Human Interaction and Emerging Technologies (IHIET 2023), Vol. 70, 2023, 77–87
https://doi.org/10.54941/ahfe1002932

The Future Impact of Digital Assistants on Aviation Safety Culture

Barry Kirwan

EUROCONTROL Innovation Hub, Brétigny-sur-Orge, 91222, France

ABSTRACT

In the coming decade, Artificial Intelligence-based **Digital Assistants** are likely to appear in operational aviation contexts, including the cockpit and air traffic control Ops room. Current scenarios for such AI support include advising flight crew during mid-flight emergencies, and executing routine air traffic duties in the Tower to reduce complexity and controller workload. The concept for Digital Assistants goes beyond today's Machine Learning-based tools, which largely offer information to human operators. Instead, the notion is of an AI-based 'colleague' that can engage in dialogue with its human counterparts. This in turn leads to the notion of a Human-AI Team and raises a host of questions about how such a team can and should function to optimise system performance and safety. One question in particular concerns how working with a Digital Assistant, and even potentially relying on one in safety critical scenarios, will affect the team's, and the parent organisation's **safety culture**, since safety culture is seen as high in the industry, and valuable in assuring passenger and crew safety. In the European air traffic network, safety culture is measured regularly in different countries using a standardised 50-item scientifically validated questionnaire. This questionnaire has been applied to the Digital Assistant concept to see which facets of safety culture might be affected. The results of this analysis have identified six high-level concerns, but also six instances where the Digital Assistant could potentially reinforce or improve safety culture, providing new '**safety affordances**'. Although the current work's focus is on aviation, the safety culture issues raised here may also pertain to other domains including health care, the energy sector, space and defence systems.

Keywords: Safety culture, AI, Digital assistant, Aviation, Safety affordances

INTRODUCTION

In aviation today, safety in terms of accident rates is seen as better than in the other three transport modes (rail, sea and road). This is in part attributed to the level of safety culture in aviation, hard won due to various accidents during the early years (1950's onwards) of commercial aviation, and following the introduction of 'glass cockpits' in the 1980's (Billings, 1996). Since commercial aviation accidents in flight are often fatal, there is both strong public awareness of air crashes (they usually make the headlines) and appropriately demanding regulation on safety across the aviation industry spectrum. Hence, safety culture has generally been strong in aviation, whether in aircraft manufacturing, air traffic control, airport operations or in the airlines themselves. Given the advent of Artificial Intelligence, already in use

in some quarters in commercial aviation, the question this paper addresses is whether future, more advanced AIs known as Digital Assistants, might help or hinder safety culture in aviation. The paper does this by first describing how safety culture has evolved in aviation and how it is assessed. It then considers examples of Digital Assistants in future (2030 and beyond) aviation scenarios, and evaluates how they could impact on safety culture dimensions and operational practices, both positively and negatively.

Safety Culture

A useful definition of safety culture is the following (ACSNI, 1993):

> "*The safety culture of an organization is the product of individual and group values, attitudes perceptions, competencies, and patterns of behaviour that determine the commitment to and the status and proficiency of an organization's health and safety management*"

Safety culture has been shown to be a key predictor of safety performance in a number of industries (including nuclear, chemical, oil & gas, and rail). It refers to the attitudes of personnel to the company's approach to safety, their perceptions about the magnitude of the risks they face, and their beliefs about how to control those risks. It affects what they say and do about safety during their daily work. Safety culture concerns everyone: the most obvious candidates are the front-line people such as air traffic controllers and pilots. But everyone contributes to the state of safety in an organisation, including managers, people in support functions, cabin crew, flight ops, drivers on the apron, fuellers, ramp operators and baggage handlers, engineers, caterers and cleaners, security and police, etc. At an airport, for example, almost all operations are inter-connected, involving multiple partners from different organisations, so it is even more important that all these companies and their staff have a positive safety mindset. The bottom line is that *people make safety*, so no one should be excluded.

Safety Culture in Aviation

The origins of Safety Culture are usually traced back to the Chernobyl nuclear power plant accident in 1986. Just as the Three Mile Island nuclear power plant accident in 1979 demonstrated unequivocally the importance of Human Factors in the design of human-machine interfaces in high systems, Chernobyl showed that the prevailing operational culture could catastrophically trump established safety procedures and processes. It appeared that if, as Management legend Peter Drucker famously said, "*culture eats strategy for breakfast*," then unfortunately it could eat safety for lunch, too. A spate of high-profile accidents from different domains (see Figure 1), including space, oil and gas, and rail – all seen as 'safety culture accidents' – only served to emphasise the enduring importance of this newly identified organisational trait. Certain high profile public enquiries into key accidents such as the Piper Alpha disaster (Cullen, 1990) and Clapham Junction rail crash (Hidden 1989), as well as key safety thought leaders at the time (e.g. Turner and Pidgeon, 1997; Reason, 1997), and a number of accidents at least

Figure 1: Safety Culture Timeline for European Aviation.

partly attributed to safety culture ever since, have ensured that safety culture has endured as a critical and non-negotiable attribute for any high risk organisation.

Figure 1 highlights the fact that safety culture wasn't initially seen as being of too much concern for aviation. This was despite the Kegworth air crash in 1989 (AAIB, 1990), which had certain safety culture aspects. The thinking at the time was that the strong training and (Human Factors-based) design in the cockpit and air traffic operations rooms, as well as Safety Management Systems (SMS) and Standard Operating Procedures (SOPs), were sufficient. In Europe, this notion was shattered with the mid-air collision over Lake Constance in Überlingen in 2002, following shortly after the Milan Linate runway collision a year earlier. As Chernobyl did for nuclear, these accidents triggered a rethinking that SMS and SOPs were not enough, and that safety culture was crucial. SMS was seen as the safety *processes*, whereas safety culture was the motivation to energise such processes for safe outcomes. Since the two accidents were primarily related to air traffic management (ATM), the safety culture assurance method development was carried out in that sector of the industry, with almost all European Air Navigation Service Providers (ANSPs) engaging in one of several independent safety culture evaluations of their organisation. This led later to safety culture spreading to some airlines and airports, though not nearly as systematically as the ongoing ANSP programme on safety culture.

Management and Designers – 'Upstream' Drivers of Safety Culture

Just as aviation safety occurrence investigation has generally shifted its focus from 'human error' to 'systemic' failure in recent years, the concept of safety culture must also consider the critical importance of **management and designers regarding safety**, based on their collective values, beliefs and behaviour, as well as the safe behaviour and attitudes of operational personnel, engineering and other support staff. As the wide-ranging investigation into the two Boeing 737 Max accidents has shown (Zwiefel and Vyal, 2021; Dias et al., 2020), even with the best engineering and a strong track record in safety performance, a compromised safety culture can lead to disaster. And whilst

most 'Human-AI Teaming' research is currently focused on human-AI interactions at the 'sharp end', i.e. operational contexts, the safety culture of the designers and developers of AI systems – the manufacturers – remains an important consideration.

Measuring Safety Culture

Since the Uberlingen mid-air collision in 2002 (Nunes and Laursen, 2004), there has been a focus on safety culture in the air traffic industry. EUROCONTROL, originally in partnership with the University of Aberdeen and more recently the London School of Economics (LSE), developed a validated survey approach (Reader et al., 2015) that has been applied to over thirty countries in Europe, with a number of Air Navigation Service Providers (ANSPs) carrying out surveys every 3–5 years. These ANSPs have found the process useful in understanding their complete risk picture, where their safety culture strengths and weaknesses lie, and how to improve them (Kirwan et al, 2021). Since 2016, this approach has been extended to deal with other sectors of aviation including airlines and airports (e.g. Kirwan et al, 2019). The EUROCONTROL Safety Culture questionnaire is based around eight safety culture 'dimensions':

- Management Commitment to Safety
- Collaboration & Involvement
- Just Culture & Reporting
- Communication & Learning
- Colleague Commitment to Safety
- Risk Handling
- Staff and Equipment
- Procedures & Training

Each of these dimensions is supported via a set of carefully worded questions, tailored to different segments of the aviation industry (ANSP, airline, airport, and airframe manufacturer), e.g. '*My colleagues are committed to safety*,' and '*If I see unsafe behaviour by one of my colleagues, I would talk to them about it*.' The answers to these questions build a picture of the safety culture of an organisation, including an understanding of how it may differ in an organisation's various sub-cultures.

AI in Aviation – Just Another Tool, or a Potential Step Change?

Safety culture is a property of the *people* working in an industrial system, and while this is largely true (without people there is no safety culture, or indeed culture), in practice it is an emergent property of the *system*, including the technology, and Artificial Intelligence (AI) will certainly be a feature of the future technology used by aviation personnel. Currently, it is common in 'culture' discussions to talk of *artefacts* (i.e. tools) that people use, with current aviation examples being flight strips or track data blocks (air traffic), various cockpit displays, alarms and electronic flight bags (airlines), weekly safety reports and safety messaging via airport community 'Apps' (airports), and drone control interfaces (drones and urban mobility), etc.

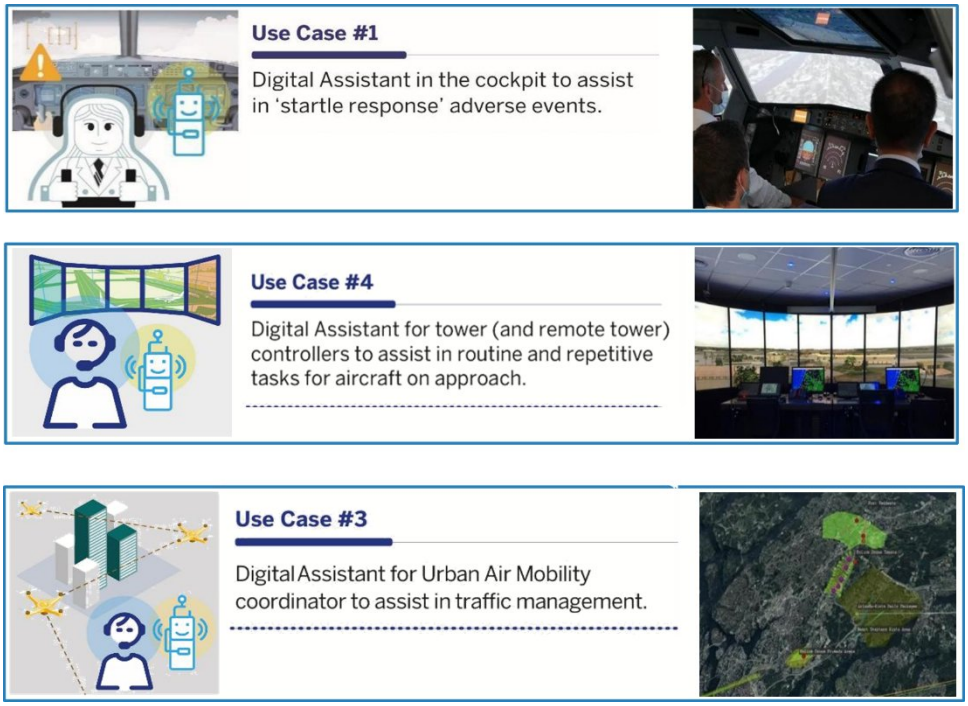

Figure 2: Three examples of use cases in the HAIKU project.

Current AI in the form of Machine Learning (ML) can provide information that can be considered as an artefact, or feeding into an artefact, and in most practical respects can be simply considered as additional automation (Kalliardos, 2022). But future (as yet to be realised) Digital Assistants (DAs), anticipated to be realised in the 2030–2040 timeframe, aim to go further. Current AI support doesn't 'answer back', whereas there will be dialogue with future DAs, elevating the role of AI from information support to team player and in some cases, even team supervisor. This means the future Digital Assistant must not only make sense of the task and its context, but it must also to a degree make sense of human culture, so it knows how to interact with humans so as to be effective. But what might such Digital Assistants be asked to do in aviation, and what roles might they take?

Exploring Future Digital Assistants – the HAIKU Project

HAIKU – *Human AI teaming Knowledge and Understanding for aviation safety* – is a three-year (2022–2025) European-funded project aimed at enhancing Human-AI teaming for future aviation systems in the 2030+ timeframe (https://www.haikuproject.eu/). HAIKU has six human-centric AI use cases, two each in the air traffic, cockpit and airport sectors, where prototype Digital Assistants will be developed (see examples in Figure 2). HAIKU aims to explore human-AI interactions and teaming in dynamic and realistic simulations of operational flight scenarios.

Assessing the Impact of the Digital Assistant on Safety Culture

Each of the 50 questions from the EUROCONTROL questionnaire was considered in the context of a future Digital Assistant, e.g. *commitment to safety* might reduce if the Digital Assistant was judged to be handling safety, and this could affect both human operators at the 'sharp end' as well as managers running an organisation. Initially eighteen potential concerns were identified, in particular with respect to *Colleague Commitment to Safety* and *Just Culture*, as well as *Management Commitment to Safety*, *Risk Handling* and *Procedures*. However, a similar number of potential safety improvement areas were also identified, wherein a Digital Assistant could potentially improve safety and safety culture. Each was further analysed and in several cases amalgamated into higher level constructs, leading to six safety culture concerns or threats (that could reduce safety culture), and six 'safety affordances' (which could enhance existing safety culture). These are presented below.

Six Safety Culture Concerns

People Make Safety, Don't They?

In aviation today, it is often said that 'people make safety', whether in the cockpit, the air traffic Ops room or tower, or on the ground at the airport. But if Digital Assistants become (in)valuable safety assets, might this change? Currently in aviation safety culture (and other domains), the notion that safety is everyone's responsibility has replaced the decades-ago refrain that 'safety is only the job of the safety manager'. Could the reliance on human commitment to safety be replaced by strong safe AI performance characteristics, and is this a trade we want to make (because there may be no coming back from it)?

Who's Flying the Plane?

At what point will passengers and businesses be happy to have an AI fly a plane or sky-taxi? The comparison with self-driving cars or automated train systems is not ideal, due to the vastly larger number of people at risk in air crashes, and the fact that there is no safe 'default' mode in an aircraft once in flight (it cannot simply stop or apply the brakes). One major 'AI-induced' air crash could delay AI-controlled flight for years. Probably AI's controlling role in other sectors outside aviation will weigh on such a decision (e.g. in self-driving cars, trucks, and trains). Aviation is also currently taking a cautious 'stepwise' approach towards full automation, e.g. AI plus single pilot during the cruise phase of flight (while one pilot is resting); single pilot plus AI for cargo operations (no additional pilot); AI plus single pilot for short-haul commercial flights; AI-piloted sky-taxis; etc.

Who's to Blame?

It is unlikely that any AI manufacturer will guarantee that its Digital Assistants will never make a mistake. It would probably be unreasonable to do so, as any AI relies on data which may be incomplete or skewed, and cannot anticipate every eventuality that could come to pass. Nevertheless, humans working with Digital Assistants may fear a double-bind: if they fail to take

safety critical advice from an AI that turns out to be correct and there is an accident, or if they take such advice and it turns out to be flawed (and there is an accident), how will they be judged? Judicial bodies have a mission that 'justice must be served,' and relatives of victims inevitably want someone to blame, someone to sue, someone to pay. This dilemma is likely to apply to many domains, and will be at the forefront of professional unions' concerns about the introduction of AIs into human teams with safety-critical roles.

Digital Cop

The AI should be able to advise, but also point out if an unsafe course of action is being taken by its human 'colleagues'. Depending how this is done, and how frequently it happens, resentment could occur between the human and Digital Assistant, as the human may feel (s)he is being 'policed'. A related aspect is that the DA could be used to check on the human's competence. This could be advantageous if it is done confidentially to the individual human user, but disadvantageous if it reports to management on the human's performance, in which case it would likely be seen as a 'snitch.'

Turn Me Up, Turn Me Down

Digital Assistants may not be solely concerned with safety; there may be other 'drivers' impacting their decision-making, such as environmental considerations, (e.g. carbon footprint, noise near airports, etc.), as well as performance considerations related to efficiency and competitiveness. Today, people in aviation already face such competing priorities, and make the necessary trade-offs. However, when safety comes to the fore, they move such factors to one side and 'put safety first', in order to minimise the risk of an accident or its consequences. Humans make this judgement call based on experience and their instincts and values. It is not clear whether an AI could do the same (we do not really know how humans do it). The problem is that the human receiving AI advice may not be able to unpack the optimising balancing act that underpins the AI's advice. It also means that the AI's own weighting of competing demands could be set by its 'owners', and with respect to the priority of safety, turned up or down.

Fewer People, More AI

If AI is automation by another name, then it must be accepted that automation inevitably leads to fewer human workers, e.g. single pilot operations after 2030, and perhaps fewer air traffic controllers as well as less ground handling staff as more automation and robotics supports ground activities. In the post-COVID world wherein several aviation sectors are finding it hard to attract staff (e.g. pilots or airport ground staff), this may seem less of a problem than before. And yet aviation has always attracted people who are passionate about their jobs, and this has no doubt supported a strong safety culture. If the job becomes less challenging, with fewer people, how will this affect motivation and commitment to safe work?

Six Safety Culture Affordances

Don't Panic…

One of the prime intended uses of AI Digital Assistants is to help pilots in what are known as flight upset conditions, in which something goes drastically wrong, and the flight crew have very little time to diagnose the event and recover the aircraft. Accidents such as AF 447 and the two B737 Max accidents (and others) fall into this category. Whilst it can take precious minutes for flight crew to work out what is happening and how to correct it, a Digital Assistant plugged in to all the sensors, with a vast database of potential conditions and recovery actions, could arguably make such a diagnosis in seconds. This use of AI (Digital Assistant in the cockpit) is therefore a priority in aviation safety research.

The Living Black Box

Commercial aircraft carry two black box recorders, recording all control inputs and voice recordings of flight crew prior to an accident. Nevertheless, investigators are not always clear from this information as to why pilots chose a particular course of action over another, and accident investigations can take months or even years piecing together what really happened. A Digital Assistant, particularly one that is interacting with a human crew, could in theory record in far richer detail the reasoning and decision-making during an event (whether recovered or not) than is possible today.

The Dispassionate Oracle

Today there are various aviation risk models and 'top 5' or 'top 10' hitlists for safety area improvement. These are always a mixture of operational data (from accidents and incidents or 'near misses') and subjective judgements from experienced safety managers or directors. Such data are always 'lagging', since they relate to the past few months or even years of operational experience, and depend on detailed and painstaking analysis that can itself take months or years. Such safety management approaches can be relatively 'blunt' as they may miss weak signals, or common underlying factors that could make a larger safety difference if they were improved. An AI that sifted through all the data, including live operational data, and 'knew' all the risk models and could do the 'safety calcs' (including use of Bayesian statistics to avoid certain human judgemental biases), could in theory determine a better way forward for safety enhancement and accident prevention. It could help aviation organisations see around the corner and take quicker action 'upstream' before incident patterns turned into accidents.

Mr Know It All

A Digital Assistant can be a vast source of knowledge for aviation flight crew, controllers, and ground staff, on procedures and hazards, and if networked into live operations and multiple teams, can give live, up to date safety advice related e.g. to weather, problems other teams might be encountering etc. Such a Digital Assistant could also store daily NOTAMs (Notices to Airmen) and other briefings, and transmit them to the teams, and ensure that they don't

miss any key points during subsequent operations. Such a DA could be seen as a valuable repository by human teams.

Looking After the Little Guy

Teamwork can have adverse effects, such that someone's advice is over-ruled, even though later it might turn out that such advice should have been listened to. A number of aviation accidents which manifested this problem led to the development of cockpit Crew Resource Management (CRM) training (and its Team Resource Management equivalent in air traffic). The Digital Assistant could in effect be integrated into CRM practices, ensuring that all voices are heard and listened to, before a decision occurs (depending on time constraints) in the scenario. The Digital Assistant could help flatten 'authority gradients' in the cockpit, and ensure that the best (safest) decision is taken.

Reducing the *Works as Done* vs. *Work as Imagined* Divide

In most safety culture surveys, some people complain that what they need to do in practice to get the job done often disagrees with what it states in the formal job procedures. Such procedures, they will say, have been written by people sitting in offices who have never done the job, or haven't done it for a long time. This leads to a gap between 'work as done' and 'work as imagined'. A Digital Assistant could dispassionately observe how the work is done in real operational contexts, and determine if it is less or more safe compared to the official procedures (including whether the latter are practicable in real working conditions). This information could be used to determine where existing procedures should be maintained, and where they need updating.

CONCLUSION

The six concerns and six affordances are very much the result of a 'thought experiment' by a safety culture expert, ratified by two other safety culture experts. In HAIKU it is planned to further explore these results with other safety culture experts, operational aviation stakeholders, and participants in the six HAIKU use cases. The roadmap for this further exploration is shown in Figure 3.

What is clear is that there is the potential for Digital Assistants to impact on safety culture in aviation. How much of an impact will be more estimable by the end of HAIKU. But already research can begin considering safeguards to put in place, whether in the Digital Assistants themselves, or in the Human-AI Teaming arrangements in future aviation operations, or in the parent organisations. At the same time the safety affordances identified are well worth investigating to see if and how they can be realised in operational aviation systems.

It is likely that similar safety culture concerns and affordances could be identified in other domains, including other transport systems, the energy sector, space and health. It may therefore be useful to carry out cross-domain research into Digital Assistance and safety culture, so that when AIs begin to learn fast how to do human jobs, we do not fall too far behind. At the end

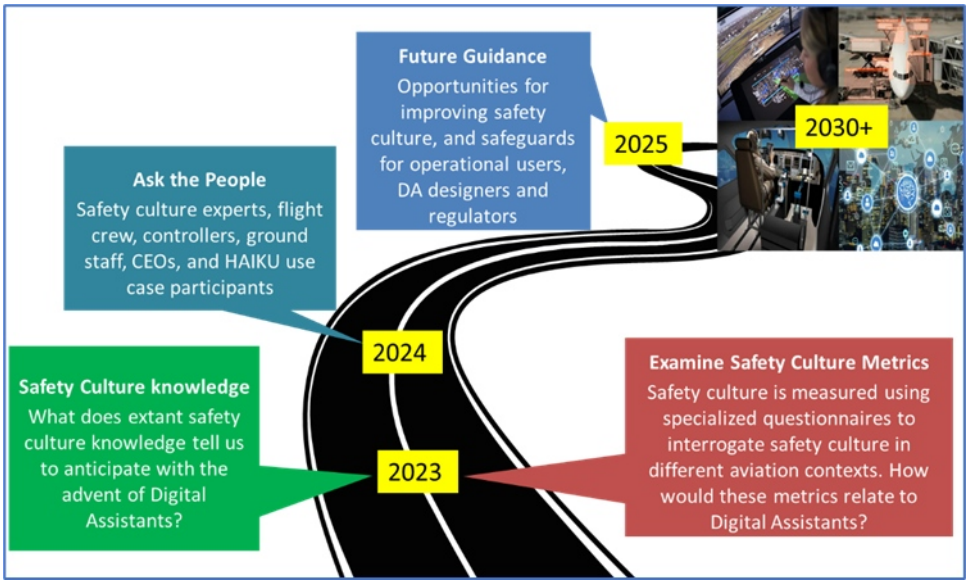

Figure 3: HAIKU Roadmap for Evaluating the Impacts of Digital Assistants on Safety Culture in Aviation.

of the day, humans will always value their (and others') lives more than a machine could ever do.

ACKNOWLEDGMENT

The author would like to acknowledge the support of Andrew Kilner and Beatrice Bettignies-Thiebaux from EUROCONTROL, for early comments on the paper and the concerns/affordances. This publication is based on work performed in the HAIKU Project which has received funding from the European Union's Horizon Europe research and innovation programme, under Grant Agreement no **101075332**. Any dissemination reflects the author's view only and the European Commission is not responsible for any use that may be made of information it contains.

REFERENCES

AAIB (1990) Report on the accident to Boeing 737–400 G-OBME near Kegworth, Leicestershire on 8th January, 1989. Air Accidents Investigation Branch Report 4/90. London: HMSO.

Billings, C. (1996) Aviation Automation: the Search for a Human-Centred Approach. CRC Press: New York.

Advisory Committee on the Safety of Nuclear Installations (ACSNI) Study Group. Third Report: Organizing for Safety. Sheffield: H.M. Stationery Office; 1993

Cullen, D. (1990) The public enquiry into the Piper Alpha Disaster. London: HMSO.

Dias, M., Teles, A, and Lopes, R. (2020) Could Boeing 737 Max crashes be avoided? Factors that undermined project safety. Global Scientific Journals: Volume 8, Issue 4, April, Online: ISSN 2320–9186.

Hidden, A. (1989) Investigation into the Clapham Junction Railway Accident. London: HMSO.

Kalliardos, W. (2022) Enough Fluff: Returning to Meaningful Perspectives on Automation. FAA, US Department of Transportation, Washington DC. https://rosap.ntl.bts.gov/view/dot/64829

Kirwan, B., Reader, T.W., Parand, A., Kennedy, R., Bieder, C., Stroeve, S., and Balk, A. (2019) Learning curve: interpreting the results of four years of safety culture surveys. Aerosafety World, Flight Safety Foundation, January.

Kirwan, B. Shorrock, S.T. and Reader, T. (2021) The future of safety culture in European ATM – a White Paper. EUROCONTROL. https://skybrary.aero/bookshelf/future-safety-culture-european-air-traffic-management-white-paper

Nunes, A. & Laursen, T. (2004) Identifying the factors that led to the Ueberlingen mid-air collision: implications foroverall system safety. Proceedings of the 48th Annual Chapter Meeting of the Human Factors and Ergonomics Society, September 20–24, 2004, New Orleans, LA, USA.

Reader, T. W., Noort, M. C., Kirwan, B., & Shorrock, S. (2015). Safety sans frontieres: An international safety culture model. *Risk Analysis, 35*, 770–789

Reason, J.T. (1997) Managing the risks of organisational accidents. Aldershot: Ashgate.

Turner, R. and Pidgeon, N. (1997) Man-made disasters (2nd edition) Oxford: Butterworth-Heineman.

Zweifel, T.D. and Vyal, V. (2021) Crash: BOEING and the power of culture. Journal of Intercultural Management and Ethics Issue No. 4, 13-26.