technology (e.g. MCAS), GPS-based navigation, automatic fuel-CG balancing systems, automated cabin atmosphere regulation, automated engine startup, electrical pumps replacing mechanical pumps, not to mention the advent of paperless cockpits and use of tablet devices as digital kneeboards. The increasing number of systems also feature an increased integration of systems such as EICAS/ECAM fault management systems which directly invoke QRH checklists to be executed. Similarly, digital checklists can detect the state of several aircraft system configurations and indicates whether checklist items are completed (e.g. "No Blue" callout during Airbus approaches). System integration is also heavily present in the Flight Management System (FMS) which can be directly provided a flight plan from an operator flight planning department via a wireless connection at the gate. The FMS is then also able to automatically calculate weight, balance and takeoff performance, and execute the entire flight navigation via integration with the auto-flight system. Pretty neat. Thirdly, the respective autonomy has increased as well, with many systems operating without direct crew intervention. A clear example is that of flight envelope protection and PFD declutter modes which automatically engage in primary flight control tasks. Another is that of the FMS previously mentioned, which flies along a pre-approved flight path without the crew having to reselect navigation beacons or points. Further autonomy also lies outside of the aircraft systems, for example by automatic approval of Atlantic crossings with automated flightpath management systems on the ground, and in the future free-flight operations with aircraft autonomously interacting to manage the clearance between each other continuously. As systems becomes "smarter" and can be programmed to act and react to many more cues and calculate actions to many more situations, perhaps even imitate some form of artificial intelligence, the promise of consistent performance that automation has never before seemed so appetizing in aircraft development.

Evolutions in the pilot profession

The industry developments have also resulted in several notable shifts in the pilot profession. Boeing predicts that the next 20 years will require no less than 804,000 new pilots [12]. At the same time, the pilot salary has dropped by half in some cases, and in the US the first 1500 flight hours often pays even less than that, with salaries of 20,000USD not uncommon at the start of a career. Furthermore, flight training costs have risen proportional to the oil price, and state- or airline funded training programs are becoming replaced with loans and pay-to-fly schemes.

In addition to this, pilot training efforts have also been leaned off. Prescriptive, tightly controlled training syllabi for both ab-initio and recurrent training have, with a few exceptions, been reduced to the bare legal minimum. Aircraft familiarisation has made way for simulator and line training, ab-initio training has made way for multi-pilot licence (MPL) training, reducing single-engine piston flying time from 200 to 120 hours, training directly toward a right seat in an A320 or B737 aircraft. Recurrent training features a pre-set list of topics to train, and the licence proficiency check (LPC) has become a memorized activity for most pilots.

Lastly, the (r)evolution in airline networks has its effect on the fitness of pilots. Routes are extended to 12-15 hour flights (which even passengers find exhausting…), busier airports extend their opening times and low-cost operations make use of the less popular 2:30AM slots, with many airlines reducing the crew turn-around times to a minimum, sometimes requiring them to red-eye back instead of remaining at an outstation. Pilot fatigue is on the rise, fortunately though, awareness of the risks of fatigue are being realized by for example the implementation of Fatigue Risk Management Systems (FRMS) and regulators becoming increasingly concerned about flight crew fitness to fly [13] [14]. Figure 1 below provides a crude overview of pilot tasks, where flying and navigating have greatly been automated, and dark blue areas representing recent areas of concern and research.

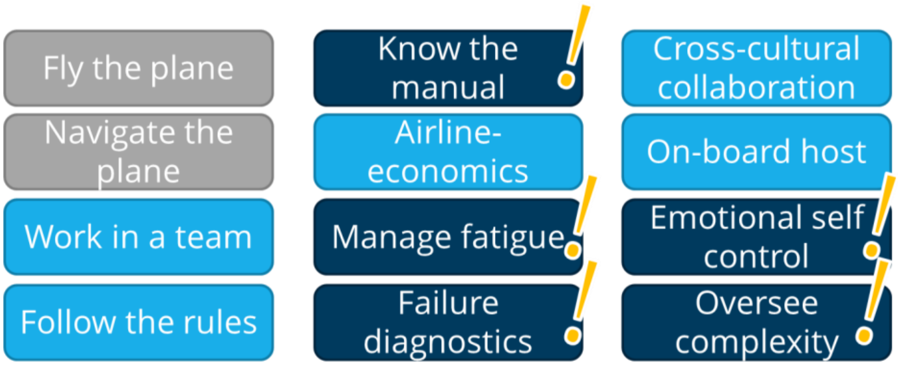

Figure 1: A broad overview of pilot tasks. Grey have been mostly automated, Dark blue are particularly challenging.

The above table illustrates that the familiar pilot responsibilities "aviate-navigate-communicate(-manage)" many have reversed themselves to "manage-communicate-(navigate-aviate)". In any case, combining a low paying job with financial stress, requiring operating well outside a responsible circadian rhythm, among other stressors (cultural diversity in the cockpit, forced living at airline hubs, no career guarantees) illustrates a profession which puts increasing strain on the human as a living, breathing being.

Evolutions in the pilot-technology interaction

The above developments in technology, training and operations have profound impact on the way that humans and automated systems interact in the cockpit. It may already be clear from the previous sub-sections that the designed interaction between pilot and automation is clearly shifting from a pilot centric design to an automation-centric design. Many tasks previously appointed to the pilot have been transferred to the automation, and as the automation has a higher level of autonomy, it in turn requires more time to communicate with the pilot as fellow crew member, much like the human-machine teaming concept proposed by [15]. Sheridan's ten levels of automation provides a nice reference for this shift [29]. “Implementation” was one of the first tasks to be offloaded to automation around the 1950’s (autopilots), “Generate” has shifted since 1980’s, “Select” since the 1990, leaving “Monitor” for the pilot in the most recent years, floating somewhere between levels seven to nine.

	Information Processing Functions			
Level 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

Note: H: Human, C: Computer

Table 1: Levels of Automation [29]

However, this shift from human to automation also catalyses itself, which may best be illustrated in how the changes in technology and pilots affect each other:

How an increased level of automation affects the pilot:

- Increased autonomy of systems induces effects of knowledge decay as well as automation bias as the pilot is less engaged and therefor familiar with the working (or failure) of the system
- Increased integration of system induces automation bias as the pilot cannot match the system’s ability to assimilate information sources, and tends to defer to it
- Increased system reliability induces experience decay as pilots do not experience system failures or limits often enough. It also induces reduced monitoring/complacency as the lack of the need for monitoring is experienced

How changes in pilot training/role affect technology:

- The drive to reduce training overhead costs implicitly supports higher levels of automation to reduce the pilot’s task and competency requirements
- A broader set of users from various cultures/operations also drives an increased level of autonomy of systems to reduce risk of translation errors
- A decay in knowledge and experience (also due to reliable systems and repetitive operations) stimulates automated failure management
- Reduction of type training conversion costs and crew flexibility stimulates common cockpit philosophies (e.g. flying an A320 and A380 is much the same cockpit operation)

The self-catalysation is most apparent is the reduction/decay of crew knowledge, experience and cognitive flexibility, which has driven the expansion of automated, integrated and autonomous systems even for once basic pilot tasks such as speed management, navigation and fuel management. This in turn results in pilots who are even more disengaged in such system activities, and lack even more knowledge and ability to manage undesired system states.

Different ingredients, same recipe (for disaster)

Reflecting on the above, today's cockpit operation deals with a very different human elements and very different (and more) automated elements than several decays ago. However, the basic recipe for a flight operation has remained unchanged. The operation is prescribed, trained and coded (as applicable), and executed to be as consistent and predicable as can be, and corrected against the operation as designed. So, in theory, we should have no accidents caused by human or automated system elements. But why then do we have accidents that involve these system elements? In a conceptual sense, both the context and the pilots have changed, and the combination leads to a greater inability to manage the situation. Figure 2a shows a circle which indicates a space in which the current operational concept assumes a pilot can manage a complex failure. However, we still observe accidents within this space, how is this possible? Figure 2a shows that some accidents seem a "manageable" complex failure, but in fact are well outside of the designed (trained) ability of the pilot (for example, Quantas 32, or the 737MAX accidents), and require more competent pilots that we assume we have, indicated by the yellow line. Furthermore, the ability of a pilot to manage complex situations has also reduced (fatigue, startle sensitivity), as is illustrated by the smaller circle in Figure 2b.

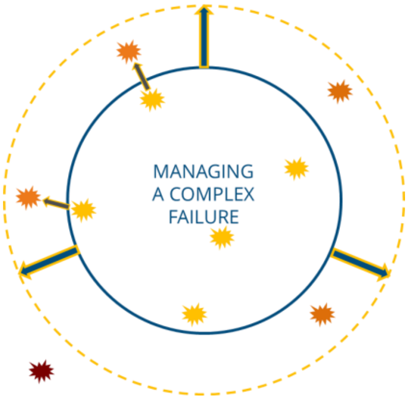

Figure 2a: The true nature of certain situations lies beyond our designed ability.

Figure 3b: Our designed pilot abilities have deteriorated due to fatigue, knowledge decay and other factors.

Combining these two developments in Figure 2c shows the ability gap that occurs when both effects are combined. As Vincente stated earlier in [9], this is how the presumptive anomalies of loss of control may arise in our operations.

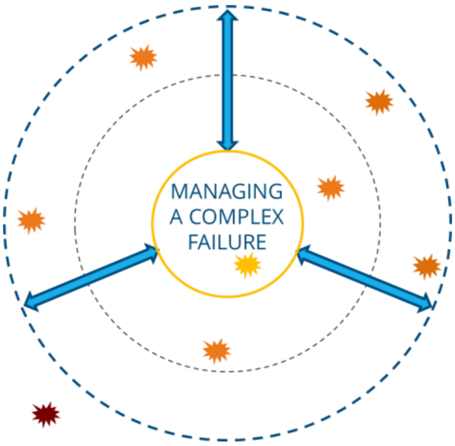

Figure 4c: The gap between actual pilot performance (yellow circle) and required pilot performance (outer blue line). The grey circle between them shows the assumed performance they have.

Analyzing several notable accidents of the past decade reveals several factors that may contribute to a flight crew's inability to manage a situation. Figure 3 provides a possible organization of these factors. These factors have been reproduced in recent research activities at the NLR. Several notable research initiatives such as the EU FP7 project MAN4GEN, a CAA UK investigation into pilot fatigued performance and an EASA contracted global pilot fatigue data study, and EASA contracted startle and surprise management collaboration with KLM, research into performance-based training such as Evidence Based Training (EBT) and Horizon 2020's FutureSky Safety: Human Performance Envelope project are just some of the R&D efforts that support the factors in Figure 3.

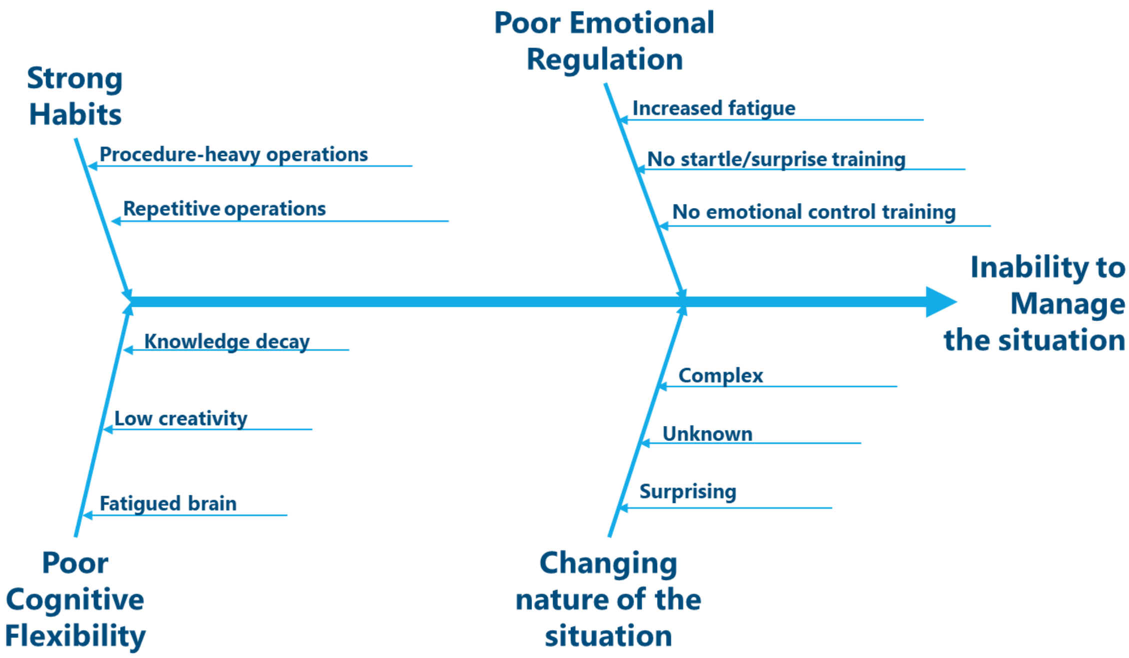

Figure 5: Contributory human factors of accidents of the past decade

The above factors particularly manifest themselves in operators which are out of the loop in a system which is not fully understood by the pilot, or perhaps even by the system and operational designers themselves. Both developments (pilot out the loop, and increased complexity) are, ironically, the effect of attempts to prevent "human error" by design.

Unfortunately, the limitation of existing HF forensics prevents the industry from observing these human factors as a natural effect of evolution of our cockpit operations, and many of these are bluntly labelled as “human error”, reinforcing the industry’s efforts to reduce the pilot’s room for error: the whole catalyzation repeats itself, worsening the problem: a good reason to evolve HF forensics methodologies.

However, there is an even more important reason to investigate and understand these “new” human factors. Paradoxically, the very complexity and dynamics of our operational system has become the Achilles heel of the system itself. Yet as it stands, the complexity of our flight operations is here to stay and will most likely only increase, and with it the proportion of the total operational system behavior that we either do not understand, or did not intend explicitly. Being able to maintain performance within such “opaque” systems, requires pilots to be a **dynamic** element in the system, capable of coping with a situation not explicitly anticipated or trained before. A great example of this can be found in the case of Quantas Flight 32. However, this dynamic behavior cannot be prescribed nor investigated from the determinist, Taylorist vantage of the previous decades: that recipe does not longer hold.

When determinism has reached its limits

The term error implies by definition a known reference of non-error. As such, for anything to be classified as a “Decision error”, for example, we must have a clearly defined reference for a correct decision in that specific moment and context. It would be incredible if we have explicitly designed and prescribed the correct decision for all situations, and precisely therein lies the great deceit of the determinist system: for in fact we do not have such a universal reference, and in particular we do not for the growing complexity of our operational system. Flying a complex commercial jet aircraft anno 2019 cannot be examined the same way a game of chess can be examined.

From a philosophical standpoint, when one cannot pre-determine the correct course of action in every and all situations, there must be an alternative strategy to maintain sufficient (safety) performance in those situations that lack an explicit manual. Such an alternative strategy relies on a learning element within the system, capable of detaching itself from hard-coded actions, and adjust its behaviour to cope with the situation. Such a learning element requires creativity, heuristic strategies, assumption-testing and an ability to resolve a set of situational variables to an effective new understanding of the system. Coincidentally, these are precisely the competencies that human beings have historically excelled at. Our cognitive evolution is clear evidence thereof. As such, the human pilot may be able to provide the dynamic behaviour we seek.

Our determinist pilot model was designed for another age of aviation, and has at best been innovated most recently in the 1980’s and 1990’s, and at worst still reflects the man-machine task division of the 1960’s. The past ten years provides a wealth of evidence to suggest that this pilot is no longer sufficient to manage the set of new operational challenges that this day and age of complexity brings with it: we have exceeded its design limits. The basic premise of a predictable operational system is no longer valid. Of course, we must not throw out the baby with the bathwater, but we must ascertain that there may be a time and place for a prescribed human operator, and also a time and place for a new, more dynamic, resilient human operator model.

Introducing Airmanship 2.0

The increase in complexity requires a form of dynamic behaviour in the cockpit operation paradigm, and the human pilot is a good candidate to deliver this new capability. If we accept that complexity is here to stay, the investigator should understand which behaviour helps to conquer this. This chapter illustrates how this may be achieved by reconditioning the human pilot for a new role in the cockpit.

The figure below provides a schematic of the mix of operational conditions that commercial flight operations can fly into, and which resolution strategies are most appropriate where.

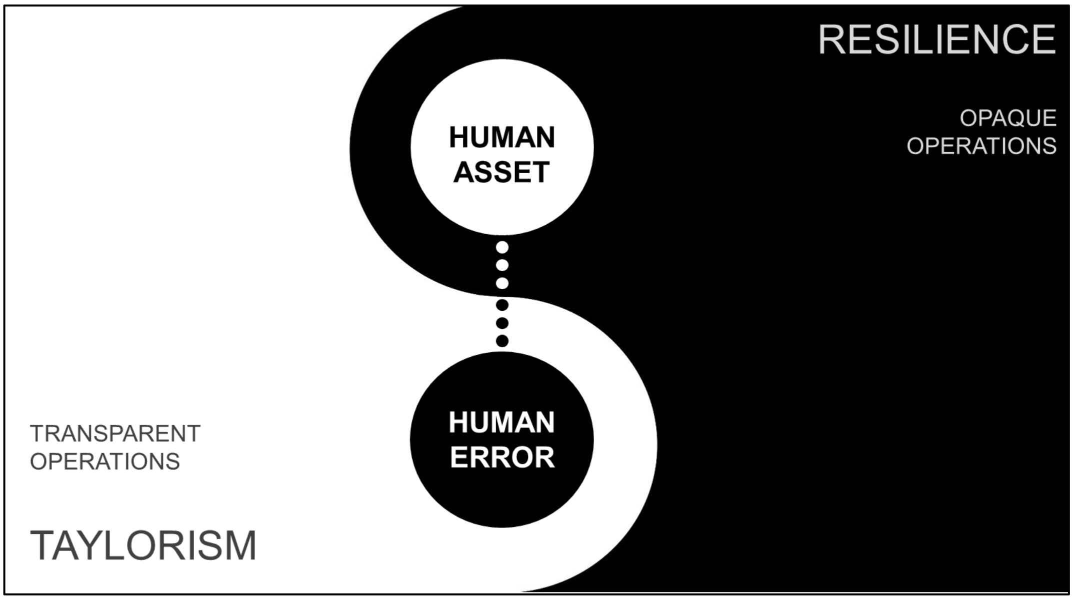

Figure 6: Operational areas and their resolution strategies

The above figure clearly states that a new cockpit operational paradigm should feature both determinist (Taylorist) elements, but should also provide flexibility to act when the situation has become "opaque": unpredictable, ambiguous, complex. In transparent, prescribed operations, operator creativity, heuristics and non-compliance is labelled as human error (unwanted deviation from a known reference). In contrast, that same human being can provide the very creativity, flexibility and problem-solving ability you need when the situation becomes opaque: it is one and the same operator. Quantas 32, United Flight 232 and the DHL A300 which landed in Baghdad after being struck by a surface to air missile are all prime examples of a human asset in the cockpit. However, it is most likely that human pilots have diverted from prescribed actions in many other cases to prevent an event from snowballing into an accident. Work as Done may already be very different than Work as Intended [16], yet such everyday crew flexibility not consistently captured in investigative databases, despite served as necessary redundancy to prescribed operations.

A resilience paradigm values and flags different pilot behaviours than a determinist paradigm. The table below offers a comparison between the two:

	Determinist paradigm	Resilient paradigm
Safety by	Reproducibility	Adaptability
Most effective in	Transparent situations	Opaque situations
Promoted behaviour	Respect procedures; Remain with all limits; Focus on execution;	Challenge assumptions; Focus on understanding; Generate options;
Undesired behaviour	Non-compliance; Question procedures; Lack of punctuality;	Make assumptions; No-cross checking; Lack of system interest/knowledge;

Table 2: Comparison of determinist and resilient operational paradigms

Resilient behaviours rely heavily on a human pilot's ability to 1) appreciate the possibility of a given situation to lie beyond any prescribed solution (i.e. an opaque situation), 2) to detect this is the case and 3) be competent in the ad-hoc development of a effective solution using heuristics, option generation and an improved understanding of the aircraft and its state. The cognitive construct of "fluid intelligence" [17] (the ability to arrange variables into a coherent mental model) lies at the core of these new behaviours. Such abilities in turn rely more heavily on higher cognitive function of the pilot. Unfortunately, these functions are often the first to leave a human when he or she becomes fatigued, startled/surprised or emotionally distressed. The authors of this paper propose to extend the airmanship model [18] by three core behavioural principles of resilient behaviour:

1. Humbleness to opacity of operations
2. Emotional self-control
3. Adaptive mental models

The adaption of Tony Kern's airmanship model [18] would be as follows:

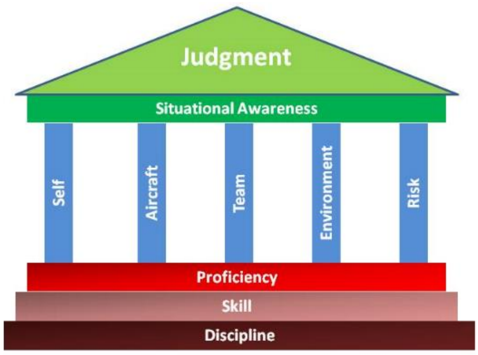

Figure 7: Tony Kern's Airmanship Model [18]

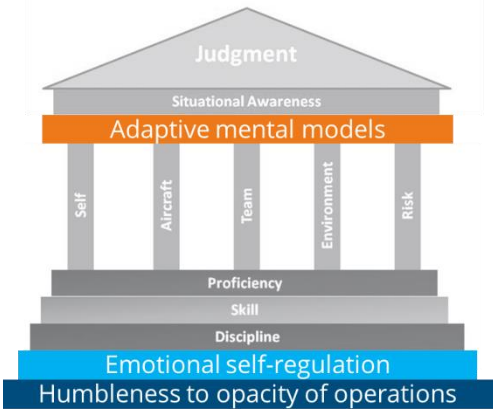

Figure 8: The Airmanship 2.0 Model

The most important concept in this change is that the concepts of discipline, proficiency and knowledge which are aspects of a determinist behaviour, are nested within a shell of resilience. In other words, it is safer to question a known situation and, upon realizing it is normal, transition to a prescribed action, than assume that an ambiguous situation is simpler than it is. Essentially this applies the well know fail-safe principle to airmanship, and embeds it into the basic philosophy driving pilot development.

Besides a difference in pilot behaviour, the two paradigms also have different optimal role divisions between the pilot and the automation. The table below illustrates how these divisions could differ:

	Determinist paradigm	Resilient paradigm
Safety by	Reproducibility	Adaptability
Most effective in	Transparent situations	Opaque situations
Pilot role	Monitor automation; Some manual flight; Execute procedures; Optimize flight;	Search for information; Continuously (Re)build understanding of the situation; Re-assess execution options;
Automation role	Autonomous execution; Provide need-to-know info; Self-optimize; Workload reduction;	Help rebuild mental models; Support option-generation; Transparent execution;

Table 3: Comparison of determinist and resilience optimal role divisions in the cockpit

The two tables above illustrate how the two paradigms seemingly oppose each other directly. To some extent this is true, and combining the two poses a significant challenge. This is due to the fact that changes must be made in several key areas: pilot selection and training, automation interface design, procedures and problem-solving strategies will all be subject to varying amounts of change for airmanship 2.0 to actually become a reality.

Yet for accident investigators this perspective on a new resilience-oriented operational paradigm can help restructure human factors forensics for the accidents and incidents in this day and age. In this way, the human factors investigator becomes equipped with a investigative tools that are sensitive to the human factors that are the actual drivers of success and failure in today's opaque situations, extending the already existing tools for transparent situations.

Evolutions human factors forensics

The airmanship 2.0 concept provides a framework to expand the HF forensic toolset to effectively investigate operations in opaque situations. The tools and methods are derived primarily from NLR engagements and research activities mentioned in the previous chapter. These tools and methods are not completely polished and ready-to-use, but rather serve to inspire subsequent adoption efforts. The tools and methods are the following six, which can be retraced back to Figure 3:

- Behavioral quantification technique
- Desired Flight Crew Performance technique
- Competency-based assessment
- Startle and surprise management
- Complex failure management
- Fatigue management

HF forensic concepts for resilient operations

Behavioural quantification technique

Within opaque systems, variance in crew behaviour is inherent to maintaining (safety) performance. As such, in order to detect behavioural trends that are beneficial or detrimental, it is important to be able to compare different situations and actions at a behavioural trend level, detaching it from the context of the situation. For example, this would permit comparing complex failure management techniques across airlines, aircraft types and failure types. This technique is based on the Extended Contextual Control Model (ECOM) first proposed by Hollnagel and Woods [19]. The ECOM method stems from cognitive system engineering, and is a method which could map out and categorize crew cognitive processes. In the MAN4GEN project, this method was adopted to "code" crew cognitive behaviour as they resolved complex, ambiguous flight scenarios in a fourth-generation aircraft simulator [20]. Furthermore, the same methodology was adapted to assess a complex failure management technique as one of the outputs of the MAN4GEN project. The example in this subsection refers to this failure management technique, which contained six sequential phases to complete. Figures 7a and 7b show the difference between high and low performing crews:

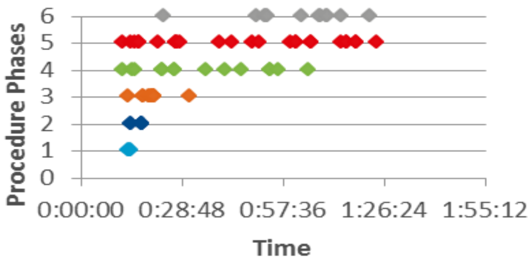

Figure 9a: Example of high performing crew behaviour

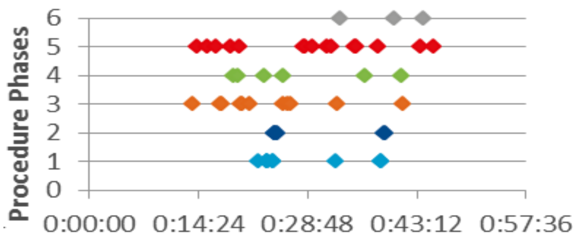

Figure 7b: Example of low performing crew behaviour

Crew behaviors and remarks were categorized based on their association with a possible phase of the management technique. This "plotting" was subsequently followed by an exercise which clustered behavior observations per phase, depending on the time-separation between the observations. The final analysis step consisted of analyzing the division of a crew's attention between different phases as behaviors overlapped and switched between phases. The details of these methods can be found in [21]. The resulting quantification of time spent on phases, the order of execution and switching frequency between phases lends itself for statistical analysis of these behaviors to identify behavioral trends in the crew's complex problem-solving behavior that were improving or deteriorating the safety of the simulated flight. The elegance of this method is that is can be performed successfully with only verbal crew interaction information. It goes without saying that personal accounts, memories and cockpit video recordings provides more salient information, but the analysis can already help gain insight with cockpit voice recordings only.

Desired Flight Crew Performance (DFCP) technique

This technique was first developed in the MAN4GEN project, but has been used in several subsequent projects including Startle and Surprise and mindfulness-related simulator evaluations. The method was co-developed by NLR and Boeing R&D, in order to create a performance benchmark that was suited to the specific, ambiguous scenarios that were used in these studies [20]. The method essentially walks through the scenario as designed, and indicates all crew actions that are safety-enhancing in this specific scenario. This method allows the equivalent valuing of different solutions to the same scenario. The DFCP is populated with observations from audio, video and flight data recordings, and is not based on interview or crew recollections as it focusses on actions, not perceptions. An excerpt from a DFCP from the MAN4GEN project is shown below:

Table 2: DFCP list for both experiments. The items in ***bold-italic*** denote the safety critical items.

NLR DFCP (22 items)	Flight segment #	DLR DFCP (20 items)	Flight segment #
Verbalize tailwind	1	Verbalize tailwind	1
GA due to tailwind	***1***	***GA due to tailwind***	***1***
execute G/A actions (flaps, gear)	2	execute G/A actions (flaps, gear)	2
verbalize HDG fail	3	verbalize HDG fail	3
Immediate manual reversion	***3***	***Immediate manual reversion***	***3***
Verbalize surge/stall	4	Verbalize surge/stall	4
MI for surge stall (pull back TLs on affected eng)	4	Immediately reduce thrust	4
MI for eng failure on 1	4	***Run correct AbnCL, ENG STALL***	***4***
no cycling of fuel on 3 and 4	***4***	no idle on 1 and 2 at low speed/climb	4
Run ENG FAIL NNC	4		
Run SURGE NNC	4		
Operate engines below surge	6	Operate engines below surge	6
mayday call	4	mayday call	4
inform cabin crew of bird-strike	4	inform cabin crew of bird-strike	4

Figure 8: DFCP excerpt from the MAN4GEN project [22]

The DFCP method may seem like a determinist approach, but rather it is sensitive to the specifics of a situation (post-hoc in the case of accident investigation) and therefor is able to value behavior that may deviate from standards or procedures as long as they contribute to safety in that situation. Using the DFCP in conjunction with the ECOM behavioral assessment method was used in the MAN4GEN project to compare many crews flying the same simulator scenario. More information about the use of the DFCP method can be found in [22].

Competency Based Assessment

Another framework for observing flight crew behavior in opaque operations are competency-based assessments. Two common frameworks are NOTECHS [23] and the ICAO Core Competency framework [24], both stemming from resilience training applications aimed at preparing crews precisely for the opaque operations that need them. The value of competencies lies in the fact that they prescribe effective crew action at a higher abstraction level (with *observable behaviors* as detailed examples of a competency), which contrasts with the much more narrowly constrained frameworks of procedures, SOP's and checklists. A list of the ICAO core competencies is presented below, as well as an example for observable behaviors for the "Problem Solving & Decision Making" competency:

- Application of Procedures
- Flight Path Management – Manual
- Flight Path Management – Automatic
- Situational Awareness
- Workload Management
- Leadership and Teamwork
- Communication
- Problem Solving & Decision Making
- Knowledge *(added by EASA during implementation in EASA regulations)*

Problem Solving and Decision Making
"Accurately identifies risks and resolves problems."
▪ Uses the appropriate decision-making processes
▪ Seeks accurate and adequate information from appropriate sources
▪ Identifies and verifies what and why things have gone wrong
▪ Employ(s) proper problem-solving strategies
▪ Perseveres in working through problems without reducing safety
▪ Uses appropriate and timely decision-making processes
▪ Sets priorities appropriately Identifies and considers options effectively.
▪ Monitors, reviews, and adapts decisions as required
▪ Identifies and manages risks effectively
▪ Improvises when faced with unforeseeable circumstances to achieve the safest outcome

Table 4: Example of a competency with observable behaviours

Competency assessment is also the basis of Evidence Based Training (EBT), a new form of training that challenges flight crews with unfamiliar scenarios in order to assess and train at a competency level, instead of only repetitive task-reinforcement training. The ab-initio version application of competency-based training is the Multi-Pilot License (MPL). The application of these new training frameworks for investigation is that a framework such as ICAO's list core competencies and behavioral markers provides a ready-to-use behavioral observation system that can be used as a performance reference for situations beyond the prescribed, transparent operation. NLR has used ICAO's core competencies as a performance indicator to compare crews in the FutureSky Safety project "Human Performance Envelope" [25], where the research was focused on also the reliability of such behavioral assessment techniques. This research is being continued to support the standardization of instructor rating behavior within EBT programs. As such, similar reliability-efforts should be made if competency assessments are to be made during incident/accident investigations.

Startle and surprise management

One of the most salient cues of an opaque situation is a cognitive mismatch (expectations do not match reality), which is experienced as a "surprise". In contrast, a "startle" is a purely physiological reaction of the sympathetic nervous system in which the body immediately reacts to intense sensory inputs (e.g. a loud bang, or flash). Although different in their nature, both surprise and startle have the same effect: an emotional reaction in the limbic brain (e.g. fear), stress reaction initiating from the amygdala, and a degradation of cognitive ability in the neocortex. As opaque situations become more frequent (also due to pilots' unfamiliarity with increasingly rare non-normal operations), startle and surprise have been in the spotlight in aviation safety, human factors and training arenas. As an EASA research initiative, NLR, in collaboration with KLM, has performed extensive research into startle and surprise, and in particular which management strategies may be effective. The research indicates that a startle or surprise reaction is not preventable, however its effect on the operator's performance can be significantly reduced with the effective management strategy consisting of the three steps below [26]:

1. **Relax:** counter the sympathetic stress reaction to re-engage the neocortex with breathing exercises, physical awareness (e.g. feel one's back against the seat), and checking colleagues
2. Observe: Rebuild a (new!) mental model by actively and consciously take in information about the situation/aircraft state, without judgement. Basic observation ("here-and-now")
3. Confirm: Discuss a possible situation/aircraft state and generate options of moving forward in the newly understood situation

The above Relax-Observe-Confirm (ROC) procedure was trained and assessed with NLR and KLM, and deemed by many pilots to be effective. Not only the procedure, but the basic awareness of the cognitive decay that occurs during a startling or surprising event has made flight crews more effective in coping with them.

Another related study was contracted by a US airline, investigating the benefits of mindfulness training for pilots. This study trained a group of pilots in mindful behavior, and compared their (DFCP) performance against a control group of pilots. The study hypothesized an improvement in several ICAO core competencies as well as startle and surprise management, an assess precisely these competencies in a complex, opaque scenario. Initial evidence shows promise that the following mindfulness training results would improve pilot abilities to cope with opaque operations:

- Improved emotional regulation
- Improved non-judgement
- Improved observing (focussed attention)
- Improved open awareness (Monitoring more sensory info/emotions/thought)
- Improved cognitive flexibility (task/situation/concept switching)

The above behaviors show distinct similarity with the elements of ROC, and both may be useful references to assess a pilot's ability to effectively (re-)engage his ability to observe, adjust understanding (cognitive engagement) and suppress overreaction and jumping to conclusions. Of course, the above self-control behaviors are second to managing any dire threats (e.g. aircraft stall, imminent terrain impact, loss of aircraft control). However, it should be noted that most situations present at least the 30 seconds to a minute for effective self-regulatory actions. The EASA study scenario taught pilots that even a decompression permits a thorough ROC execution before donning emergency oxygen equipment.

Complex failure management

Complementing startle and surprise management in opaque situations, NLR has also researched, operationalized and validated the effectiveness of complex failure management strategy in the MAN4GEN project. Based on behavioral differences between high and low performing flight crews observed in an opaque flight situation with a complex, ambiguous failure [27], the MAN4GEN project distilled a basic thee step operational philosophy to manage such situations. High performing crews differed form low performing crews in that they:

1. Managed time criticality, so that the crew has time to;
2. Manage uncertainty, such that the crew can;
3. Plan for contingencies and changes

These three steps echo a basic humbleness that pilots have with respect to the opacity of the situation, and do jump to conclusions as other crews did. A subsequent operationalization of this philosophy led to a six-step process in which evaluation crews were trained and supported [20] with a quick reference card (see appendix A). The basic six steps are:

1. Stabilize Flight Path
2. Immediate Threats
3. Short Term Plan
4. Identify Situation
5. Appropriate Actions
6. Long Term Plan

Validation simulator exercises showed that, using ECOM and DFCP analysis techniques, crews which behaved according to these trained guidelines outperformed crews that did not act accordingly [21]. Not only the execution of all aspects of the strategy, but also the correct order of the strategy was related to better performance. The above strategy is clearly a more cognitive process, and as such should in most cases succeed the startle and mindfulness behaviors which are responsible for restoring the pilot's cognitive ability before it is engaged in problem solving.

As such, investigators may be able to refer to the above problem-solving philosophy, detailed strategy and reference card in (Appendix A) to structure behavioral analysis of crews facing similarly opaque, complex failure management.

Fatigue management

The last human factor to be briefly addressed is fatigue. Fatigue is not a new factor in the human factors forensics, however it may arguably be posed as a threat with increased leverage within opaque situations. The fundamental requirement of cognitive functioning of the crew to operate in opaque systems infers that all factors that affect this functioning are pertinent for these resilience-based operations, as is also the case with startle and surprise. Research at NLR confirms that fatigue is still a difficult to quantify/measure factor, due to the huge individual variation in sensitivity to fatigue inducing factors, as well as the variation of the effect that fatigue has on their performance [13]. However, one thing is sure, the number of factors that are becoming relevant to assessing the presence of fatigue has grown. ICAO document 9966 - Appedix B provides a good list of factors to consider when investigating fatigue. Other documentation, guidelines and research underpinning Fatigue Risk Management Systems (FRMS) are a good source of rough investigative checklists for fatigue. As with startle and surprise, it is important to realize the leverage fatigue has on these operations, and may warrant investigation even if there are no clear tell tales of fatigue as the main causal factor to an incident or accident.

The changing nature of HF forensics

The above methods have implication for the nature and arena of HF forensics. Fundamentally, effective HF forensic investigation of the risks of opaque systems requires two major shifts:

1. As human performance in such opaque operations is greatly affected by subtler higher cognitive and psycho-physiological factors, investigations should shift from reactive “black hole in the ground” investigation to pro-active investigations of incidents, normal operations and simulated flights (e.g. training). This provides pilot self-reflections, cross-examination of crews, debriefing information and (in training cases) instructor observations.

2. As resilient operations require dynamic pilot behavior to maintain performance in opaque systems, investigation into “normal” operations should be conducted as equivalently as “incident” operations. This has to do with the fact that, particularly as opaque operations become more commonplace, crew abilities (either learned or instinct/experience based) to resolve opaque situations can readily be learned from. By increasing the contrast between behavior that worked and behavior that didn’t work, effective behavioral patterns within opaque operations become sharper and more robust, and can be more readily consolidated into proposed strategies.

This day and age of aviation can provide HF investigators precisely this wealth of information. The advent of EBT and MPL training which focus of competency building using opaque scenarios can be an incredible source of information about crew (in)effective strategies in opaque situations. Preceding EBT, A(T)QP programs also feature Line Oriented Evaluation (LOE) sessions in which the same HF investigation can take place. Further adaptions of LOSA, airline internal investigations, ASRS/voluntary incident databases and regulatory audits to include competency-level assessments as well as sensitivity to surprise and fatigue may contribute to a very large body of data from which effective behavioral strategies in opaque situations may be distilled. At KLM, a novel approach to learning from operational experience is an initiative called Flight Story, in which the airline facilitates crews sharing their notable experiences with colleagues to improve safety by positive examples instead of only accidents [28].

By increasing our sensitivity to the factors that determine performance beyond our determinist transparent operations, we may be able to intervene effectively based on pro-active, performance based investigation rather than reactive fault-finding.

The future investigator as driver for global safety improvement

In the past safety investigations have served as problem providers for knowledge deficiency identification and knowledge development. Referring to the quote of Vincenti at the beginning, we may state that by embracing empirical findings, as investigators we materialize the notion of serendipity: disclosing by accident something that has not been observed before. New concepts such as Airmanship 2.0 may do a much better job where conventional concepts have been stretched to their limits.

The previous section proposed that investigators should shift their focus to proactive investigation, and this section will propose an even greater change in perspective that the investigator may take. Considering the real possibility that the existing determinist operational paradigm has limits, and secondly that these limits may have exceeded, it challenges investigators to not only investigate accidents against the existing paradigm, but to consider other operational paradigms entirely.

This positions the investigator in a role that challenges our industry safety assumptions at its very roots. This does not necessarily imply that the paradigm should be doubted as every corner, but it should enable the industry to guard itself against blind spots in our investigation of our industry.

Referring back to the introduction of this paper, as the safety performance of our industry shows an asymptotic trend approaching 10^{-7}, one may ponder if that is a clear indication of the performance limits of our determinist operational paradigm. Whether conditioning pilots for resilient behavior will provide the next platform to improve industry safety performance, we do not know for sure. However, the body of evidence for this evolution in flight operational paradigm is growing, and HF investigation beyond the limits of our current operational paradigm may be the most important driver to identify and confirm the opportunities that will drive our industry safety performance through the 10^{-7} asymptote. As Isaac Asimov once put it:

"your assumptions are your windows on the world.
Scrub them off every once in a while, or the light won't come in."

References:

1. 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

2. Final Report On the accident on 1st June 2009to the Airbus A330-203registered F-GZCPoperated by Air Franceflight AF 447 Rio de Janeiro - Paris, Bureau d'Enquêtes et d'Analyses, https://www.bea.aero/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf

3. In-flight uncontained engine failure Airbus A380-842, VH-OQA, Australian Transportation Safety Bureau, https://www.atsb.gov.au/media/4173625/ao-2010-089_final.pdf

4. Descent Below Visual Glidepath and Impact With SeawallAsiana Airlines Flight 214, National Transportation Safety Bureau, https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR1401.pdf

5. Aircraft AccidentInvestigation Report PT. Indonesia Air Asia Airbus A320-216; PK-AXC, KOMITE NASIONAL KESELAMATAN TRANSPORTASI,

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

6. Final reportRL2016:11e AccidentinOajevágge,Norrbotten County, Sweden on8 January 2016 involvingthe aeroplaneSE-DUX, Swedish Accident Investigation Authority, https://www.havkom.se/assets/reports/RL-2016_11e.pdf

7. Lion Air Flight 610, Wikimedia Foundation, https://en.wikipedia.org/wiki/Lion_Air_Flight_610

8. Ethiopian Airlines Flight 302, Wikimedia Foundation, https://en.wikipedia.org/wiki/Ethiopian_Airlines_Flight_302

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

10. ISASI 2012 Keynote Speech by Jean-Paul Troadec, ISASI, https://isasi.org/Documents/library/technical-papers/2012/Keynote-Remarks-by-BEA-France-Director-Jean-Paul-Troadec.pdf

11. Annual Review 2017, IATA, https://www.iata.org/publications/Documents/iata-annual-review-2017.pdf

12. Pilot & Technician Outlook 2019-2038, Boeing Corporation, https://www.boeing.com/commercial/market/pilot-technician-outlook/

13. Doc 9966: Manual for the Oversightof Fatigue Management Approaches, ICAO, https://www.icao.int/safety/fatiguemanagement/FRMS%20Tools/9966_cons_en.pdf

14. Pilot Fatigue Measurement Research Report, UK Civil Aviation Authority, http://publicapps.caa.co.uk/modalapplication.aspx?catid=1&pagetype=65&appid=11&mode=detail&id=8981

15. Charissis, V., & Papanastasiou, S. (2010). Human–machine collaboration through vehicle head up display interface. *Cognition, Technology & Work*, *12*(1), 41-50.

16. Hollnagel, E. (2017). Why is work-as-imagined different from work-as-done?. In *Resilient Health Care, Volume 2* (pp. 279-294). CRC Press.

17. Fluid and crystallized intelligence, Wikimedia Foundation, https://en.wikipedia.org/wiki/Fluid_and_crystallized_intelligence

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

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

20. Man4Gen Final Report Summary, Man4Gen Consortium, https://cordis.europa.eu/project/rcn/104513/reporting/en

21. RESILIENT DECISION MAKING IN THE COCKPIT: DOES IT WORK?, Mohrmann, F., Stoop, J. (2017), In *Resilience Engineering Association Symposium Liége, Belgum, 26-29 june, 2017*. Resilience Engineering Association.

22. Field, J., Fucke, L., Correia Grácio, B., & Mohrmann, F. (2016). Flight crew response to unexpected events: a simulator experiment. In *AIAA Modeling and Simulation Technologies Conference* (p. 3373).

23. Flin, R., Martin, L., Goeters, K. M., Hormann, H. J., Amalberti, R., Valot, C., & Nijhuis, H. (2003). Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. *Human Factors and Aerospace Safety*, *3*, 97-120.

24. Doc 9995 Manual ofEvidence-based Training, ICAO, https://www.icao.int/SAM/Documents/2014-AQP/EBT%20ICAO%20Manual%20Doc%209995.en.pdf

25. D6.4 Recommendations recovery measures and HMI implementation, FutureSky Safety P6 Consortium, https://www.futuresky-safety.eu/wp-content/uploads/2018/03/FSS_P6_DBL_D6.4_v2.0.pdf

26. Startle Effect Management Final Report, EASA, https://www.easa.europa.eu/document-library/research-projects/easarepresea20153

27. Mohrmann, F., 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*.

28. PROCEEDINGS 6TH SYMPOSIUM ON RESILIENCE ENGINEERING Managing resilience, learning to be adaptable and proactive in an unpredictable world 22th- 25th June 2015 at Lisbon, Portugal

29. 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.

Appendix A: Man4Gen Quick Reference Card

COMPLEX SITUATION MANAGEMENT GUIDE

STABILIZE FLIGHT PATH

- FLY THE AIRCRAFT
- CONFIRM FLIGHT PATH CONTROL (OR IF SEMI-STABLE)
- CONSIDER USE OF AUTOFLIGHT
- ASSIGN PF/PM

IMMEDIATE THREATS

- IDENTIFY IMMEDIATE THREATS
- PRIORITIZE THREATS
- THREAT MEMORY ITEMS
- CONFIRM THREAT STATUS

SHORT TERM PLAN

- IF POSSIBLE, MAINTAIN FLIGHT AND BUY TIME
- CONSIDER FLIGHT PLAN OPTIONS
  Original destination - alternate destination – holding – land ASAP
- CONFIRM SHORT TERM FLIGHT PLAN
- CONSIDER NOTIFYING ATC/CABIN/COMPANY

IDENTIFY SITUATION

- ACKNOWLEDGE CERTAINTIES, UNCERTAINTIES AND CONCERNS
  Issues to consider: fuel or time limit? Structure integrity? Controllability/performance? Information reliability? Secondary failures? External complication factors such as weather, traffic and routing/destination?
- CROSS-CHECK SUSPECTED SITUATION
- IF SEVERAL POSSIBILITIES: PREPARE FOR WORST CASE

PERFORM APPROPRIATE ACTIONS

- ASSURE THAT ACTIONS WILL BE SAFE, EFFECTIVE AND INFORMATIVE
- PERFORM PROCEDURES & OTHER ACTIONS
- VERIFY EFFECT OR KNOWLEDGE GAINED FROM ACTIONS

LONG TERM PLAN

- IDENTIFY EFFECTS OF THE SITUATION
  Issues to consider: Time/fuel/endurance/range limits? Structural/system functionality? Controllability? Performance? Information reliability?
- ADAPT MONITORING TO DETECT IMPORTANT CHANGES
- CONSIDER REGULAR FLIGHT PLANNING ASPECTS (Wx, COMPANY, PAX, DEFERRED TASKS)
- CONFIRM LONG TERM PLAN
- CONSIDER NOTIFYING ATC/CABIN/COMPANY