MASTER THESIS

Course code: BE323E

Name / Candidate no.: Roy C. Solgård/82 & Jarle B. Opheim/04

Flight Safety Implications associated with the use of Electronic Flight Bags by Norwegian Operators

Date: 15.5.2019

Total number of pages: 101

2019

Flight Safety Implications associated with the use of Electronic Flight Bags by Norwegian Operators

Roy C. Solgård & Jarle B. Opheim

https://brage.bibsys.no/xmlui/handle/11250/92968

Print: Nord University Print Centre, Nord University

Sammendrag

Luftfartsbransjen er kjent for å ha høy grad av konkurranse mellom ulike aktører, som påfølgende skaper store incentiver for å kutte kostnader og øke effektiviteten. Overgangen til digitale løsninger er et eksempel på innovasjon som bringer med seg lovnader om slik økt produktivitet. Som følger av en stadig økende pålitelighet og datakraft, har håndholdte elektroniske hjelpemidler fått økende oppmerksomhet fra ledelseshold. En Electronic Flight Bag (EFB) utnytter dette potensiale ved å øke ytelsen og forbedre sikkerheten relativt til hvordan oppgaver forbundet med administrasjon av flyvningen tidligere var løst i cockpit. Tiltros for at denne økningen i ytelse er godt dokumentert, forblir utslaget på organisasjonens effektivitet i form av virkningen dette har på flysikkerheten mer tvetydig. Bekymringer knyttet til EFB-bruk og flysikkerhet har pekt på at interaksjonen mellom bruker og plattform ikke er optimal og kan utgjøre en sikkerhetsrisiko. I denne studien undersøker vi hvorvidt innfasingen av EFB'er hos Norske operatører har hatt et utslag på flysikkerheten, gjennom å spørre pilotene selv hvordan de opplever disse digitale plattformene relativt til sin mer tradisjonelle cockpit. Videre, så undersøker vi hvorvidt det finnes mulige organisatoriske forskjeller mellom operatørene, og setter søkelys på enkelte menneskelige faktorer som kan ha en effekt på hvor vellykket denne implementeringen har vært. Vi bruker variasjonsanalyse rettet mot å sammenlikne variasjonen innad og mellom ulike grupper piloter, og statistisk modellering til å finne den kombinasjonen av ulike faktorer som best forklarer denne variasjonen. Våre data antyder at opplevd flysikkerhet har forbedret seg innenfor de fleste sikkerhetsområdene som ble inkludert i denne studien. Denne opplevelsen hos pilotene varierer signifikant mellom de ulike operatørene, noe som antyder at organisatoriske forskjeller kan ha en innvirkning på hvor vellykket implementeringen har vært. Menneskelige faktorer hadde også en innvirkning, men disse funnene er mindre robuste. Alt i alt viser våre data at elektroniske hjelpemidler i cockpit har ført til økt effektivitet, og at denne effektiviteten varierer mest ut fra hva slags selskap pilotene flyr for.

Abstract

The aviation sector is highly competitive, driving managers and business leaders to great lengths in order to cut costs and increase efficiency. The ongoing digital revolution affecting almost every aspect of our society, has also gained a foothold with aircraft operators. With the increased reliance and computing power of personal electronic devices, aviation managers are seeking to exploit this potential. The Electronic Flight Bag (EFB) is fast become popular, promising to increase efficiency and improve safety by replacing many of the flight management tasks previously performed on paper or other more cumbersome platforms. Although the efficiency gains are well documented, the affect on organizational effectiveness in terms of flight safety improvements are more ambiguous. Inquiries into EFB usage and flight safety concerns has showed that the human-machine interfaces can pose a threat. Here we set out to test whether the implementation of EFBs by Norwegian operators has had an impact on flight safety, by asking the pilots themselves how they perceive these devices. Furthermore, we investigate whether organizational differences between operators and certain human factors has an effect. We use statistical inference methods aimed at comparing variation between different groups, and statistical modelling to find the best combination of predictors to explain this variation. Our data suggests that flight safety has improved since the implementation of EFBs for almost all the safety areas included in our study. In addition, this perception of flight safety seems to vary between operators, suggesting that organizational effectiveness is not unison across different companies. Human factors were also found to have an impact, but these results are less statistically robust.

Preface

Completing this master’s thesis in Aviation Management at Nord University has been challenging and hard work. The experience of gaining knowledge in new subjects has been exciting, and by utilizing experience from the aviation industry, with routines and structured plans, we managed to balance family, work and studies to complete the thesis. When we now deliver our master’s thesis, it marks the end of our study.

Without the help of LO and Parat unions, along with Widerøe and Helikopter Service Pilot unions, the distribution of the survey would have been quite difficult. It was distributed to all major airlines in Norway, which we greatly appreciate, and a special thanks to all respondents who helped us by answering our survey.

Frode Soelberg (Ph.D.) at North University, have supported us with constructive and detailed guidance. Your feedback throughout the master’s thesis has been of great support and assistance, thank you. We are also grateful to our colleague Bernt Christian Helén (M.Sc.), for his invaluable support and critical questions completing our survey. It has inspired us throughout our work on this thesis.

Finally, a special thanks to our wives Therese Aas Vegsundvåg and Marianne Solgård, for their patience and assistance. Completing this master’s thesis without their support, would not have been possible.

We hope our thesis will be off interest for operators, and inspire to further research on the subject. We have acquired new knowledge through the process, and this has been an exciting project. We want our research to contribute to a positive development through the introduction and use of the EFBs in the future, as well as to focus on areas to be aware of in further development and use.

Roy C. Solgård and Jarle Bøyum Opheim May 2019.

List of figures

List of abbreviations

AC	Advisory Circular
AIC	Akaike’s Information Criteria
AMC	Acceptable Means of Compliance
ANOVA	Analysis of variance
BIC	Bayesian Information Criteria
CAA	Civil Aviation Authority
CASS	Commercial Aviation Safety Survey
CRM	Crew Resource Management
CVA	Canonical Variance Analysis
DOT	U.S. Department of Transportation
EASA	European Aviation Safety Agency
EFB	Electronic Flight Bag
FAA	United States Federal Aviation Administration
ICAO	International Civil Aviation Organization
JAA	Joint Aviation Authority
KS-test	Kolmogorov Smirnov test
MANOVA	Multivariate analysis of variance
NAA	National Aviation Authority

OLS	Ordinary Least Squared
OM	Operations Manual
PFB	Paper Based Flight Bag
POH	Pilots Operating Handbook
PED	Portable Electronic Device
PPM	Policy and Procedures Manual
PM	Pilot Monitoring
PF	Pilot Flying
QMS	Quality Management System
SARP	Standards and Recommended Practices
SMS	Safety Management System
SOP	Standard Operating Procedure.
TGL	Temporary Guidance Leaflet
TØI	Transportøkonomisk institutt
VBM	Value-Based Management

Table of content

(Intentionally left blank)

1 Introduction

1.1.1 Background

The aviation industry is putatively a highly volatile sector. In order to stay competitive and increase organizational efficiency, aviation operators are constantly looking for new and innovative ways to cut costs. The availability of more reliable and powerful software and hardware solutions integrated in various electronic platforms, represents one area which air operators are seeking to exploit (Scott, 2007; Carey, 2013). The increased efficiency resulting from safer information distribution, larger data handling capabilities, and automation, contribute to a more efficient organization (Airplanes, 2005; Higdon, 2017). Although more constrained by regulations, aircraft manufacturers and suppliers of various on-board systems has followed the general trend towards the more paperless society facilitated in part by these devices. The term *Electronic Flight Bag* (EFB) refers to the digital revolution concerning flight deck management, where EFBs can replace many of the features and functions that traditionally has been performed on paper (CASA, 2013). Since the introduction of Electronic flight bags (EFBs) into civil aviation around 15 years ago, these digital solutions have now also become a common tool for flight crews in many major Norwegian air carriers. Their increased popularity is fuelled in part by the increased demands placed on aviation businesses to be cost effective in an ever more competitive marked. EFBs allows operators to cut costs through increased organizational efficiency (e.g., Ates, 2017). Specifically, this is achieved by placing more tasks on flight crews which previously required more expensive and often too inflexible auxiliary services (e.g. third-party calculation of mass and balance and various performance aspects of the flight). In addition, EFBs provides more effective means for the operational management branch to distribute information such as bulletins and procedural updates, and it provides flight deck crews with more pertinent access to relevant documentation. The whole philosophy of the EFB concept is to replace the traditional and cumbersome flight bag, and by doing so, facilitate for a more effective organization and a safer conduct of flight operations.

While the benefits related to EFBs are putatively beneficial from an efficiency standpoint, its impact on organizational effectiveness and flight safety are more ambiguous. Concerns about task saturation, increased workload, more “head down” time, and various other potential hazards needs to be fully understood in order to mitigate risks (Chandra et al., 2003). Numerous publications have addressed how pilot-EFB interaction might affect flight safety (Chandra et al., 2009; Volpe, 2010; Volpe, 2014; Tump et al., 2014). In Norway however, there has not been conducted any large-scale studies into the potential safety impact of EFBs, despite the widespread usage of these digital solutions in all major Norwegian carriers. The ability of EFB solutions to mitigate safety threats is ultimately the measure of success for a typical safety-oriented organization (e.g., Wong et al., 2005). As the reliance on these digital solutions only increase, a more in-depth understanding of its impact on flight safety is arguably warranted.

1.1.2 The Electronic Flight Bag

The traditional flight deck contains large volumes in paper. Requirements imposed by regulators, as well as the complex nature of safe aircraft operation, meant that cockpit crews needed to have access to this information while on duty. Collectively this material and the luggage it was carried around in, was referred to as flight bags (i.e., *paper-based flight bags*; PFBs). Documents containing operational procedures, aircraft operating manuals, navigational charts, flight plans, and other information pertinent to a particular flight, made these PFBs cumbersome since they added to considerable weight. More importantly, the effort of finding information, processing it, and produce a reliable and accurate result took considerable time and was more prone to errors (Airplanes, 2005). This was particularly so in tasks such as the calculation of critical flight parameters (e.g., mass and balance, and performance calculations). If one defines efficiency as the elapsed time between the desired results and the efforts employed to achieve that outcome, PFBs represented a great potential for improvement. Subsequently, pilots and operators in the 1990s stared to employ various electronic platform to help administer flight deck duties. The term electronic flight bag emerged to describe these electronic display systems intended primarily for cockpit or cabin use.

As development in technology has progressed and demand from operators increased, regulators have adapted the rules of the air accordingly (fig. 1). As there are a diverse number of digital solutions with different levels of aircraft integration, the U.S. Federal Aviation Administration (FAA) published Advisory Circular (AC) 120-76A, Guidelines for the Certification, Air- worthiness and Operational Approval of Electronic Flight Bag Computing

Devices. The EFB hardware was divided into three subcategories, Class 1, 2, and 3. The hardware required different approvals off the authority depending on their class.

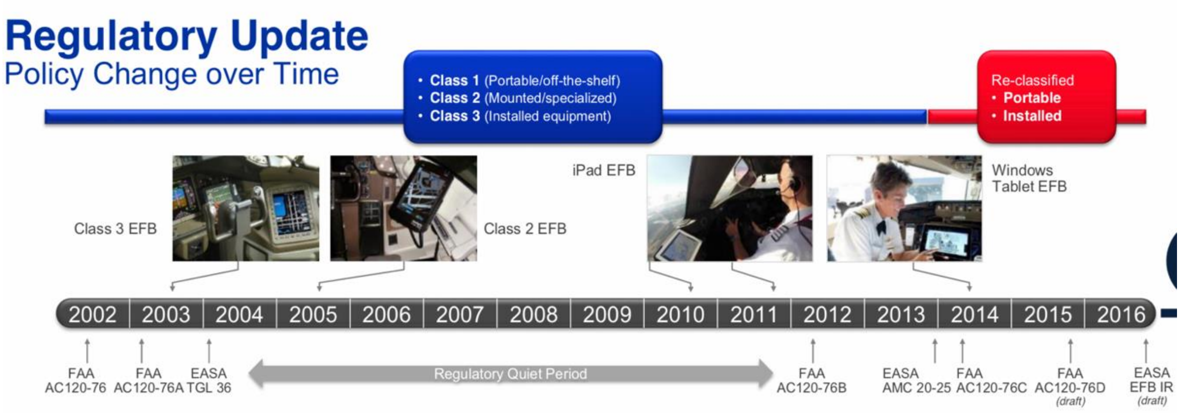


Figure 1. EFB policy timeline. Regulators in both the US and Europa has issued a series of advisory circulars in order to help manufacturers, users, and management in handling the EFB transition more safely.

Class 1 EFBs are defines as standalone laptops, PalmPilot’s etc., normally used to support the operation during per-flight preparations but not necessarily used during flight. Class 2 EFBs are technically un-installed items but could be mounted in the flight deck. Most Class 2s will be able to receive data from the aircraft but could not send any data to the aircraft. Subsequently, new airworthiness-requirements is not required to be implemented. Class 3 EFBs are despite class 1 and 2, considered installed as a part of the aircraft hence requiring airworthiness approval (FAA, 2002). In Europe, The European Aviation Safety Agency (EASA) developed similar regulations for commercial aviation. EASA has worked with modernizing the rules governing EFBs, as well as harmonizing them with other worldwide standards and best practices. EASA make the distinction between portable platform, which is not part of the certified aircraft configuration, and installed platforms, an EFB host platform installed in the aircraft and considered as an aircraft part, covered, thus, by the aircraft airworthiness approval (EASA, 2014). Furthermore, EASA classify software solutions by the consequences of failure, where type-A applications imply no safety effects, while type-B application failure would impose a safety threat.

Today, modern EFB devices can display a variety of aviation data: checklists, navigation charts, pilot’s operating handbook (POH) or perform basic calculations (e.g., performance data, fuel calculations). The scope of the EFB system functionality may also include various other hosted databases and applications (Flight safety foundation, 2005). The EFB enables electronic documents to be reached and used in flight, which can even be integrated with airborne avionic systems. As the varieties of applications used and the content management and distribution systems in EFB changes, the specific benefits of EFBs are also changing and updating respectively. In more recent years, the EFB systems have become even more cost effective as more off the shelve hardware has been able to provide the necessary computing power and software integrations (McKenna, 2013).

> *"According to United Airlines, a conventional paper-based flight bag contains an average of 12,000 sheets of paper. The airline estimated that deployment of an EFB system running on Apple iPads would save the airline nearly 16 million sheets of paper a year, as well as save 326,000 gallons of jet fuel due to the reduced weight on board the aircraft".*
> (Computer Science and Information Technology 5(4): 128-134, 2017)

Modern day uses of EFBs include everything from access to information, to providing a primary source of navigation in a fully integrated digitized flight deck. Today, EFBs are in some form or another, implemented by most major western carriers.

1.1.3 Organizational and human factors

From an organizational standpoint, managers need to consider both organizational and human factor elements when implementing new procedures or technologies (Smith et al., 1995). The factorial output of the organization has become safety itself, where effectiveness is measured in frequency of unwanted events (Flouris et al., 2009). Subsequently, the organization is only viable to the extent it can mitigate threats and sustain acceptable levels of safety violations. To achieve this level of safety, an evaluation of both formal and informal organizational elements and their contributions could improve effectiveness (Jacobsen & Thorsvik, 2013). As there are strict rules and regulations to be followed, a clear definition of responsibility, and a high degree of formalization, the typical aviation operator's organizational typography often resembles a machine bureaucracy structure (Jacobsen & Thorsvik, 2013). There is a clear horizontal and vertical specialization into different areas of expertise, and there is a widespread use formal element to dictate organizational behaviour. There is a major reliance of standardized operating procedures (SOP) in place that facilitates strict control of operations, often accompanied with a hierarchical managerial structure. This formal organizational structure has historically received the most attention with operators as it was believed to affect the behaviour of individuals the most (Jacobsen & Thorsvik, 2013). The individual freedom to choose how to perform tasks has been sacrificed for the need for standardized procedures and uniformity of operations. In recent years however, the informal elements have received increased attention, particularly within the aviation sector (Helmreich & Merritt, 2017). These informal elements consist of norms, traditions, social networks, human emotions and attitudes, culminating in an organizational culture and how this affects effectiveness. This has also been recognized in aviation safety management, as a very important part of a well-functioning Safety Management System (SMS) wish is the reliance on safety reporting, and that it encourages and facilitates a safety culture based on *just culture* (Reason, 2016). The way you are treated in a just culture, contributes to trust and respect, which, according to ICAO's descriptions of safety culture, is dependent on. Just culture is crucial in achieve a high level of reporting, which is the prerequisite for a well-functioning SMS system. The counterweight to just culture is *blame culture* or fear culture, where one is looking for a placement of guilty subsequently blaming the individual for the incident (Woods, Decker & Cook, 2010). Blame culture often reflects poor leadership, where leaders will hold others accountable when something goes wrong rather than taking full responsibility and encouraging and seeking actions that in the future will be similar events. A third variant is *no blame culture*, which is an organizational approach that is categorized by sensible attitudes to errors and near misses, based on the assumption that there is no perfect system. It is important to create an atmosphere of trust where employees are encouraged and rewarded to report potentially dangerous catastrophic errors (Proverva, Montefusco & Canato, 2010).

As we have seen, the effectiveness of a typical modern aviation organization is dependent on a multitude of different elements. However, business managers also must consider not only the effectiveness of the organization, but also its efficiency (Ostroff et al., 1993). An organization can be quite effective at producing flight safety, but it can be quite inefficient in doing so, using up vast sums of resources in the process. Any rational organization will seek to limit its expenditure, maximizing its results using the least amount of recourses required. Without such an approach, the organization will not be sustainable in the long run (Jacobsen & Thorsvik, 2013). Hence, in the interaction between efficiency and effectiveness in aviation organizations, there lies a potential for conflict of interest (ICAO, 2012). An organization is ultimately forced to define its safety goals by balancing production towards an acceptable level of risk. The aviation industry has long recognized this, and ICAO describes this as “The management dilemma” (ICAO, 2012). In practice, a safety buffer in the form of a “safety space” is established. In order to maintain this barrier, important contributors include training, internal processes and procedures. In addition, the development and implementation of new technology can farther strengthen the barrier by making the organization more efficient. Through increasing efficiency, the organization can maintain the desired level of effectiveness through the spending of fewer resources. The electronic flight bag (EFB) represents one such effort, promising to increase efficiency by providing a reliable digital platform intended to replace many of the tasks traditionally handled on paper.

1.1.4 Typical integration of EFB in an organization

Business managers in Norway has also recognized the potential efficiency gains to be had from the digital flight deck management transition, as most major operators have now incorporated some sort of EFB solution to their respective flight decks. A typical integration of EFBs can be considered as a sequence of four phases (Borgen, 2018):

In the first phase of the EFB implementation, it was only maps, and other documentation that were transferred to a digital platform. Previously, the pilots used maps, as well as operative aircraft operation manuals in paper version. Required maps used in the flight operations were taken out of the company's routing manual and were subject to updates once a week, where pilots had their own personalized route manual responsible for carrying out mandatory updates. On average, the pilot's former manual (EFB) weighed between 15 and 20 kg. By digitizing it, it was possible to meet the first cost savings through reduced weight on the aircraft. In addition, the chart subscription for each pilot was less expensive, as the transmission of these per letter item ceased. The first phase of the change process therefore included only limited training and use. The first approval for the EFB project, which was issued by the Civil Aviation Authority (Norway), assumed that the pilots should have access to maps and documentation both electronically and in paper format.

In the second phase, the EFB was clarified to be able to process the operational flight plan digitally. Traditionally, the work on the operational flight plan was very primitive. The operational flight plan contained information on which route and altitude the flight was scheduled with, and the pilots themselves had to use their skills to calculate the necessary fuel and choice of alternative airport if the flight could not be carried out as planned. Otherwise, all flights were planned with a fixed route and altitude, regardless of weather conditions, traffic, airspace closure and other factors that could affect the flight. In addition, the choice of speed in the different phases of flight was determined by the pilot himself, but there existed a culture to always fly at maximum speed. The operational flight plan was printed in paper format at the moment the pilot checked in to the company's computer systems. Since the operational flight plan now are processed digitally, there were opportunities that did not exist before. By digitizing this platform, it was now possible to handle large amounts of parameters. These parameters provided the basis for issuing extremely precise operational flight plans, taking into account factors such as weather, traffic, airspace and aircraft performance. The result was that each flight was planned according to the parameters that the company itself chose to use. In Widerøe, all flights are planned based on conditions that provide the most cost-optimal flight. In practice, the pilot gets access to this information in the operational flight plan and is used in the further planning. This, however, requires the pilots to change their established routines and work patterns and acquire attitudes that match the new regime. This phase of the change process therefore included changes in working methodology, procedures, culture, competence and attitudes. These factors were important to handle properly by operational management to ensure the desired result.

In the third phase, weight and balance calculation were included in the EFB system. Previously, this was handled by qualified ground staff at the departure gate. The work consisted primarily of planning the aircraft's cargo consisting of passengers, luggage and freight. Through computer systems that only the ground staff had access to, then the weight and the balance of the aircraft were calculated before the cargo report was handed over to the pilots before each departure. The pilots had to study new knowledge about weight and balance calculation while at the same time change their working methods.

In the fourth and final phase, the calculation of the aircraft's services during departure and landing is integrated and transferred to the EFB platform. These were calculations that were previously performed with paper tables and a simple electronic performance calculator. The data used were partly roughly rounding figures, which caused limiting departure weights that could have consequences for the number of passengers and luggage and, likewise, increased engine power setting during a flight. These were factors that were expensive. After a new digital version, the pilot now operates with dynamic and detailed figures for the entire flight from A to B. This gives a better overall picture for the pilot.

After the digitizing in to EFB, it became possible to look at this operation in a whole new spectre. The idea was to transfer this responsibility to the pilots, so that the human resources in this operation were significantly reduced. The consequences were that, as stricter requirements for calculations and calculations are made before departure are filled in correctly. In order for one to fail, the pilots do their own calculations before comparing the data and then following a fixed procedure for programming the aircraft. Finally, a copy of the calculations is submitted to the company's database before each departure. Key elements that were affected here were therefore training, competence and change in relations to leadership.

1.1.5 Study purpose and research questions

"We are investigating whether the implementation of EFBs by Norwegian operators has resulted in a change in flight safety."

The EFBs has putatively improved efficiency to such an extent that it remains highly unlikely that the industry will devolve to a paper-based system. However, the ultimate measure of the success of EFBs, will be to what extent it has strengthened or weakened the acceptable risk buffer. With improvements in efficiency comes promises of improvements in effectiveness, as safety margins are allegedly increased. Indeed, one of the strongest arguments supporting an EFB-transition are flight safety gains. The EFB-implementation by Norwegian operators thus provides an excellent opportunity to evaluate if efficiency translates to effectiveness in high-risk organizations. Here we set out to test the effectiveness of the EFB-transition by quantifying how pilot's perception of flight safety has changed since the implementation of EFBs to their flight decks. We investigate whether flight deck crews experience any difference in their ability to conduct a safe flight after the introduction of EFBs. Furthermore, we analyse if organizational and human factors have any impact on how the individual pilot perceives the transition from a pre-EFB flight deck environment. These insights might help both managers and regulators in gaining a better understanding of how the implementation of EFBs and similar large-scale operational transitions is affected by organizational and human factors.

1. Has the implementation of EFBs by Norwegian operators increased organizational effectiveness?

In aviation, effectiveness is ultimately measured in terms of flight safety gain or loss. We use the individual pilot experience as a unit to measure the effectiveness of EFB implementation across all major Norwegian operators.

2. Are there organizational differences between operators?

From an organizational standpoint, both informal and formal structures affect the effectiveness of the organization. We propose that the degree of variation between operators acts as a proxy for differences in various organizational elements between large Norwegian operators that could be investigated further in order to make similar implementations more successful and effective.

3. What is the importance of pilot background, experience, or attitudes?

From a managerial perspective, it is not only important to evaluate from an organizational standpoint, but also to understand how everyone contributes to the system. Particularly in aviation, where a lot of responsibility and trust is put on each individual, valuable insights could be gained through a better understanding of the human factors involved. Individual differences in employee background, experience, or attitudes could contribute to how pilots perceive the EFB transition. We hypothesized that some of these human factors could help explain potential differences in answers between participants across all operators in Norway.

1.1.6 Scope

This study aims at using statistical methods to infer relationships between a set of predetermined predictor and response variables. These variables are selected for their relevance in the literature and based on the authors combined experience from aviation. Hence, the study is limited to discussing the relationship between this particular draft of potential independent and dependent variables. Other conclusions might result if another set of variables where to be included. Data is gathered and quantified from self-reported responses from a questionnaire distributed to pilots in Norway speaking one of the three Scandinavian languages. The study incorporates responses only from major commercial aviation operators in Norway, and has no data of minor operators, the business aviation sector, or the private/aero sport community. The study concerns itself with flight safety and EFB usage, focusing particularly on the importance of organizational affiliation and attitudes, background, and experience of the pilot respondents.

1.1.7 Reading guide

The study proper is divided into five sections. In **section 2**, we outline the theoretical framework focusing on the development of flight safety theory and how this is incorporated at the organizational level. We then go on to review some important progress and findings made in flight safety research from an EFB perspective. These insights are used to design the 16 response variables in the questionnaire (fig. S1), so that flight safety can be accurately quantified according to our EFB oriented perspective. **Section 3** is devoted to discussing the methods used. A brief description of the study design, survey design, and question selection regime make up the first part of this chapter. The second part is reserved for the analysis itself, divided into three sections each of them representing one of the three research questions. This latter structure is mirrored in **section 4**, where the results are presented. This welcomes the reader to easily evaluate the methods used and the results attained by them, while being reminded about the research questions as they form the headline of each section. In **section 5** we proceed with a discussion of the findings. We elected not divide up this section into sub-sections, as we believe that the results are best interpreted as a coherent whole. Finally, **section 6** make up the conclusion where we summarize the main findings by revisiting each of research questions. We explain some limitations and propose potential ways forward in exploring the relationship between EFB usage and flight safety.

The reader will be advised to refer to the Appendix at the very end of this study for support material. Several figures and tables can be reviewed here for more in depth details of the statistical analysis in particular.

2 Theoretical framework

"If you think safety is expensive, try an accident."

(Dr. Trevor Kletz)

This chapter presents the framework that will form the theoretical foundation needed to pursue our research questions. To evaluate flight safety with regards to EFBs, four main topics were deemed crucial; *access to information, distraction, situational awareness,* and *workload.* The survey that the participants completed, would give them the opportunity to compare how they consider flight safety after the implementation on EFB, to how they experienced the traditional flight deck. To better understand how this is affected by the entire organization, we refer to models and systems that are used to measure safety management within the organization, as well as highlight the field of human factors that is so central in flight safety theory and research.

A crucial step in answering our research questions involves quantifying flight safety perception in pilots. As with any study that involves human behaviour, a key challenge is to design study questions that can measure responses from the survey participants. To do this as accurately as possible, we based our questions on developments in flight safety theory, outlined below.

2.1.1 The evolution of flight safety

Traditionally, flight was regarded as a high-risk activity, but the industry has over many years undergone major changes with technological developments and the introduction of modern passenger aircraft, as well as a better understanding of the interaction between people and technology. This due to unfortunate experience from accidents and near-accidents, led to improving routines, procedures, education and training, and the focus on the human factors has been strengthened. The United Nations International Civil Aviation Organization (ICAO) looks at developments in aviation safety as three eras (fig. 2).

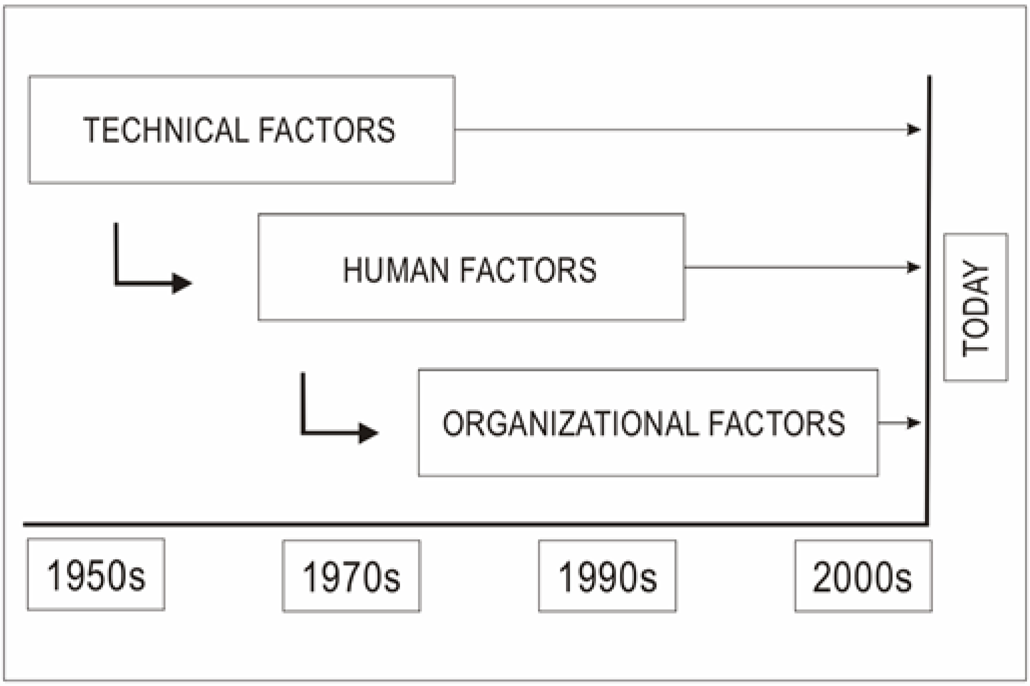


Figure 2. The evolution of flight safety. There has been a steady development towards incorporating other aspects than purely technical malfunctions etc., to the point of recognizing that organizational behaviour influence flight safety profoundly.

The technical era, from the early 1900s to the late 1960s, aviation developed into a form of mass transport, where safety deficiencies were discovered, primarily due to technical factors and technical errors. The safety focus was therefore to investigate and improve the purely technical. In the 1950s, technical improvements led to a gradual decrease in the frequency of accidents, which led to safety work being extended to regulatory compliance and oversight (ICAO. Int).

The subsequent era from the early 1970s to the mid-1990s is characterized by the human factors, dubbed aptly the *Human factors era*. The number of aviation accidents in this era was significantly reduced due to major technological advances and improvements in safety regulations. Aviation became a safer form of transport, and aviation safety was expanded due to human factor related issues, including human and machine interaction. This led to a search for safety information beyond what was generated by previous research. Despite changing focus and investment in new aviation safety tools, human performance continued to be mentioned as a recurring factor in accidents. The use of new methods within the flight safety tasks aimed at human factors (Human factors) focused on the individual, without completely considering the operational and organizational context. It was not until the early 1990s that it was first recognized that individuals operate in a complex environment, which includes several factors that have the potential to affect behaviour (ibid).

Hence *the organizational era* from the mid-1990s to present, began. During this period, one started to look at safety work in a system perspective, which included organizational factors as well as human benefits and technical factors. The term "the organizational accident" was introduced and one began to look at the impact of organizational culture on risk management (ICAO, 2012).

2.1.2 Flight safety models

Much of the theoretical framework stems from years of research into accident prevention and mitigation. A well-known risk analysis in aviation is the Swiss Cheese model, developed by James Reason of the University of Manchester (see Reason et al., 2006 for an overview).

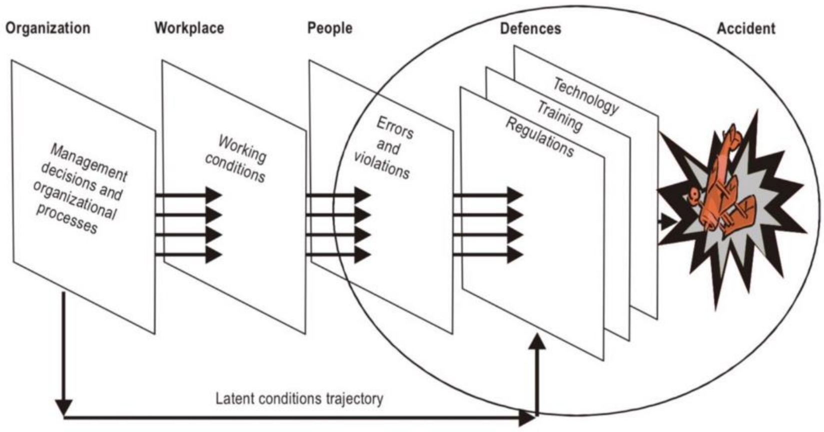


Figure 3. The Swiss cheese model. In order of accidents to occur, events must pass through a series of "holes" in a "swiss cheese".

The model graphically illustrates the complexity of accidents and is used in risk analysis and risk management in aviation safety. It deals with human causes of accidents, symbolized as slices of Swiss cheese set up behind each other, where each slice represents a risk level. Reason's hypothesis was that most accidents could be traced back to 4 causal levels induced by four cheese slices, with the first three layers representing latent causes (errors), while the last being active errors that trigger the actual event and accident. The latent causes are attributed to organizational influence, Unsafe supervision, preconditions for unsafe acts, on which the latter represents unsafe acts. Types of defence are layered behind each other and protect against the risk of developing into an accident. Former air safety efforts focused largely on identifying the active failure that triggered the accident, such as pilot failure or technical failure. This model represents latent states that may be present for a long time before an event results in an accident. In the model, an original defence against accidents is illustrated as a series of barriers represented as slices by the Swiss cheese. The holes in the cheese slices represent the individual weakness of the system and are variable in size and position. The system as a whole will produce errors when the holes in each of the layers are aligned and coincide so that the accident passes. (Reason, 2016). One way to use the model can be to reduce the possibility of human error by inserting several barriers into the system, as this will cause more things to go wrong before the accident is a fact. We also have such barriers in the form of legislation and regulations, supervision and at the organizational level. If one had managed to eliminate all the risks, such protective layers had been 100 percent tight, but in reality, they are more like slices of Swiss cheese with several holes. The holes in the cheese slices represent latent errors and defects in the different layers. Although there are several holes, a threat will often stop in one of the other barriers to avoid an unwanted event. But if circumstances cause holes in many layers to fall in line, there will be the possibility of an accident.

Another influential framework is represented by the SHELL model (Edwards, 1988). This is a conceptual model of human factors that clarifies the scope of aviation human factors and assists in understanding the human factor relationships between aviation system resources/environment (the flying subsystem) and the human component in the aviation system (the human subsystem). The SHELL model was first developed by Elwyn Edwards (1972) and later modified into a 'building block' structure by Frank Hawkins (1984). The model is named after the initial letters of its components (software, hardware, environment, liveware) and places emphasis on the human being and human interfaces with other components of the aviation system. (ICAO, 2019).

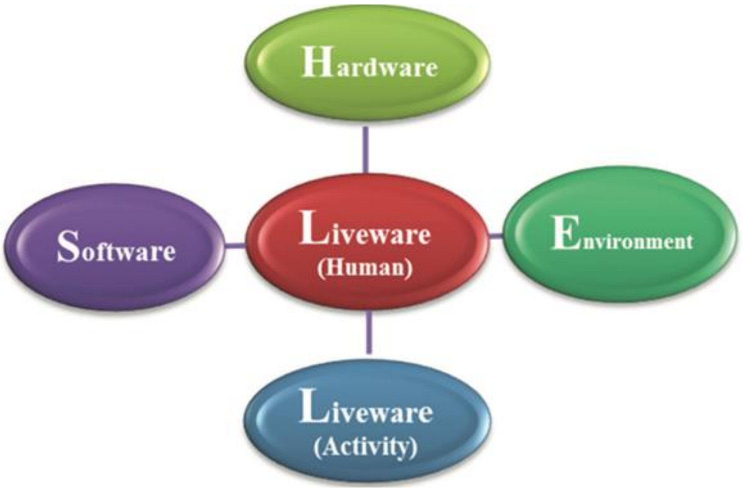


Figure 4. The SHELL models**.** *Software* - the rules, procedures, written documents etc., which are part of the standard operating procedures (SOP). *Hardware* - the Air Traffic Control suites, their configuration, controls and surfaces, displays and functional systems. *Environment* - the situation in which the L-H-S system must function, the social and economic climate as well as the natural environment. *Liveware* - the human beings - the controller with other controllers, flight crews, engineers and maintenance personnel, management and administration people - within in the system.

The SHELL model adopts a systems perspective that suggests the human is rarely, if ever, the sole cause of an accident. The systems perspective considers a variety of contextual and task-related factors that interact with the human operator within the aviation system to affect operator performance. As a result, the SHELL model considers both active and latent failures in the aviation system. Each component of the SHELL model represents a building block of human factors studies within aviation. The human element or worker of interest is at the centre or hub of the SHELL model that represents the modern air transportation system. The human element is the most critical and flexible component in the system, interacting directly with other system components, namely software, hardware, environment and liveware. However, the edges of the central human component block are varied, to represent human limitations and variations in performance. Therefore, the other system component blocks must be carefully adapted and matched to this central component to accommodate human limitations and avoid stress and breakdowns (incidents/accidents) in the aviation system. To accomplish this matching, the characteristics or general capabilities and limitations of this central human component must be understood.

2.1.3 The safety organization

As we have seen, in the field of safety research over the last 30-40 years, a realization has been made that human actions including mistakes cannot be understood in isolation but must be regarded as part of a system (summarized by TØI, 2005). To understand how the pilots will be influenced by this system; i.e., the organization they are employed in, and part of - we need to look at how the safety organization is structured. According to Jacobsen & Thorsvik (2013), organizations are regarded as production systems. Because organizations have one common goal of producing something, we distinguish these from other social groups. This means that all organizations must produce something. Organizations are designed in a way that some believe is the most effective way to solve the task. In aviation, effectiveness is measured in how the organization handles risks. Years of flight safety research has led to new and innovative ways to promote flight safety in the organization, increasing its effectiveness through a deeper understanding of how both formal and informal organizational traits affect flight safety. Subsequently, we view the safety organization in the context of the flight safety management approach (ICAO, 2012).

The formal elements of a flight safety organization relate how the organizational structure is adapted to handling risk. In addition to traditional data collection and analysis, limited to data gathered by accident and serious incidents, the safety work has been extended with a new proactive approach. Proactively collecting and analysing data was initiated routinely, and not as previously only reactive to monitor known safety risks and to detect emerging safety issues. This improvement provided the basis for moving towards a "safety management" (SMS) or safety management approach (ICAO, 2012). SMS is a comprehensive safety system that is required by law for all airlines that are registered in and operate in a country affiliated with the EU. The system was introduced in connection with the renewal of a common European regulatory framework in 2014 by the EU agency EASA. This system is designed to continuously improve safety by Identify hazards, collect and analyse data, and continually assess safety risks. Through SMS, one proactively seeks to control or reduce risk before it results in accidents and incidents. (ICAO, 2012). The system is also intended in providing the operator with a framework for establishing and developing a positive air safety culture in the organization. According to Stoltzer and Goglia (2008) SMS is defined as: “(…) *a dynamic risk management system based on quality management system (QMS) principles in a structure* *appropriately on the operational risk, applied in a safety culture environment."*

The scope of the SMS is based on the ICAO framework and EASA regulations for SMS and highlights how flight safety saturates and defines the way in which organizations needs to be structured. It consists of four pillars of safety management (ICAO, 2012; fig. 5).

Figure 5. The four pillars of safety management.

Safety Policy and Objectives; every business must have set guidelines, procedures, and organizational structure to achieve its goals. The most important thing is that safety must be rooted in the organization, through the management's defined strategy, methods, processes and organizational structure to achieve the goals (FAA, 2015). Safety Risk Management; safety risk management is the analysis and elimination, and mitigation to an acceptable level of the safety risks of the consequences of identified hazards. Safety Assurance; safety assurance activities focus on assessing the health of the organization, with an emphasis on safety. safety Promotion; operators must identify safety training requirements for each level of management and for each employee group.

The safety organization must set guidelines, procedures, and organizational structure to achieve its goals. The system needs procedures in place that describe responsibility, authority and expectations. The most important thing is that safety must be rooted in the organization, through the management's defined strategy, methods, processes and organizational structure to achieve the goals (*ibid*). The commitment of the senior management to safety is reflected in a policy statement, which is signed by the Accountable Executive. Safety management activities must be documented appropriately and be available to all employees. In addition, a system for Safety risk management must be implemented. This entails the analysis and elimination, and mitigation to an acceptable level of, the safety risks of the consequences of identified hazards. The objective of risk management is a balanced allocation of resources to address all safety risks and viable safety risks control and mitigation. It is a data-driven approach to safety resources allocation. This leads to *safety assurance* activities, which focus on assessing the health of the organization, with an emphasis on safety. Specific goals for improvements in all areas should be set for all senior operational managers. Safety assurance should include monitoring of external sources of safety information and include participation in regional safety groups or safety data sharing organizations. External or internal changes may introduce new hazards to operational activities.

Processes must exist to manage organizational responses to regulatory changes, major changes in operational procedures, or new activities such as new airport destinations. Safety reporting systems should have processes established to identify new risks and actively monitor performance in new areas of the operation. Finally, the ICAO and EASA regulations recognize the need for *safety promotion*. Operators must identify safety training requirements for each level of management and for each employee group. Safety training for operational personnel should address safety responsibilities, including complying with all operating and safety procedures, recognizing and reporting hazards, and ultimately ensuring that employees have the knowledge and skills to safely complete work activities. Communication of safety information is a key responsibility of the Safety Manager. Continuous improvement and learning are accomplished through the sharing of lessons learned from investigations, hazard report analysis, and operational safety assessments. Feedback to operational personnel, such as examples of procedural improvements as a result of safety reports, is an essential feature of safety communications.In addition to these formal elements, there also seems to be broad consensus that it is possible to distinguish between different organizations or businesses based on its safety culture. According to Pidgon and O´Leary (1994), safety culture is defined as:

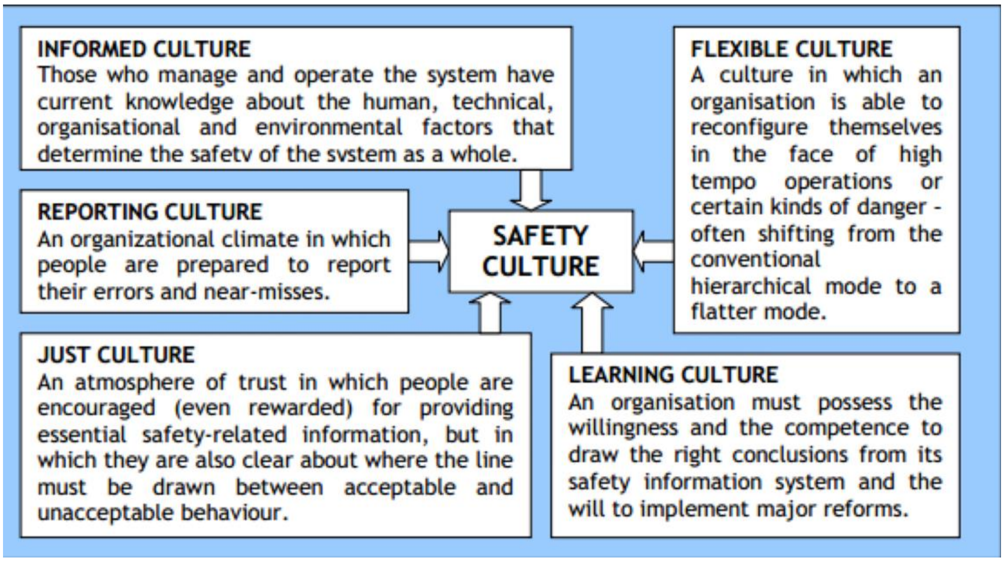
> *“(...) the set of beliefs, norms, attitudes, roles and social and technical practices within an organization which are concerned with minimizing the exposure of individuals both within and outside an organization to conditions considered to be dangerous (...)”*

In Jacobsen & Thorsvik, 2013, it is argued that a strong organizing culture will gather employees in a common experience of belonging and community. Organizational culture can have a strong impact on members of the organization, such as the formal structure, and appear to have five general effects on behaviour. Culture will create belonging and community (Mayo, 1945), which will create a sense of safety, and that will make groups act and think relatively equally. The community and identity that a strong organizational culture can provide will be motivating (Deal & Kennedy 1982; Lincoln & Kalleberg 1990). With common goals and values in the culture, it will create motivation to perform, and one can get the feeling of working as a team, rather than alone, which has also proven to be a strong effect on motivation (Latham & Pinder 2005).

Between management and employees, trust is an important factor. By a strong culture, the stronger the better it is claimed, confidence will lead to the need for less monitoring. The employees will also act in accordance with the norms of the organization / group they are members of (Fukuyama 1995; Gold & Campbell 202). When we have studied that culture creates belonging, experience of fellowship and trust between those in the culture, this will then lead to collaboration and coordination (Jacobssen & Thorsvik 2013) becoming easier. This will apply between people, groups and hierarchical levels. When you trust each other and feel a kind of mutual interdependence, you want to work for this community rather than for your own interests. One will be able to develop a common language, a common way of communicating, to facilitate collaboration (Chatman & Spataro 2005). The culture will often provide *recipes* on how employees should act in given situations (Ray 1986; O`Reilly 1989; Pfeffer 1997), a kind of governance or structure. What we could call *sense-making*, because culture then contributes to the employees emphasizing conditions that are considered important in culture (Perrow 1986; Barley et al. 1988; Huang & Wu 1996). Organizations with extensive trust between managers and employees can be more effective than an organization using traditional bureaucratic-administrative management tools (Peters & Waterman 1982; Fukuyama 1995).

Indeed, a very important part of a well-functioning safety management system is safety reporting and that it encourages and facilitates a safety culture based on J*ust* C*ulture*. The way you are treated in a just culture, contributes to trust and respect, which, according to ICAO's descriptions of safety culture, is dependent on. Just Culture is so important to achieve a high level of reporting, which is the prerequisite for a well-functioning SMS system. The counterweight to just culture is blame culture or fear culture, where one is looking to find the place that is guilty of an unwanted event and then blame the person for the incident (Woods, Decker & Cook, 2010). blame culture often reflects poor leadership, where leaders will hold others accountable when something goes wrong rather than taking full responsibility and encouraging and seeking actions that in the future will be similar events. A third variant is no blame culture, which is an organizational approach that is categorized by sensible attitudes to errors and near misses, based on the assumption that there is no perfect system. It is important to create an atmosphere of trust where employees are encouraged and rewarded to report potentially dangerous catastrophic errors (Proverva, Montefusco & Canato, 2010).

Benefits from having a just culture versus a blaming culture would be increased safety reporting, thrust building, and a more effective safety and operational management.
A just culture support learning from unsafe acts in order to improve the level of safety awareness through the improved recognition of safety situations and helps to develop conscious articulation and sharing of safety information (Flight safety, GAIN Working Group E 2004; fig. 6).



INFORMED CULTURE
Those who manage and operate the system have current knowledge about the human, technical, organisational and environmental factors that determine the safety of the system as a whole.

REPORTING CULTURE
An organizational climate in which people are prepared to report their errors and near-misses.

SAFETY CULTURE

FLEXIBLE CULTURE
A culture in which an organisation is able to reconfigure themselves in the face of high tempo operations or certain kinds of danger - often shifting from the conventional hierarchical mode to a flatter mode.

JUST CULTURE
An atmosphere of trust in which people are encouraged (even rewarded) for providing essential safety-related information, but in which they are also clear about where the line must be drawn between acceptable and unacceptable behaviour.

LEARNING CULTURE
An organisation must possess the willingness and the competence to draw the right conclusions from its safety information system and the will to implement major reforms.

Figure 6. The safety culture complex. A positive safety culture is itself dependent on strong subcultures

2.1.4 Flight safety and EFBs

Safety research and how it relates to EFB usage was from the beginning mostly performed by various civil aviation authorities. In 2003, the Federal Aviation Administrations issued an advisory circular which was the first set of guidelines on this subject produced by a civil aviation authority (FAA, 2003). Here, the FAA was especially concerned with how the human-machine interface could affect flight safety. The EFB needed to be as good as an existing paper-based system, and that it did not result in unacceptable levels of flight crew workload. Significant contributions have been made by Chandra and colleagues, as well as Volpe (a U.S Department of Transportation subsidiary) (Volpe, 2018). These publications help in providing a framework to evaluate EFB functionality, and aid operators, regulators, and users in how to deal with some of the challenges posed by EFBs. Subsequently, they form much of the scientific anchorage behind the Federal Aviation Administration advisory circulars (ACs) and are especially concerned with the human factor considerations (e.g., Chandra et al., 2000; 2000a; 2002). Chief among these are concerns about automation and efficiency.

Automation complacency arise when a computer lulls a user into a false sense of security (Carr, 2013). Having confidence in the computers abilities to handle a task, the user will allow heirs attention to drift or not maintain as much vigilance as the situation might require. *Automation bias* occurs when users put too much faith into information arising from automation, that they begin to ignore other sources of information (Carr, 2013). The EFB also represent a potential source of distraction on the flight deck. Chandra makes the point that when the design of a new system enters the flight deck, it’s important that the pilot’s expectations of how the aircraft operate are not violated by the EFB. If the EFB is incompatible with the cockpit design of philosophy, pilots will have trouble learning to use it, and are subsequently more prone to errors (Chandra, 2003). Another source of distraction is noted as the legibility of text and screens (Chandra, 2003). Furthermore, operators may not use a system based on automation it they believe it to be *untrustworthy*. Automation is often problematic because people fail to rely upon it appropriately. Because people respond to technology socially, trust influences reliance on automation (Lee & See, 2004). Parasuraman (1997) notes that, “a factor in the development of trust is automation reliability. Several studies have shown that operators’ use of automation reflects automation reliability” (p. 237). Therefore, “trust often determines automation usage. Operators may not use a reliable automated system if they believe it to be untrustworthy. Conversely, they may continue to rely on automation even when it malfunctions” (Parasuraman, 1997,). In addition to the problems posed by automations, the issue of efficiency of EFBs are also prominent. Pilot interaction with the EFB was significantly slower compared to paper according to one study (Hamblin, 2004). This is also the finding of another study, in addition the participants found the EFB data input to be more frustrating (Cahill, 2006). Flight safety Foundation (FSF) raise concerns about the increase in workload as a result of inefficient deigns of software and hardware, and a decrease in flexibility (Flight Safety Foundation, 2005). In addition to these, other approach seeks to explore the prevalence of incidents involving EFBs in various accident and incident data bases (e.g., Chandra et al., 2009; Chase et al., 2014; Tump et al., 2014). Findings here reveal that EFBs can be a source of in-flight distraction, hence emphasizing the need for EFBs to be integrated so that they correspond with the general flight deck design philosophy.

After a review of this literature, we deemed four main topics crucial to evaluate: access to information, distraction, situational awareness, and workload. *Access to information* relates to the promise of EFBs to facilitate better for information flow, as well as providing better access to the information the flight crew needs in order to perform their duty. *Distraction* highlights the requirement of EFBs to not divert attention from critical flight related duties. *Situational awareness* is crucial for pilots, and a loss of situational awareness is putatively regarded as one of the most frequent culprits of accidents and serious incidents. The potential for increased *workload* is frequently mentioned in the research literature and is a prime concern for regulators. These topics deduced from the literature was transcribed into sixteen (16) response variables, representing our proxies for flight safety.

3 Methods

3.1 Data sampling

3.1.1 Study design

As we wanted to explore relationships between different variables, this study was designed around statistical analysis of quantitative data. Subsequently, we use statistical methodology and test statistics as are our primary means of inference. The study is extensive as it is aimed at providing a general overview of how EFB-usage in Norway has affected flight safety (Johannessen et al., 2011).

Quantitative studies are most commonly used in the natural sciences, but it is also frequently used within social sciences. Although there are inter-disciplinary differences, these designs follow a more standardized format than other comparable approaches. Quantitative designs can be either descriptive or experimental (Johannessen et al., 2011). The former measures associations once, while the latter incorporates subsequent measurements for comparison in order to deduce causality (Johannessen et al., 2011). Here, we use a descriptive approach aimed at establishing associations.

In order to measure associations accurately, there is a need for large quantitates of data. What is "enough" is highly debatable (see Kotrlik et al., 2001 for an introduction). Realistically, it is impossible to sample the entire sample space. Hence, the concept of error-margins is introduced (Alwin, 2007). In essence, the error margin is a tolerable probability one accepts of obtaining the observed results by chance. Convention normally puts the error margin at 5%, categorizing results as "statistically significant" when the applicable test-statistics produce values which have a 1 in 20 chance of occurring by pure accident. We follow this convention with some adaptations to accommodate for discrepancies rendering us more prone to type-I errors.

Outliers are a problem in any meaningful inference (Hampel et al., 1986). There is controversy surrounding whether to treat outliers or not and there is no clear convention (Gosh et al., 2012). One must weigh the disadvantage of capturing important exception against the disadvantages of having a minority of samples affect the overall direction of the data. Ultimately, we decided upon correcting for outliers. Subsequently, an outlier detection was performed using Cook's Distance on pairwise squared mahalanobis distances (D^2) between participants (N = 571), calculated from all response variables (N = 16). Mahalanobis distances is a commonly used method for multivariate outlier detection (Aggarwal, 2015). Outliers above Cook's threshold was removed, resulting in missing data values. We also had missing data from incomplete sampling. We elected to imputate rather than remove these sample points from the analysis all together, as we were dependent on a large a sample size as possible to conduct a more meaningful downstream analysis. The disadvantage of having "normalized" the data would only serve to decrease our chances of rejecting the null hypothesis, hence making us less prone to type-I errors. We proceeded with the imputated using a k-Nearest Neighbours approach ((Fix and Hodges, 1951)). The choice of the k-Nearest approach was a mostly practical, as it eased the complexity of the software coding. The choice of the number of "neighbours" to look for when calculating the missing values is also a matter of debate (see Ghosh, 2006). Again, there is no clear convention. Setting to few neighbours will increase the influence of noise, rendering the results less generalizable. To many, and one could miss important local variations. We set k=20, more concerned with not having generalizable data than capturing local variations.

All statistical analysis was done using the open source statistical software platform *R* (Team, 2017).

3.1.2 Survey design

We conducted an electronic questionnaire type survey distributed to most major Norwegian operators (fig. S1). Our target population were professional pilots with careers in commercial aviation in Norway. These individuals operate multi-pilot high performance aircraft intended for passenger or cargo transport. We used non-probability sampling based on volunteers. The problem of self-selection bias was mitigated by correcting for pre-existing attitudes towards the topics included in the questionnaire.

A pilot survey was conducted prior to the main distribution. The questionnaire was given to a selection of 14 cohorts from our target population in order to provide feedback on the study design. We requested feedback on visual presentation, wording, spelling, format consistency, clarity, and relevance. In addition, we wanted to check for known biases, such as social desirability, order effects, and fatigue effects (Kalton et al., 1982). The former was mitigated by making the survey anonymous. Order and fatigue effects were not reported by the test population.

The distribution was done using an internet link from which the respondee had the opportunity to complete the survey once. The only user data collected was the local internet protocol address, in order to avoid multiple responses from the same device. This user data was not accessible to us, and only stored temporarily by the survey manufacturer. The questionnaire was open for approximately two months, from December 1st, 2018 until February 3rd, 2019.

The questionnaire was designed with objective response, close-ended questions, and divided into three main sections: (1) background, general (predictor variables); (2) background, PED familiarity (predictor variables); (3) participant EFB experience (response variables). Sections 1 and 2 was mainly composed of single answer multiple choice questions. Section 3 was a series of statements using a graphical rating scale ranging from 0 (complete disagreement), through 50 (neutral/indifferent), to 100 (complete agreement).

3.1.3 Question selection

A total of 16 statements where designed to address these areas and represents the response variables in the study design. These statements make up section 3 of the questionnaire and based on the flight safety literature outlined in the *Theoretical Framework* section of this study. Participants were asked to rate to what degree they disagreed or agreed with a series of statements, ranging from 0 (completed disagreement) to 100 (complete agreement). A graphical rating scale was used. The statements called for the participants to compare how they perceive flight safety *after* the implementation of EFB, to how they experienced the traditional flight deck.

A secondary objective was to evaluate if the perception of flight safety varied across all major Norwegian operators. Differences in practice, culture, and procedures are well known to accident investigators as being crucial for the safe conduct of flight. Any variation between different operators in how key aspects of the safe conduct of flight operation is performed since the implementation of EFBs, could be indicative of informal or formal structural dissimilarities organizational. Subsequently, we asked participants which operator they worked for.

A third objective was to see if we could identify any human factor predictors which could potentially help explain any variation in the response. From a managerial perspective, it is not only crucial to identify changes, but also understand what contributes to that change with a perspective of the individuals background, experience, and attitudes. To this effect, we asked the respondents about their general background and attitudes in section 1 of the questionnaire, and about their background and attitudes relating to PEDs in section 2.

As with the design of the statements in section 3, attempting to cover all possible predictors that could potentially help explain variation in the data is inherently difficult when dealing with human behaviour and human factors. Our goal here is to try and identify some human factors which we hypothesize could help explain some of the variation in how individuals perceive flight safety. Age and flying experience are putatively known to be relevant with regards to pilots and their attitudes towards flight safety (e.g., Li et al., 2003; Taylor et al., 2007; You et al., 2013). Flight deck role, whether they work as flight officers or commanders, was included because of the known difference in how they perceive or relate to flight safety. Commanders have the ultimate responsibility of the aircraft, while flight officers’ responsibilities in that regard are less stringent. This entails a difference in perspective with regards to many aspects and tasks of operating an aircraft. In recent years, atypical employment schemes and uncertain employer-employee relationships has emerged. Variation in how happy the individual pilot is with his working situation could affect how he copes with the implementation of new procedures such as EFBs, and subsequently affect how he perceives the change from the traditional flight deck. To this effect, we asked extent the participants approved of his current working situation.

As we are investigating the effects of an electronic device on flight safety, not only human factor predictors relating to the specifics of operating an aircraft should be included, but also questions that address how participants relate to such devices. Studies have shown that experience and familiarity with electronic devices affect both performance and attitudes towards them (e.g., Kang et al., 2008). Subsequently, we asked questions about user behaviour and whether touch screens where regarded as easy to work with in general. In addition, the FAA advisory circular emphasized the importance of sufficient training during EFB implementation, how automation could be a source of stress, the flight deck ergonomy with the EFB (i.e., the interaction between human and device), and type or category of EFB, as important aspects that should be evaluated (FAA, 2003). Lastly, we wanted to check if differences in general attitudes towards a more digitalized society could be contributing to how individuals responded to the implementation of EFB, which subsequently could affect how they regard the transition from the more traditional flight deck.

3.2 Data analysis

3.2.1 The effectiveness of EFB implementation in Norway

In aviation, effectiveness is ultimately measured in terms of flight safety gain or loss. Subsequently, we used the individual pilot experience as an evaluation procedure to measure the effectiveness of EFB implementation across all major Norwegian operators.

To evaluate whether EFB usage had any effect on the individual's perception of flight safety as opposed to a traditional flight deck, we wanted to compare the density distribution of the responses (H_1) to a normal distributed probability density function (H_0). Any significant deviation from the normal density function were interpreted as a flight safety impact. A negative flight safety impact would skew the response distribution towards the left of the mean. A positive flight safety impact would skew the response distribution towards the right of the mean.

As a first step we wanted to visualize graphically the responses and its potential deviation from the normal. We superimposed the density distribution of the responses on a normal density distribution using the package *ggplot2*. The parameters of the H_0 distribution (i.e., standard deviation and variance) was obtained by 1000 random samplings of the same number of integers as the number of survey responses (N = 572). The integer range was the same as the range of the graphical rating scale (0-100). This was done using the *sample* function in R (Team, 2017). In addition, we used the *qqPlot* function from the package *car* (Fox et al., 2012) in order to gauge the correlation of H_1 against a "best-fit" line representing a gaussian distribution (i.e., H_0).

While its useful to visualize the data, it is often not precise enough to draw more accurate conclusions. Skewness and kurtosis are used in statistics to derive numerical values or representations of data distribution. Skewness is a measure of asymmetry about distribution mean. A negative skewness is associated with a tail to the left of mean. A positive skewness is associated with a tail to the right of mean. For our study, a negative skewness entails a decline in flight safety. Conversely, a positive skewness would be indicative of an increase in flight safety. Kurtosis is a measure of the rate the tails reach zero. A kurtosis value of 3 represents any univariate normal distribution. Values less than 3 signifies a flat tail. A value above 3 is associated with tails that reach zero more slowly. A distribution with slow rate tails would translate into more outliers than expected from H_0. We calculated these parameters for each independent variable using the *moments* package (Komsta et al., 2015).

As a final step in the analysis, we investigated the probability of obtaining our results by chance. In order to determine statistical significance of any deviation from the H_0-distribution, we performed a two-sample Kolmogorov–Smirnov test (KS test) using the function *ks.test* (Team, 2017). The two-sample version KS test is suitable as it compares if two samples are drawn from the same probability distribution. We also performed a one-sample Shapiro-Wilks normality test (Team, 2017). The Shapiro-Wilks test is generally regarded as more powerful than the KS test, and was included to increase robustness. For both tests, the α-value was set to .05.

3.2.2 Organizational differences between operators

From an organizational standpoint, both informal and formal structures affect the effectiveness of the organization. When it comes to implementing new procedures in a high-risk environment, these structures play a crucial role in the success of this endeavour. A measure of how much the perception of flight safety varies between operators is a proxy for effectiveness in organizations with risk mitigation and flight safety as the desired outcome. Variations here might help reveal organizational differences between large Norwegian operators that could be investigated further in order to make similar implementations more successful and effective.

To investigate if pilots perceive flight safety differently depending on which operator they work for, we first performed a canonical variance analysis (CVA) using the R-package *Morpho* (Schlager, 2016). It is used to separate known groups in the data and provides an ordination that maximizes the separation of the group means relative to the variation within groups. A common challenge with CVA is that the number of variables is high relative to the sample size, thus restricting full-rank covariance matrix. Our group sizes were large enough to include all 16 statement response variables from section 3, however, as we had poor response frequency from some operators (i.e., $n < 20$), we elected to combine them into one group called "others" to avoid any singularities.

The CVA was run with a more conservative cross validation algorithm that mitigates overestimation of group separation (Mitteroecker & Bookstein, 2011). The CVA results were analysed graphically in two dimensions along the first two CVA-axes, representing most of the variation deduced from multivariate space. 95% confidence ellipsoids with was calculated in order to assess if the operators where significantly differentiated in multivariate space. In addition, we calculated the degree to which the CVA was able to categorize participants into the different operators. A large percentage of accurately categorized individuals means that the operators are highly differentiated. Finally, we assessed statistical significance of the CV-axes using Wilk's Lambda (λ). If the CV-axes are to similar, the subsequent results could be more unreliable.

To complement this approach, we extracted Mahalanobis distances (D) between all operators from the CVA results. In the multivariate case Mahalanobis distance is relative to the amount of variation in the direction of the difference. Subsequently, this distance measure is appropriate compared to other geometric distance measures, as it accounts for within group co-variance. Mahalanobis distances can be considered as statistical distance, i.e., it's a measure of the degree of statistical overlap between group samples. The greater the distance, the less likely it is that participants from different operators could be misidentified. For our case, greater distances translate to an operator where individuals are more likely to perceive flight safety differently. We ran a jackknife cross-validation procedure with 1000 iterations to investigate if groups were significantly differentiated.

CVAs are useful as a first step ordination method and to visualize multivariate space. The overlap, shape, size and placement of the confidence ellipsoids along the CV-axes gives an indication of which operators are likely to vary in how participants perceive flight safety to have changed since the implementation of EFBs. However, these methods are not as robust in evaluating which operators are statistically different from one another and hence significantly diverged. We therefor followed up with an *omnibus* stepwise analysis of variance (ANOVA) procedure using MANOVAs, ANOVAs, and finally a post-hoc was included using Tukey's honest significance test (Tukey, 1949) to find the difference in mean between all operator pairs. A logic where only significant MANOVAs are followed by ANOVAs with adjusted *p*-

values to the number of independent tests ("Holms approach"), reduces the risk of inflation error (Type-I) and considers inter-correlations between the different aspects of flight safety in explaining the total perception of flight safety (Holms, 1979).

A common difficulty in multivariate factorial designs is satisfying the presumptions of the statistical methodology employed to analyse the data. Classical multivariate analysis of variance (MANOVA) pose strict constraints that often lead researchers to use only univariate techniques with a subsequent loss of meaningful inference. More specifically, data which lack multivariate normality and equal covariance matrices are known to perform poorly in the sense that there is a greater risk of type I errors (i.e., rejecting a true null hypothesis). For univariate tests, there are also several assumptions that must be met before the *p*-values can be trusted (Anderson, 2001). Unbalanced designs and heteroscedasticity are a common real data features, as is the case with our study. Upon inspecting our data and using various test-statistics to check the assumptions (i.e., Box-M tests using the function *BoxM* (Genz et al., 2008) and Levene's test using *levenes.test* (Fox et al., 2012)), we discovered deviations that needed to be mitigated. MANOVA is robust against violations of homogeneity of variance-covariance matrices assumption if group sizes are sufficiently large (n>30). Our group sizes are larger. For the MANOVA we used the Pillai's trace test statistic (Λ) (i.e., an F-statistic approximation method), as this as shown to be robust against unbalanced designs and deviations from normality. For univariate tests we used $\alpha = 0.01$ to make sure we only rejected the null hypothesis if the tests showed very high significance levels. To alleviate heteroskedasticity, we used White-Huber covariance (a.k.a. "White" corrections) matrices available in the *car* package (Fox et al., 2012). Although these mitigating measures cannot guarantee erroneous conclusions, and non-parametric approaches exists, we elected to follow Allen & Bennett (2007) and proceed as parametric tests are more powerful and produce easier inference.

3.2.3 The importance of human factors

From a managerial perspective, it is not only important to evaluate from an organizational standpoint, but also understand how everyone contributes to the system. Particularly in aviation, where a lot of responsibility and trust is put on each individual, it is important to understand the human factors involved. Subsequently, individual differences in employee background, experience, or attitudes could contribute to how pilots perceive the EFB transition. We hypothesized that some of these human factors could help explain potential differences in answers between participants across all operators in Norway. This could provide management with useful insights as to which of these human factors aspects are important in explaining how flight safety perception has changed as a result of EFB implementation compared to the traditional flight deck.

To address these questions, we designed a model selection scheme using a combination of ordinary least squared regressions (OLS) and a downstream omnibus step wise analysis of variance (ANOVA) using various test statistics to determine significance. We considered reducing dimensionality in order to reduce the number of models. However, it was difficult to deduce what the different variance axes represents. In addition, the distance measure does not give us directionality, nor the opportunity to accurately assess the relative contribution of each predictor on the presence of others. Subsequently, we elected to analyse the raw data produced by the graphical scaling scores for each response variable (N =16). Our goal was to build one model for each response variable that also considers the multidimensionality (i.e., the correlation) of the data, subsequently analysing these to deduce which predictors are most important. If each response variable is not independent, but correlated with the others, regressing each response variable on a set of predictors will not account for this. A multivariate regression will be better to the extent that it can learn more from this correlation between these various measures of flight safety. As we have more than one predictor variable, the appropriate method is a multivariate multiple regression (Afifi et al., 2003).

For the linear regressions, we used quantile-quantile plots using the *qqplot* function in order to check for heteroskedasticity, outliers, and deviations from normality (Fox et al., 2012). For all 16 response variables, we found some minor deviations to normality and heteroskedasticity. An option is to fit the models using other distributions than gaussian. However, there is doubt to how much this matters if your deviations are relatively small (see Johnson, 2016). A distribution that better emulates our residual error distribution would yield better model fit, but following Johnson (2016), we chose to stick with the traditional linear gaussian regression for simplicity. For the multivariate and univariate analysis of variances as part of the down-stream inference, we used the same mitigating approaches as those described previously in this section.

After checking the various assumptions, we proceed with building a model space from which to draw the best model explain for the data. First, we checked for possible significant interactions between the different predictors that should be included in the downstream analysis. In order to determine which predictors to include in this interaction test, we only selected significant MANOVAs performed on each predictor independently using the *Manova* function the car package (Fox et al., 2012). For categorical predictors, the approximated F-statistics (i.e., Pillai's trace) tests within group variation to between group variation. For continuous predictors, the same statistics tests if the slope of the model is significantly greater than zero. We then proceeded by analysing all possible pairwise interactions of significant predictors. We used the *anova* function to conduct a likelihood ratio test on the "full" model (i.e., with the interaction) against the "reduced" model (i.e., without the interaction) (Team, 2017) . Significant tests ($p < .05$) were finally included in a maximized multivariate multiple regression model. As our data is not a subset of data we could have collected (e.g. there are no more airlines, age groups, or experience levels we could have sampled), we did not include any random variables, and all predictors and their interactions was included as fixed effects.

In order to find the best model from the model space provided by this maximized model, we continued with a step-wise regression (i.e. all factors and significant interactions included) to reduce it in a step-wise manner by removing interactions and or factors one at a time. Each step was ranked using the second-order Akaike's information criteria (AIC_c) to penalize over parameterization more severely (McQuarrie & Tsai, 1998).). Models with a deviance in AIC_c by less than two from the best-fitted model was regarded as comparable in explain the variation. If more models fell within this interval, we selected the most parsimonious model as the best model. AIC_c is one of two commonly used model selection methods, the other being the Bayesian information criteria (BIC). Other methods exist, but they are asymptotically equivalent to AIC and BIC. BIC and AIC have been shown to result in different model conclusions (for an overview, see Weakliem 2016, p. 29). We elected to use AIC, as BIC are more dependent on sample size to be accurate (Raffalovic, 2008).

The resulting model is the best multivariate model explaining variation in the perception of flight safety, while taking into consideration the possible correlations between response variables. The next step was to deduce the impact of each predictor has on flight safety. Specifically, this involved calculating the relative importance of each predictor as a function of how much they contribute to the variance explained by the entire model (i.e., model fit, R^2). In addition, the directionality of the predictors; if they contributed to a reduction or an increase in the perception of flight safety, needed to be deduced. This requires an inspection of the regression coefficients and standardizing them, so they become comparable between the different predictors. Subsequently, it became necessary to evaluate each univariate model making up the multivariate model space. As for the multivariate case, we conducted a stepwise AIC_c ranked regression of the multivariate model for each response variable. This produced a best model for the individual response variables, including all or some of the predictors included in the multivariate model.

In order to determine if a predictor had a positive or negative impact on the mean value of flight safety perception estimated for each univariate model (N =16), we calculated standardized coefficients using the *standardize* function in the package *arm* (Gelman & Hill, 2006). This is like the regression coefficient, beta (β). However, interpreting the comparison between binary categorical variables (i.e. flight deck position) with numerical variables in terms of standard deviations is problematic. To overcome this, the standardize function corrects numerical variables by two-times their standard deviation which is equivalent to going from one end to another. The relative contribution to explained variance was calculated using the *calc.relimp* function from the *relaimpo* package (Grömping, 2006). This calculates the relative importance of each predictor in the presence of other predictors using the R^2 contribution averaged over orderings among regressors, cf. e.g. Lindeman, Merenda and Gold 1980, p.119ff or Chevan and Sutherland (1991).

Summarized, the average predictor regression coefficients and relative importance measures represents the relative impact of these various measures of human factors on flight safety. The above described approach considers that each response variable is part of something larger, i.e. the correlation between these responses, thus producing a multivariate space that needs to be accounted for. By only including the predictors which turn up in the best multivariate model as a basis from which to deduce the best univariate models, we avoid the caveat of assuming that each response variable can be evaluated separately. Our measures of flight safety are not independent, but rather a collective whole that needs to be considered.

4 Results

4.1.1 The effectiveness of EFB implementation in Norway

Overall, participants agree more than disagree with the statements in section 3 of the questionnaire (fig. 7). Subsequently, pilots perceive a positive shift in flight safety since the implementation of EFBs to their flight deck. There are examples of certain areas where pilots perceive little change to the traditional flight deck (e.g., fig. 8: Head down time (6); Admin. duties (7); Communication (9)) , however there is only one instances where they on average respond negatively to the statements (i.e., fig. 8: Flight deck tasks (16)). The improvement to flight safety is largest, particularly in areas that concerns the ease of access to information (e.g., fig. 8: Access to OMs (14); Information access (1)) and the handling of abnormalities (fig. 8: In-flight abnormalities (11); Ground abnormalities (12)).

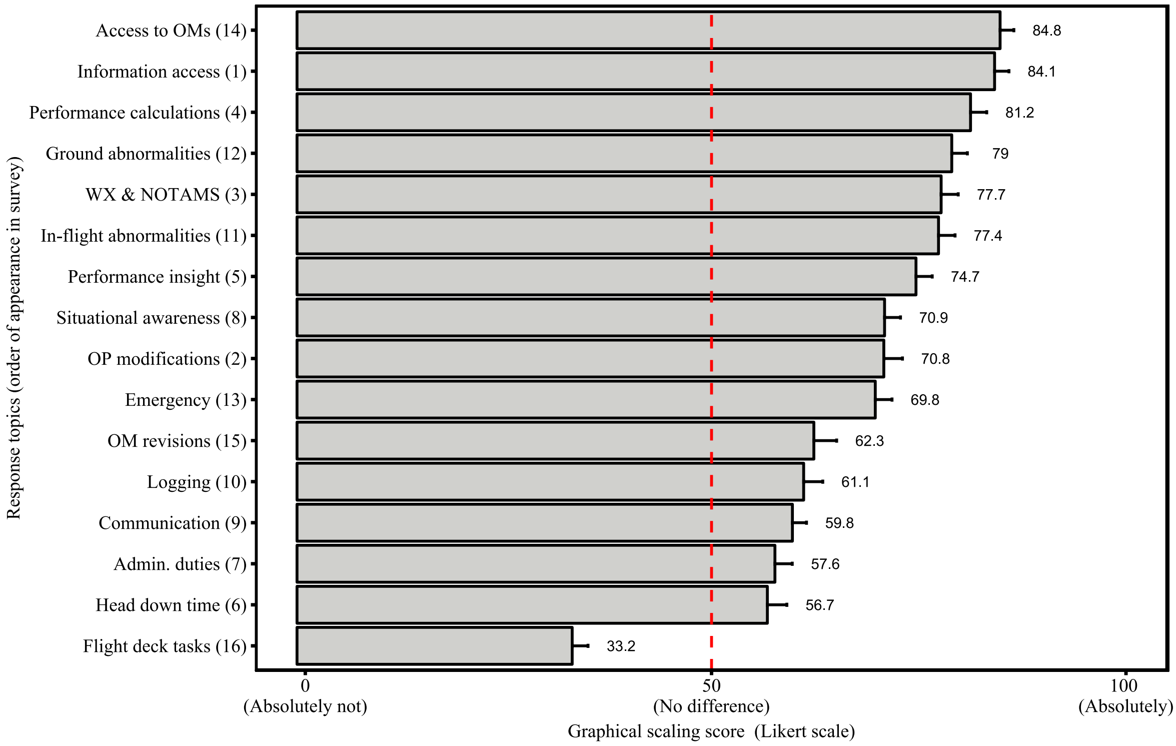


Figure 7. Mean graphical rating score of all response variable statements ($N = 16$), with 95% confidence intervals. The number at the end of each bar represents the mean score. Participants on average agree with the statements, and no statements have confidence levels that overlaps the red stapled line representing a state of no difference from the traditional flight deck.

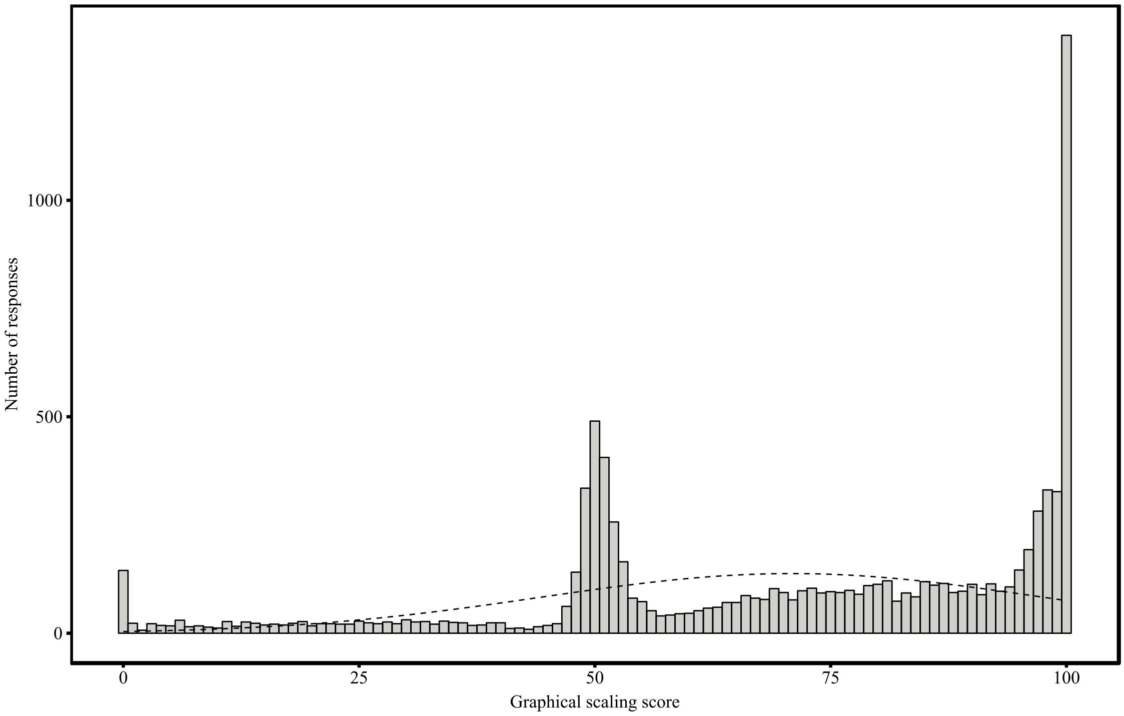


Figure 8. Frequency plot of all responses (N = 9 136) across all response variables (N = 16). The stapled black line represents a density distribution curve. The density distribution of the responses is non-gaussian and skewed to the right of mean. The pooled response from all 16 statements shows that participants perceive largely a positive shift in fight safety.

Looking at the frequency of pooled responses across the graphical rating scale, lends support to the claim that participants tend to agree with the statements (fig. 8). There are peaks around a scale value of 50, and towards the higher end of the scale. This indicates that individuals perceive flight safety to be improved, but also there is a significant proportion of participants who regard little or no change in flight safety for some response statements. The density distribution for the pooled responses is non-normal as it shows a skewness to the right of mean (fig. 8; table S1: skewness = 2,80), with a heavier tail (fig. 8; table S1: kurtosis = -0,74). The KS-test is also highly significant (table S1: $p > 0{,}001$). This result is mirrored by the quantile-quantile plot (fig. 9). The sample quantiles do not follow the gaussian quantiles to any significant degree. Specifically, there are much more responses than expected, particularly to the right of mean. This supports the claim that participants tend to agree more with the statements than expected if there where fewer differences between the traditional flight deck and the EFB flight deck.

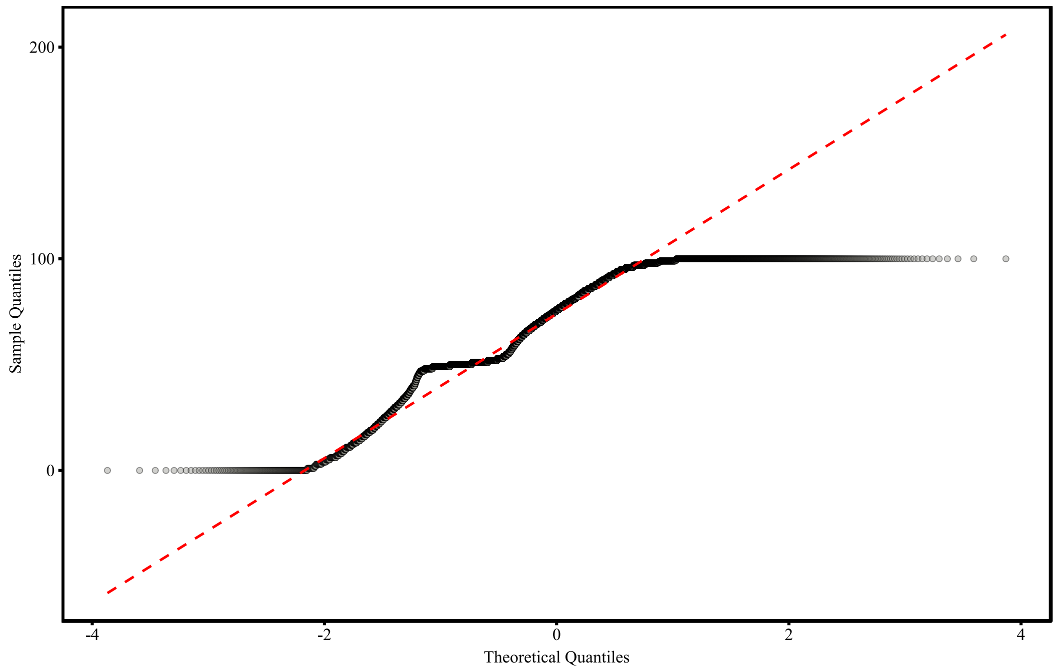


Figure 9. Quantile-quantile plot of all responses (N = 9 136). The red stapled line representing a theoretical expectancy following normal distribution. The grey filled dots are the individual responses. The response quantiles do not follow a normal distribution, with more participants than expected at extreme quantiles, particularly to the right of mean.

The results for each of the 16 statements response variables largely mimic the pooled results (fig. 10; table S1). Visually, there are clear deviations from normal with density probability distributions with skewness to the right of mean. Participants tend to answer with higher graphical rating scores than expected from a normal distribution. From visual inspection participants tend to perceive the least changes from the traditional flight deck for 'head down time' and 'in-flight logging'. Looking at the statistics for each response statement, all responses show significant deviations from normal with both the K-S tests and Shapiro Wilks tests *p*-values well below the error margin of .05 (table S1). All responses except 'Flight deck tasks' have a mean significantly in the "agreement" region of the graphical scale (table S1: $\overline{x}$ = 33.19). 'Flight deck tasks also has a positive skewness', indicative of a more than expected response distribution to the left of mean (table S1: Flight deck tasks (16), Skewness = 0,39). This is also true for 'Communication'; however, this suggests a very close to the normal expectation of skewness (table S1: Communication, Skewness = 0.04). It appears that pilots tend to perceive an improvement compared to the traditional flight deck, but an increase in the tasks demanded of them.

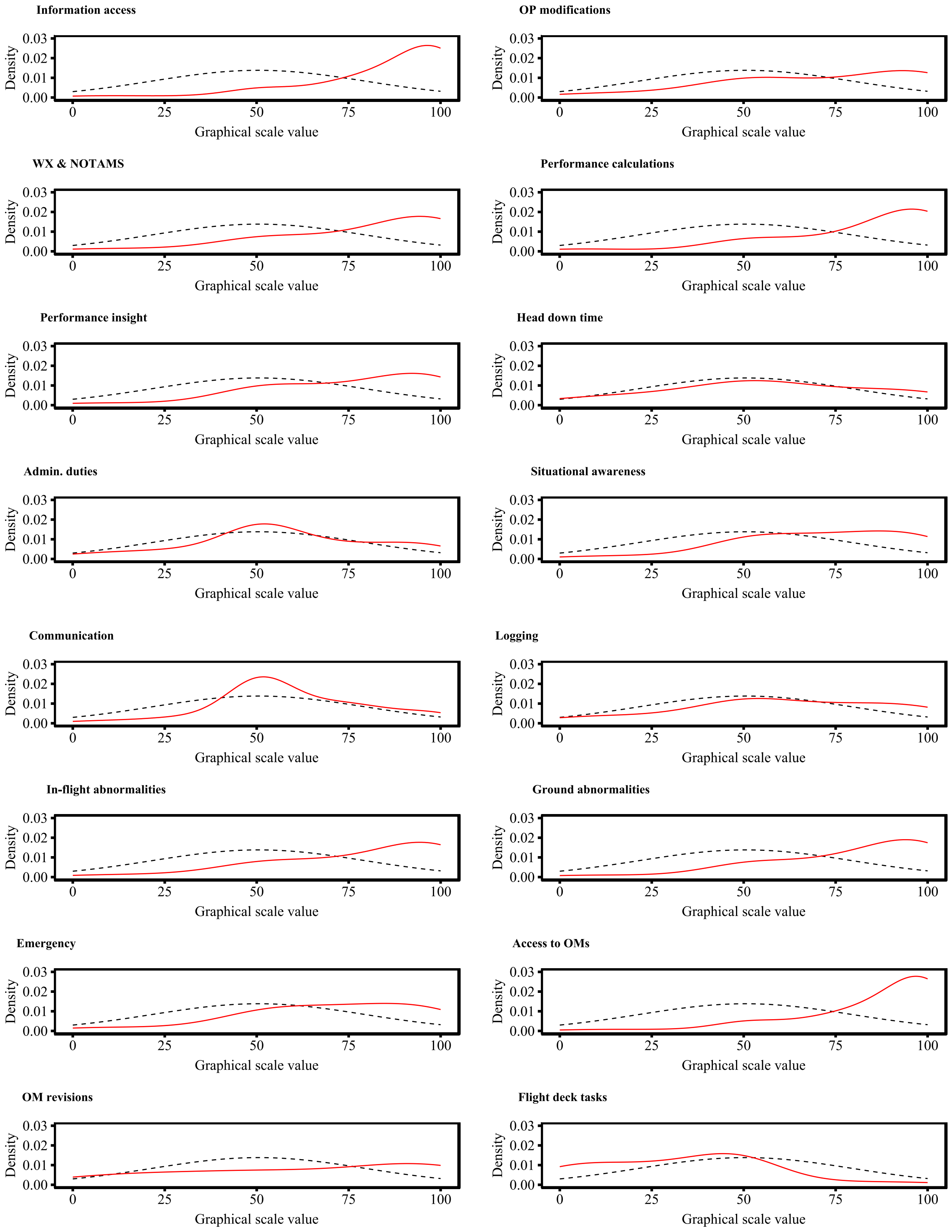


Figure 10. The probability density distribution of all 16 response variables (red line) plotted against a hypothetical normally distributed density distribution. For all statement responses except *Flight deck tasks*, the density distributions are skewed to the right of mean, with more than expected agreement from the participants.

4.1.2 Organizational differences between operators

Overall, pilots perceive flight safety differently depending on which operator they work for. This holds true for both results derived from dimensional decomposition approaches and variance analysis on the “raw” data. The extent of this variation can be deduced by studying the degree of overlap, separation, and shape of the confidence ellipsoids derived from multivariate space (fig. 11), statistical distance (fig. 12; table S2), and the amount of variation within different operators compared to the amount of variation between operators (table S3; table S4), and the difference in mean value between operator pairs for each statement response (fig. 13; table S5).

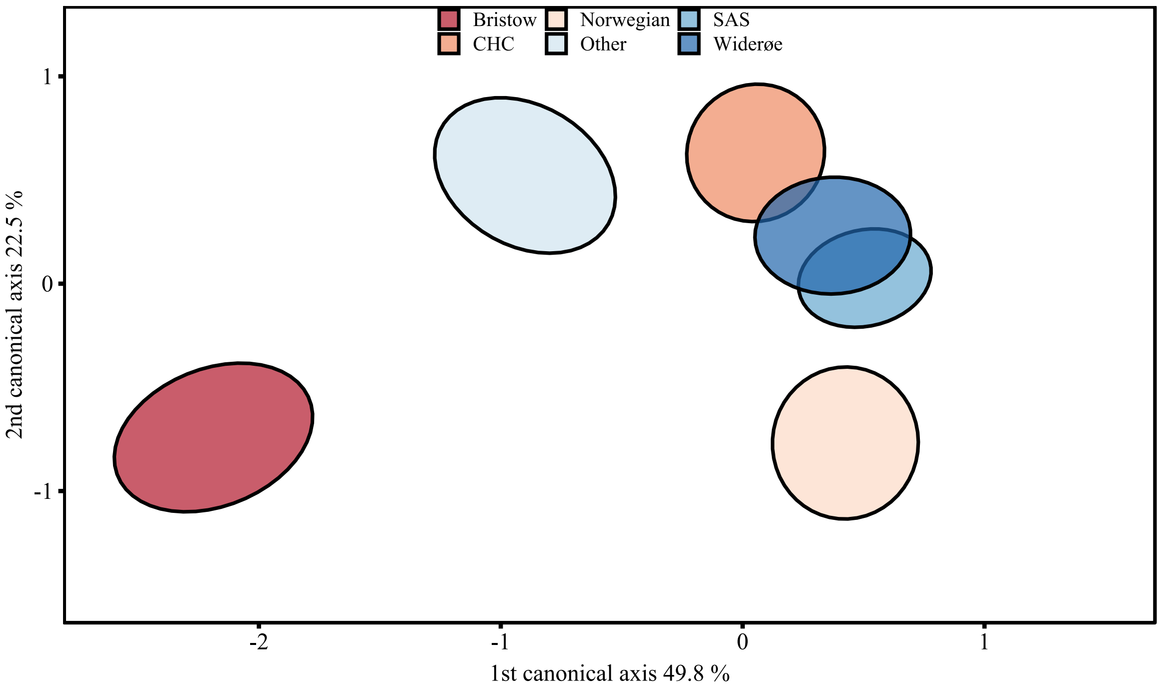


Figure 11. Canonical variates analysis for all 16 response variables grouped by the various operators. The colored circles represent 95% confidence ellipsoids. In multivariate space, there are regular occurrences of non-overlapping ellipsoids, indicating that participants perceive flight safety to have changed differently depending on which operator they work for.

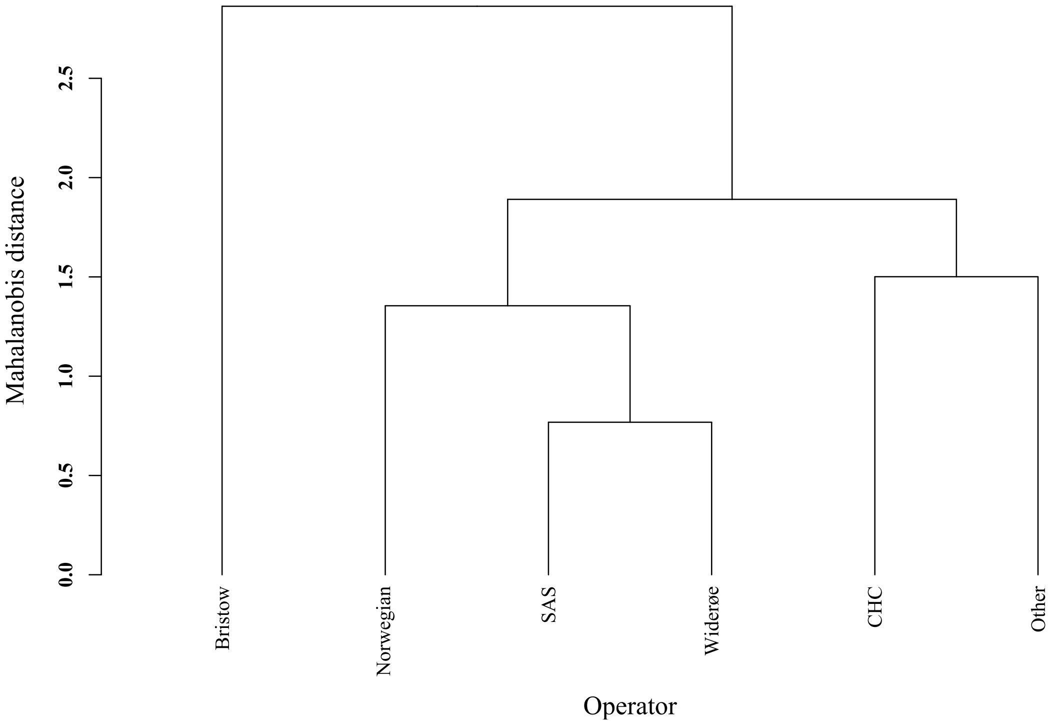


Figure 12. Bifurcating dendrogram based on mahalanobis distances derived from the CVA analysis. Each terminal branch represents a different operator. Nodes are the distance for which the different groups split. There are four families: Bristow; Norwegian; SAS, Widerøe; CHC, Other.

Bristow

Participants working for the operator Bristow regard the transition from the traditional flight deck to the EFB flight deck more differently compared to participants working from other operators. In multivariate space, Bristow is distinctly separated from the other operators (fig. 12). The distance in multivariate space is also the largest (table S2: $D^2 = 2.14$ to 2.86; fig. 12). This operator has the second highest frequency of significant different mean scores between operator pairs (table S5: $n = 23$) and exhibit the most instances of larger absolute differences (fig. 13).

Norwegian, SAS, Widerøe

Pilots working for these three operators perceive the EFB transition relatively similarly compared to other groups of operators. In multivariate space, these operators are clustered close together, particularly SAS and Widerøe (fig. 11; fig. 12). Looking at the analysis of variance, these operator pairs occur less frequently in conjunction than the others (table S5: *n*

= 5). They also have relatively low significant differences in mean for these occurrences (fig. 13).

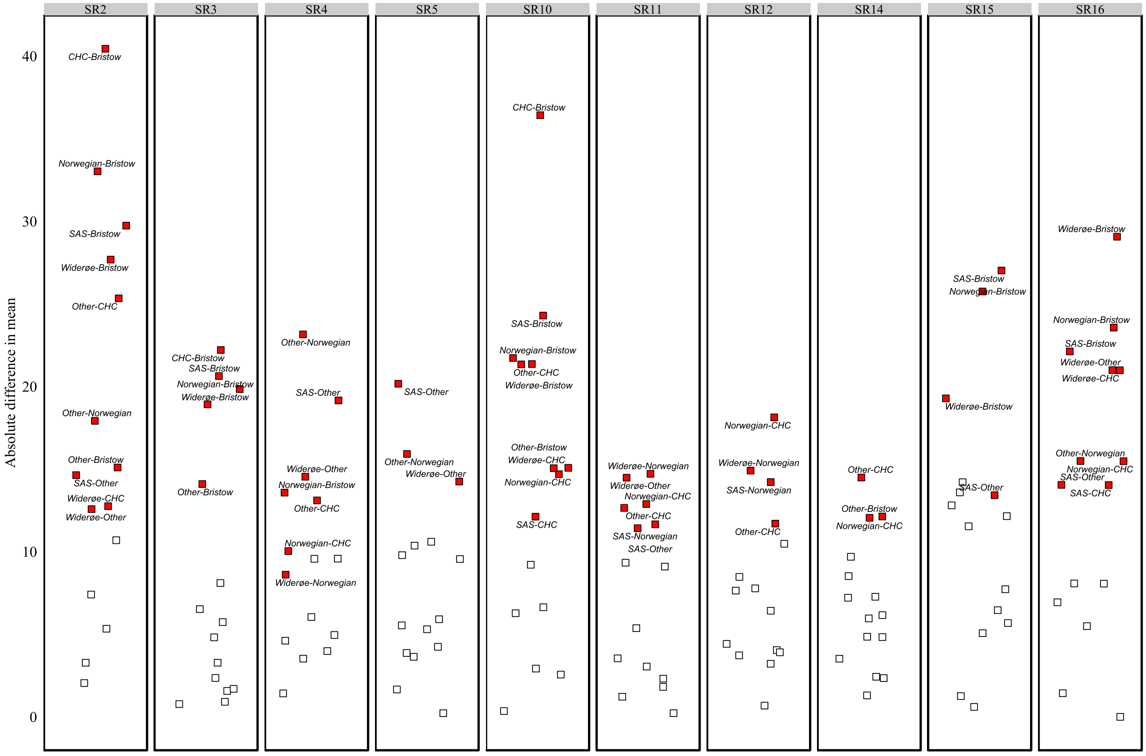


Figure 13. Result of the stepwise analysis of variance. Plot shows the frequency of occurrence and the absolute difference in mean between all significant population pairs (red squares) and insignificant pairs (white squares), sorted by significant statement response (SR) ANOVAs (i.e., 10 out of 16 SRs where significant).

CHC

In multivariate space, the operator CHC is closest to Widerøe and SAS (fig. 11). This is confirmed when looking at statistical distances (table S2: D^2(Widerøe) = 1.46, D^2(SAS) = 1.34). However, the dendrogram clustering algorithm groups it closest to Other (fig. 12). Looking at the variance analysis, CHC has 20 instances where it is has a mean value significantly different from other operators (fig. 13).

Other

This category includes participants from various operators (NLA, Thomas Cook, other minor operators). Subsequently, one expects larger within group variation compared to other operators. Looking at the confidence ellipsoids (fig. 11), confirms this as it is larger than most others. After Bristow, it is also the group that is at the greatest distance to the other operator groups (fig. 12; table S2: $D^2 = 1.50$ to 1.89). This category has the highest frequency of different mean scores between operator pairs (table S5: $n = 25$).

4.1.3 The importance of human factors

Not all human factors were found to have a significant impact on the variation in the data (table S6; table S7; table S8), but pilots background, experience, and attitudes appears to varying degrees to account for variation in the perception of flight safety. However, organizational differences have the most impact as to how pilots experience the EFB flight deck compared to the traditional flight deck (fig. 14). It has the highest frequency being present in almost all models (table S11; *n*(operator) = 15) and has the highest relative explanatory power (fig. 14: R^2 (operator) ~ 0.061). Of the human factor variables, the ergonomy of the EFB on the flight deck has the highest frequency of occurrence (models (table S11; *n*(EFB ergonomy) = 13) as well as the highest explanatory power (fig. 14: R^2 (Ergonomy) ~ 0.047)).

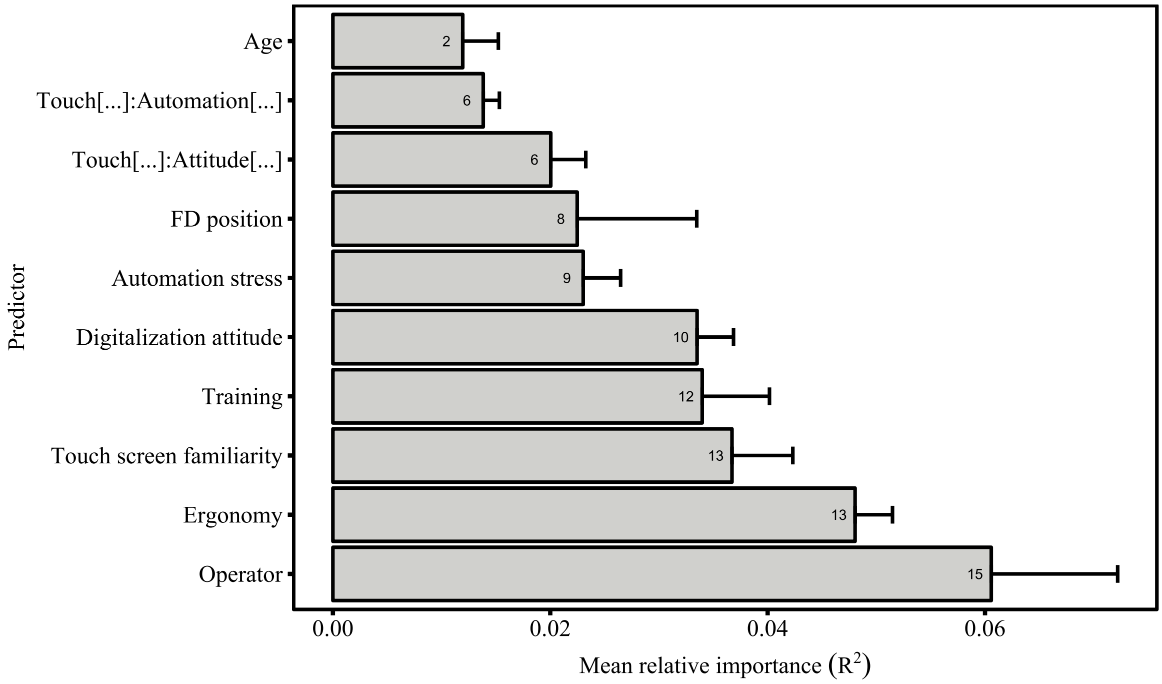


Figure 14. The mean relative importance of all predictor variables included in the best univariate models (*N* = 16). The error bars represent the standard deviation across all models, and the number inside the bars are the frequency of occurrence of each predictor across all models. Both significant and insignificant predictors are included. Operator is the predictor that is present in most model, as well as having the highest explanatory

Looking at the directionality of the impact of the human factors, most of them have a positive effect on the mean (fig. 16: green error bars). An almost equal proportion has a more varying impact depending on what aspect of flight safety that is being measured (fig. 15: black error bars). However, the mean impact across the graphical scaling score is in the positive (fig. 15: grey filled dots). The only two human factor aspects that consistently cause a decrease across the graphical scaling score mean, are the interaction between a participant's positive attitudes towards digitalization and heirs familiarity with touch screens (see fig. S4 for an example), and the interrelation between automation stress and the familiarity with touch screens (see fig. S5 for an example).

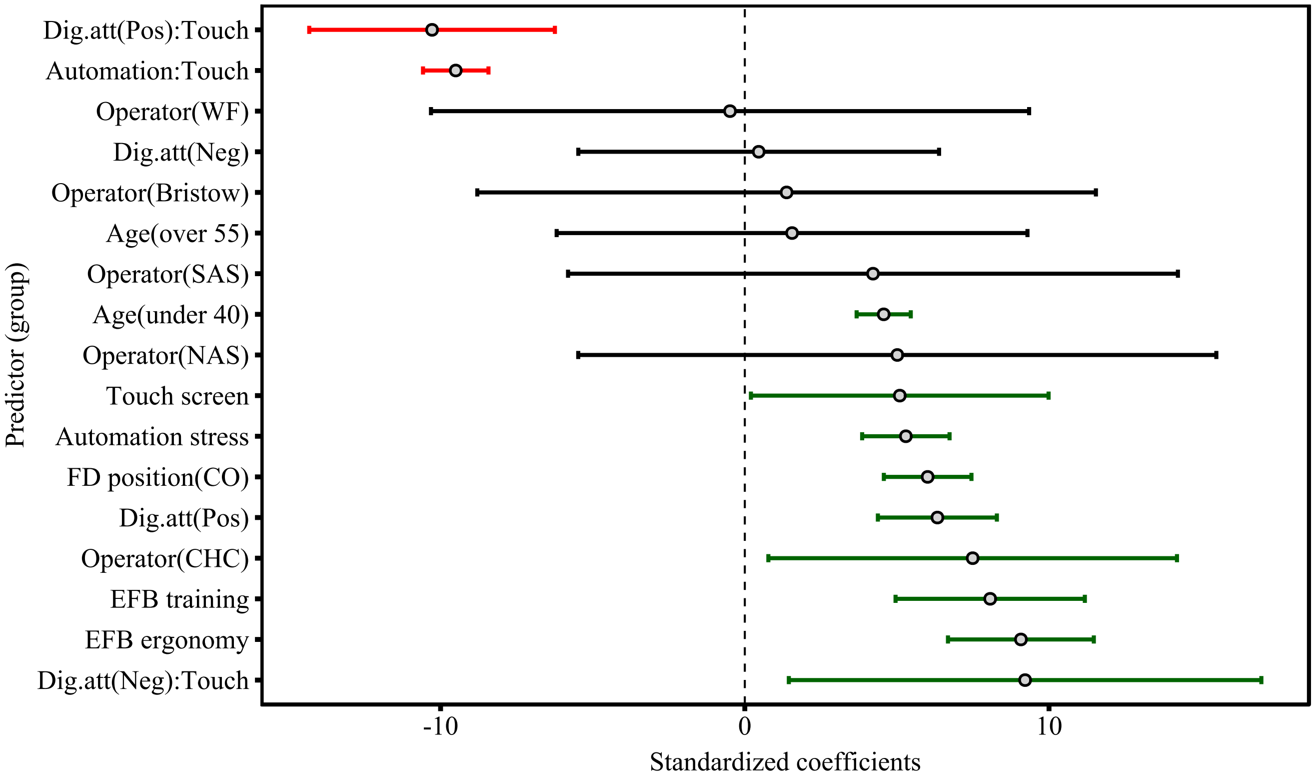


Figure 15. Mean standardized regression coefficients (β) for each predictor and its respective groups, across all univariate models (N = 16). The reference level is: Age; "40-55 YOA", Operator; "Other", Digitalization attitude; "Indifferent", Flight deck position; "FO". The error bars represent the standard deviation and is colored for their directional impact on the mean graphical scaling score across all models (red = always (consistently) negative, black = varying directionality, green = always (consistently) positive. As expected, most predictors and their groups have a positive impact.

power.

Age

The impact of a pilots age on the perception of flight safety is the smallest compared to other human factors (fig. 14). It appears in the fewest models explaining variation (table S11: *n* (Age) = 2), and the lowest average relative importance (table S13: mean R^2(Age) = 0.015). The lowest age category (under 40 YOA) has the most positive impact on the mean (fig. 15). If considering only statistically significant data, age only has a reliable effect on how pilots perceive the change in the amount of flight deck tasks (table S12).

The interaction between touch screen familiarity and one's tendency to become stressed by automation

This relationship/interaction has the second lowest impact (fig. 14). However, it appears in a fair proportion of models (table S11: *n* (Touch[…] : Automation[…]) = 6). The mean relative importance is very low despite (table S13: mean R^2(Touch[…] : Automation[…]) = 0.014). A close look at the interaction reveals that if one become stressed by automation, one's perception if the EFB flight deck becomes more negative despite one being progressively more familiarity with touch screens (fig. S5). The interaction has a reliable affect as it is statistically significant across in all models it appears in (table S12).

The interaction between touch screen familiarity and one's attitudes towards digitalization

This interaction as a relatively low impact (fig. 14), and a low relative appearance in the best models (table S11: *n* (Digitalization[…] : Touch[…]) = 8). The mean importance across those models is among the lowest (table S13: mean R^2(Digitalization[…] : Touch[…]) = 0.020) Interestingly there seems to be an inverse relationship between a positive attitude towards digitalization and its interaction with touch screen familiarity at its consistently decreasing the mean along the graphical scaling score (fig. 15). However, a closer look at a plot of the interaction shows that those with a positive attitude do not change how they perceive the access of information to have changed since the implementation of EFBs even as their familiarity with touch screens vary (fig. S4). If considering only reliable effects of this interaction, the positive digitalization attitude is significantly different from the reference level for statement responses 1, 3, and 13 (table S12). The negative attitude is significantly different for statement responses 12 and 14 (table S12).

Flight deck position

The role the participants have on the flight deck has an effect in exactly half the models (table S11: *n* (FD position) = 8). Its relative importance is relatively low (table S13: mean R^2(FD position) = 0.022), but it varies relatively greatly across the models ($R^2(\sigma)$ = 0.011). If you are first officer, you perceive a greater improvement from the traditional flight deck compared to commanders and this is impact is consistently positive (fig. 5). This predictor is reliable across all model it appears in (table S12).

Automation stress

The main effect of automation stress appears in over half the models (table S11: *n* (Automation stress) = 9), with almost the same relative importance as flight deck position (table S13: mean R^2(Automation stress) = 0.023). However, this importance is more consistent across the model it appears in ($R^2(\sigma)$ = 0.003). No matter how stressed you are by automation, you do on average consistently perceive an improvement from the traditional flight deck (fig. 15). This effect is reliable across all the model's automation stress appears (table S12).

Attitudes towards a more digitalized society

The main effect of attitudes towards digitalization appears in a fair number of models (table S11: *n* (Digitalization attitude) = 10). However, its groups are only significantly different from one another in 3 models (table S11: statement responses 1, 2, and 12). Its average importance is fair jump up from the until here mentioned human factors (table S13: mean R^2(Digitalization attitude) = 0.034). If you are positive towards a digitalized society, then more you consistently perceive an improvement to the traditional flight deck (fig. 15). There only one statistical difference between negative and indifferent attitudes (table S12).

EFB training

The degree of training participants perceives to have received has an impact in 12 of the models (table S11: *n* (EFB training) = 12). However, its only present in 4 models if one imposes a strict error tolerance (table S11). Its mean relative effect on the explained variation is almost the same as 'attitudes towards digitalization' (table S13: mean R^2(EFB training) =

0.034). Across all models, this predictor consistently contributes positively to the mean, no matter how the perception of the degree of training one has received changes (fig. 15)

Touch screen familiarity

The main effect of how familiar participants are with touch screens appears in a high number of models ((table S11: *n* (Touch screen familiarity) = 13)). However, imposing a strict error tolerance reduces that number to 1 (table S11: ‘Situational awareness (8)’). Its average importance including non-significant appearances is among the highest (table S13: mean R^2(Touch screen familiarity) = 0.037). No matter how familiar participants are with touch screens, they perceive an improvement to the traditional flight deck (fig. 15).

EFB ergonomy

As previously mentioned, how the ergonomy of the EFB has the highest relative importance of all human factors (table S13: mean R^2(EFB ergonomy) = 0.048). As is the case with other predictors, employing a strict error tolerance renders the number of appearing models to only 3 (table S11: ‘Operative flight plan (2)’, ‘WX & NOTAMs(3)’, ‘Situational awareness (8)’). The directional impact on the mean average graphical scaling score is consistently positive (fig. 15).

Operator

Differences between operators, our approximation of organizational differences, was included to gauge the importance of human factors, compared to another non-related variable. Variation between operators has the highest mean relative contribution to the explain variation across all models (table S13: mean R^2(Operator) = 0.061). It varies relatively much between models as well ($R^2(\sigma) = 0.012$)), contributing at its highest with 18% of the explained variation in the data (table S13: R^2(Flight deck workload (16)) = 0.18). In addition, even when applying a strict error margin, its appearance in the best models only decrease by one (table S11). If one takes the group ‘others’ as the reference level, participants working for CHC are the only ones who consistently perceive an improvement to compared to the traditional flight deck (fig. 15).

5 Discussion

We found that pilots readily perceive their EFB flight deck an improvement compared to the traditional flight deck. It is hence likely that flight safety has been strengthened with regards to the different aspects pursued in this study. Subsequently, the implementation of EFBs by Norwegian operators has been effective. Furthermore, we wound evidence of reliable differences between the various operators as to how pilots regard the transition. This suggest that there are structural organizational differences which has had an impact on the effectiveness, in terms of flight safety, of the EFB transition. In addition to the organizational aspects, we found that pilots background, experience, and attitudes has an impact on how they (the pilots) relate to the statements in this study. These human factors influence how pilots perceive flight safety to have changed since the EFB implementation, making it likely that these factors play a part in how change is implemented in the organization.

We found that there has been change across all the flight safety proxies included in this study. More often, this change was a positive one. Pilots seem to regard the EFB as a useful tool that aids them in their daily safe conduct of flight operations. This mirrors other studies which has looked at the Pilot-EFB interaction. For instance, Haddock & Beckman (2015) found similar support when he asked a group of pilots on how the EFB performed. Lytle (2015) found that pilots perceive the EFB positively, with only a small percentage of participants in disfavour of the device over a number of survey questions. However, most inquiries into the safety impact of EFBs has looked at reporting patterns in the safety data bases (e.g., Chandra et al., 2009; Chase et al., 2014; Tump et al., 2014). The EFB do appear in the incident statistics to an increasing degree (Tump et al., 2014). However, as Chase pointed out, it must take into consideration the increased popularity of these devices (Chase et al., 2014). In addition, it is often difficult to distinguish if the EFB is a causal factor in promoting high work load situations, or if these situations would arise regardless. Our findings seem to suggest that EFBs provides an improvement to typical high work load scenarios. This is achieved mainly through facilitating for more prompt access to documentation or information about their flight. The EFB also performs well in such high work load environments as when preparing for departure, as pilots is finding it easier to make the necessary calculations of flight parameters and deal with any abnormalities.

The EFB is meant to make the operation more streamlined and efficient. Critical in this, is the flow of information and facilitation of access to this information. Judging by the pilot's own experience, this is mainly what the EFB achieves. Subsequently, a desired outcome of less frequent safety breaches is a likely result. There is a vast amount of documentation and information that needs to be reviewed before every flight. In addition, it is critical that this information is up to date and reliable because of the many revisions to procedures and regulations that haunts the airline industry. If not, the margins of error decrease because the flight might be calculated using the wrong parameters, dispatches with out of date performance data bases, erroneous fuel consumption estimates, and a plethora of other potential hazards impeding on the robustness of the operation. From a managerial perspective, the transition from the traditional flight deck to an EFB solution has improved the flow of information and provided a reliable and useful platform as which to distribute this information. Subsequently, there is evidence suggesting that organizational effectiveness have improved as a result. Pilots themselves find it easier to process the stream of information, and this would free up attention that could be divide to more direct flight related tasks.

Impeding on this apparent gain in safety margins and subsequent increase in organizational effectiveness, is the danger of task saturation (Endsley, 1999). Task saturation arises when the number of tasks required to be performed, outweighs the time or recourses required to solve them. Task saturation is a well-known phenomenon in aviation, and pilots are trained to plan ahead and anticipate upcoming tasks that needs to be taken care of. Pilots are required to perform numerous tasks both before, during, and after flights. In addition, pilots must undergo numerous training curriculums to keep current, keep up to date with revisions, and prepare themselves for flight reviews etc. Regulatory authorities impose a steady stream of new directives and requirements that needs to be absorbed by the operator. The manner in which all these aspects are handled from an organizational standpoint is crucial to the effectiveness of that organization in terms of safety (e.g. Ates, 2017). If management is not careful, pilots, the end producers of flight safety, might get swamped with information and tasks they need to process. The EFB is partly meant to alleviate or mitigate this by making the information more readily available (Flight Safety Foundation, 2005). However, with increasing effectiveness comes the potential for diverting more tasks to that area of the organization. In addition, recent years development of PEDs and the improvement in human interface and computing power, has granted new possibilities with regards to handling calculations that were previously done by auxiliary support personnel.

Our data indicate that despite the increase in effectiveness in terms of information access and flow, pilots experience an increase in the number of tasks required of them, compared to the pre-EFB era. In addition, the amount of time spent with their “heads down” in the cockpit, time spent on administrative duties, and the pattern of maintaining a closed communication loop, are not perceived to be that different from the traditional flight deck, albeit in a positive direction.

Accident statistics is full of example of the dangers of getting preoccupied with other duties besides piloting. In a classical accident scenario, Pilots are distracted by other administrative or auxiliary tasks at the wrong time, contributing to a loss of situational awareness once a threat occurs, startling them, causing an inadequate response to resolve the situation. In addition, the threat of task saturation means that the pilots have a lesser chance of dealing with a situation successfully (Martin et al., 2015). Being outside the “loop” of communication, spending too much time with your “head down” (i.e., not directly controlling the aircraft), and fulfilling administrative duties are key areas of concern in aviation safety (Chase, 2014). Our findings seem to suggest that these are areas where the EFB is not performing as well compared to other areas of use. However, we cannot say if there is a causal link between EFB usage and a less favourable result. The increased task demand for instance, might be due to other factors which are not the EFB itself and would have arisen regardless. Future studies should try and resolve whether the EFB is contributing negatively to these areas of concern by being such a convenient and powerful platform for many of the tasks required on board an aircraft, that managers or regulators overexploit the capacity of pilots by further burdening of their tasks.

If one is to measure the effectiveness of the organization, this can be done by looking at whether an organization achieves its goals, and by using its resources to do so. An organization must be clear about what tasks or functions it should fulfil. When constructed, it will depend on plans for what to do, and how to achieve it. The organization then sets goals and designs strategies (Jacobsen and Thorsvik, 2013). By setting goals, the organization could have a motivating effect on the employees (Stajkovic, 2001), have a governing function by guidelines and limitations (Kaplan & Norton, 2000), provide support and legitimize the organization with regards to the environment (Powell & DiMaggio, 1991), and serve as evaluation criteria for the work performed by the organization and the individual (Scott, 1987). A typical airline organization resembles a mechanized structure, with strong independent subunits of heretical relationships (Jacobsen and Thorsvik, 2013). Typically, there is an operative branch dealing with flight operations. They are concerned with the direct task of ensuring flight safety, and effectiveness is directly measured by the number of safety events (Thomas, 2012). Hence, in terms of the effectiveness of the airline safety organization, the EFB does provide an improvement according to the pilot themselves. As in many other institutions, the EFB is an example of how information technology increase effectiveness and efficiency. This digital revolution has made many of the classical ideas of organizational structure, strategy, tasks, etc. obsolete (Jacobsen and Thorsvik, 2013). Although the debate is still ongoing as to exactly what effect this digital revolution has had on organizational structure, at least for the flight deck, the implementation of digital platforms appears to sit well with the people operating aircraft on board Norwegian air carriers. Subsequently, the Norwegian aviation sector could arguably be recognized as a sector where organizational effectiveness has increased as a result of a digital transition.

Having established that the implementation of EFBs provide favourable gains to organization effectiveness, we now turn to look at whether there are differences between Norwegian operators. We find that pilots experience the EFB-flight deck differently depending on which carrier they work for. This finding suggests that Norwegian operators has had varied success with the digital flight bag transition. There can be many causal origins nested in organizational differences between operators, and we do not aim to untangle these relationships as it would require in depth knowledge of each operator, which is beyond the scope of this study. What is plausible however, is that the implementation of a critical new safety feature like an EFB poses formidable challenges to any safety organization that would require a successful change management strategy. According to Jacobsen & Thorsvik (2013), changes within an organization can take place in the formal and/or informal elements. The formal elements consist of technology, goals, strategy and structure. These formal elements are the part of the organization that can be easily written down. The organization's informal elements contain organizational culture and power relations. These elements are not written down and adopted in writing but have grown in the organization over time. For instance, Borgen (2018) evaluated EFB-implementation in Widerøe within the theoretical framework of change management (Todnem, 2005). He found that a combination both informal and formal elements, e.g. the use of correct change management strategy, leadership trust, and employee involvement etc., where crucial factors in the success of the transition. In addition, organizational structure was deemed inadequately adapted to the intended change. Hence, it is likely that some of these aspects also explain some of the variation between the different operators.

Another probable aspect is revealed by accident reports and flight safety research involves organizational culture which in turn is affecting behaviour (ICAO, 2012). In the broadest sense, it is argued that a strong organizing culture will gather employees in a common experience of belonging and community (Jacobsen & Thorsvik, 2013). Organizational culture can have a strong impact on members of the organization, such as the formal structure, and appear to have definable effects on behaviour, although its often ambiguous as to how to measure it reliably (Flannery, 2001).With common goals and values in the culture, it will create motivation to perform, and one can get the feeling of working for a larger whole, rather than alone, which has also proven to be a strong effect on motivation (Latham & Pinder 2005). The effects of organizational culture on behaviour was long overlooked in aviation (ICAO, 2012). The causal factors leading up to an event has undergone an evolution from looking at purely technical aspects, to the human factors, all the way to including organizational or systemic factors. During this period, one began to look at the safety work in a system perspective, which included organizational factors as well as human benefits and technical factors. In addition to traditional data collection and analysis, limited to data collected by accident and serious incidents, the safety work was extended with a new proactive approach. Proactively collecting and analysing data was initiated routinely, and not as previously only reactive to monitor known safety risks and to detect emerging safety issues. This improvement provided the basis for moving towards a "safety management" or safety management approach (the International Civil Aviation Organization ICAO, 2012).

In aviation, the concept of *Just or Blame culture* are often often used. Reason (1998), describes Just Culture as an atmosphere of thrust in which people are encouraged, even rewarded, for providing essential safety-related information. But in which they are also clear about where the line must be drawn between acceptable and unacceptable behaviour. A ”no-blame” culture is neither feasible nor desirable. In a Just Culture environment, the culpability line is more clearly drawn. Benefits from having a Just Culture versus a blaming culture would be increased safety reporting, thrust building, and a more effective safety and operational management. A Just Culture support learning from unsafe acts in order to improve the level of safety awareness through the improved recognition of safety situations and helps to develop conscious articulation and sharing of safety information.

In Norway, the professional culture among pilots can be said to be fairly uniform. However, this sort of culture requires time to build up as it is passed from senior to junior members of the pilot corps. Most operators have also a strict no-blame policy, which seemingly give pilots a *carte blanche* authority in matters of flight safety (ECA, 2016). However, recent years have seen a shift in the industry with more creative employment schemes and more unsecure terms of employment also with Norwegian operators, making the lines more ambiguous as to what is acceptable or not in terms of pilot autonomy (ECA, 2019). The consequences of this on the pilot mindset is still being investigated (Jorens et al., 2015), but it is likely that this also has an impact on the informal levels of an organization. Hence, the safety culture within each operator might be thought of as a partly a product of that operator's business model. These formal organizational elements might impede on good informal elements (Jacobsen & Thorsvik, 2013). For instance, if there is rapid changeover of personnel with little time for vertical transfer of attitudes, no strong sense of belonging due to e.g., contract-based employment schemes, or a general misbalance between work and spare time, this can lead to a breakdown of good safety culture (Jorens et al., 2015).

There is some evidence in our data that differences in organizational elements causes pilots to relate to the EFB-transition differently. The operators SAS and Widerøe are the two operators which are least distinguishable in how they perceive flight safety to have changed compared to the traditional flight deck. Interestingly, they share a long and common history. SAS used to be the main shareholder of Widerøe. SAS-management philosophies were likely incorporated into the Widerøe organizational structure. Widerøe has operated under SAS call-signs and has historically been feeding to SAS-hubs. Widerøe and SAS crews have shared break rooms, pre-flight briefing rooms, and other arenas where cultural exchanges might have occurred. Hence, it is possible that some of the pattern in our data is explained by a SAS-Widerøe commonality, subsequently lending support to the claim that organizational informal and formal elements might affect pilot's perception of flight safety also in this study. If that is the case, it would not be a unique finding. Numerous investigations into pilots' attitudes towards flight safety has pointed out the importance of organizational factors (reviewed by Helmreich, R. L., & Merritt, A. C. (2017), and it's a concern for many pilot-interest groups (e.g., ECA). Future research could be aimed at implementing a boarder approach using the Commercial Aviation Safety Survey (CASS) (Gibbons et al., 2006). This model is aimed at using five critical aspects of organizational structure to measure safety culture. Applied in the case of EFB-implementation, the score from such a survey could be used to see directly if variations in score affected the way pilots relate to the various safety statements in this study. In addition, key formal organizational predictors could be included. This was done by Gibbons and colleagues, which for instance recognized the need to account for differences in business models between operators as the management-employee relationship is different in many cases than the traditional setting where the concept of safety culture arose. For our study question however, we are content with recognizing that there is reliable evidence that suggests that the perception of flight safety in relation to the EFB-transition does vary between operators, and that valuable insights could be gained by looking at how the EFB-implementation was handled by these different entities.

In addition to potential organizational differences, we found evidence suggesting that certain human factors are explaining differences in how pilots perceive flight safety to have changed. Depending on their difference in background, attitudes, or experience, pilots have disparate ways of regarding how their flight deck has evolved. This result is hardly surprising. The human psyche is putatively notoriously diverse, causing our perceptions of reality to vary in a multitude of ways. The experience of flying aircraft is no exception, and gains in knowledge about the human-machine interface has eventually led to an integration of human recourse management in aviation (Helmreich, 1999). A deeper understanding of how a pilot's prior beliefs, attitudes, personal life, state of mind, mood, etc. could cause different responses to the same situation, has helped the design of everything from aircraft systems to the development of operational procedures. Hence, the implementation of new procedures or systems to a cockpit environment must be conducted with intimate knowledge on how pilots will respond or interact with it.

Of the human factors included in our study, the data indicate that there are some which are more important for managers to consider than others when evaluating the effectiveness of the EFB-transition. Specifically, how comfortable the EFB is to work with and the seamlessness of its operation is of most importance according to pilots participating in our study. This suggests that the interaction between the EFB and the pilot should be optimized in order to make the EFB-solution more advantageous to flight safety. The other top causes of pilots altering perception, was variation in how familiar they were with touch screens, the degree of training they have received, and their differences in attitudes towards a more digitized society. Interestingly, some of these finding is similar to results elsewhere using more objective data sources. Chandra and Kendra (2009) used the ASRS to analyse in-flight errors committed while using EFBs. One of the top causes was found to be “display configuration” issues, indicating that pilots struggle with using the device as intended. In addition, lack of proper training was reported to be an important issue. Chase and colleagues (2014), using a similar objective data source, found that crew members consistently ran into similar ergonomical problems as they had trouble manipulating the EFB displays, contributing to a misinterpretation of performance data. Evidently, the interaction between pilot and device is crucial, both when considering self-reporting data sources and incident reports. These results are perhaps not surprising considering the fact that aircraft and avionic manufacturers for years have tried to make the human-machine interface as streamlined as possible. Indeed, one of the leading contributors to serious incidents and accidents are misinterpretation or the failure to interreact properly with instrumentation (Wiegman & Shappell, 2001). What neither we nor the other studies investigating EFB-pilot interaction address directly, is whether the occurrence of these factors in incident statistics and pilot self-reporting, is because of the physical interface itself, or the act of mastering the “mental mode” implemented in the EFB architecture and logic. However, in our data, the familiarity with touch screens rank as the second highest factor in explaining variation. This might suggest that it is the physical interaction that is more important, since the degree to which one is used to working with the same sort of equipment one is confronted with on the flight deck, significantly affects perception.

Another valuable insight is perhaps that all these top human explanatory factors present in our models, as a large has a positive effect on the perception of flight safety. That is to say that no matter how unfamiliar pilots are with touch screens, how well they regard the training they received, or how they perceive the EFB interaction, they still view the EFB flight deck an improvement compared to the traditional cockpit. Indeed, this is true for the majority of factors. For our best representation of the relationship between different independent and dependent variables, the only human factors that show a persistent negative contribution to the perception of flight safety, is the interaction between touch screen familiarity and a negative attitude towards digitalization, and between touch screen familiarity and the tendency to be stressed by automation. The results of this might be difficult to interpret without a deeper understanding of the underlying mechanism, for example by conducting in depth interviews with the pilots themselves. However, some clarity can be had by examining the interactions. For instance, it appears that pilots with a positive attitude towards digitalization always have a good perception of the EFB, even if they have a poor familiarity with touch screens, and that this causes them to view it easier to access information in an EFB cockpit. For automation and touch screen usage, we see that pilots who are more easily stressed by automation are more distracted by administrative duties as touch screen familiarity increase. To speculate to the exact reasons behind these patterns is beyond the capabilities on this study based on the data we've collected. It is interesting that there might complex interactions between different aspect of a pilot's background and abilities that can affect how he might perceive flight safety, subsequently affecting the implementation of EFBs. If this is something that can be accounted for my management remains unlikely and considering that the relative contribution to the explained variation in the data remains low, it would be perhaps wasteful to pursue it.

Often, it can be valuable to review the findings one expected to appear but did not present itself. In aviation, the amount of flight hours a pilot has have been shown to affect that individuals' attitudes towards flight safety. Research shows that the degree to which pilots harbour so called "dangerous attitudes", aptly named for their negative impact on safety, vary according to their experience level (Hunter, 2005). The curve is bell shaped, where the least display of these unfavourable attitudes aggregates at the ends, i.e.; that those with little and very much experience have the most attractable attitudes. Those who have moderate amounts of experience tend to be overconfident. Subsequently, we expected to see that flying experience also would affect how pilots relate to at least some of the various statements in this study. It did not. However, the flight deck role the pilot has does appear to affect perception, and normally commanders are more experienced than flight officers. It is possible that some of the experience effect we expected is nested here. In any respect, we see that commanders to a larger extent perceive the EFB flight deck an improvement compared to first officers. Another perhaps surprising aspect, is that commanders are often older of age. Collectively, it is often believed that younger individuals are more positive towards change as a whole , and particularly when it comes to PEDs, they are more competent users as they are more likely to adopt new technologies (discussed by Caprani et al., 2009). There might be issues with nesting here as well, as the main effect of age shows a different pattern. Here it is the youngest age group which is the only category which consistently view the EFB-flight deck an improvement. Higher resolution of the data is needed to fully resolve this. Also, there is more than age to being a first officer or commander, as the two positions have been shown to perceive the operations of aircraft differently as a factor of their respective roles (Fischer & Orasanu, 2000). One of the weaknesses with any model design is the possibility of nested factors. Our model design did not include a correction for the possibility of nested factors, and so we run a higher risk of measuring effects more than once. Our recommendations for future studies using similar approached would be to include cross tabulation techniques to help mitigate this. In the context of this studies aim as to provide a general overview of the effectiveness of EFB implementation in Norway, we would focus on the finding that both age and flight deck role affects flight safety, and that experience does not appear to contribute as much as expected.

Human factors are influencing the effectiveness of the EFB transition; however, the operator affiliation of pilots still account for most of the observed variation when comparing to each single human factor. This suggests that there are aspects of working for a particular operator that outweighs the impact of any one measurement of experience, attitudes or background. Combined, the human factors have a larger impact, but one must be careful with the interpretation of that apparent finding. This is because regression modelling, even though it is possible to correct for over-fitting, will still reward the inclusion of more potential predictors to some extent. We have only one measurement of organizational differences, and that proxy is also extremely general and could be subdivided into many thinkable subcategories that would appear to explain the data better as a whole. Subsequently, it is more correct to compare each human factor predictor to the operator affiliation predictor. Furthermore, the problem of nesting might also be present for the operator predictor. It is possible that some of the other predictors, specifically the quality of training and the attitudes towards one’s employer, are masked by the operator category. As mentioned previously, a more careful model design could help towards increasing the resolution of the data, providing conditions for more meaningful and accurate inference.

6 Conclusion

In conclusion, the EFB implementation by Norwegian operators appears to have been effective. Hence, the promises of the EFB in providing both a more efficient and effective organization seems to hold true. Almost all our 16 measurements of flight safety in the participants showed that pilots regard the EFB an improvement to the traditional flight deck. However, in some areas the apparent safety gain is perhaps not as high as expected. Particularly, concerns about in-flight distractions and overburdening crews with tasks is present and should be evaluated more carefully in order to mitigate potential threats.

This perception of flight safety was not uniform across all operators, indicating that organizational informal and formal traits could be responsible for this observed variation. This difference in the perception of flight safety regarded as a function of which operators the pilots work for, provide promising evidence for the importance of the organization itself and how it affects flight safety. Untangling different organizational traits and how they influence effectiveness is a natural next step of analysis. Further findings here could help business managers allocate resources better and increase the potential gains in efficiency when implementing new technologies, building more effective safety organizations in the process.

The impact on flight safety by most of the human factors investigated here are not statistically robust enough to be conclusive. This is surprising considering the extent to which the flight safety literature is occupied with these aspects. A possibility is that our human factors are not as relevant to the EFB-flight safety complex as anticipated. However, the strict statistical criteria employed here could possibly have resulted in Type-II errors, erroneously accepting the null hypotheses. Larger data sets, and a more careful selection of appropriate statistical methodology could provide a resolution to these concerns. Statistical analysis can be a potent tool in providing useful insights, but it also has its limits, especially when one is dealing with human behaviour in an uncontrolled environment such as this one. Subsequently, we recommend that future inquiries exploring the same area to supplement with more qualitative approaches.

References

Afifi, A., May, S., & Clark, V. A. (2003). *Computer-aided multivariate analysis*. CRC Press.

Aggarwal, C. C. (2015). Outlier analysis. In *Data mining* (pp. 237-263). Springer, Cham.

Airplanes, B. C. *Paperless Cockpit*.

Allen, P. J., & Bennett, K. (2007). *SPSS for the health and behavioural sciences*. Thomson Learning.

Alwin, D. F. (2007). *Margins of error: A study of reliability in survey measurement* (Vol. 547). John Wiley & Sons.

Anderson, M. J. (2001). Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian journal of fisheries and aquatic sciences*, *58*(3), 626-639.

Ateş, S. S. (2017). Electronic Flight Bag in the Operation of Airline Companies: Application in Turkey. *Computer Science and Information Technology*, *5*(4), 128-134.

Barley, S.R, G.W. Meyer & D.C. Gash. (1988). Cultures of Culture: Academics, Practitioners, and the Pragmatics of Normative Control. *Administrative Science Quarterly*, 33: 24-60.

Bjørnskau, T. (2005). *Flysikkerhet i Norge*. Retrieved from https://www.toi.no/getfile.php?mmfileid=865

Borgen, M. (2018). Ti Electronic Flight Bag i Widerøe´s Flyveselskap AS - en casestudie om organisatorisk endring ved overgang til digitale løsninger (Master of Business Administration (Thesis), Nord University). Retrived from https://nordopen.nord.no/nord-xmlui/handle/11250/2566834

Cahill, J., & McDonald, N. (2006). Human computer interaction methods for electronic flight bag envisionment and design. *Cognition, Technology & Work*, 8, 113-123.

Caprani, N., O'Connor, N. E., & Gurrin, C. (2012). Touch screens for the older user. In *Assistive technologies*. IntechOpen.

Carey, S. (2013). *iPads help airlines cast off costly load of paper*. Retrieved from https://www.wsj.com/articles/SB10001424127887323998604578567720762762606

Carr, N. (2013). All can be lost: The risk of putting our knowledge in the hands of machines. *The Atlantic*, *11*, 1-12.

Chandra, D. C. & Mangold S. J. (2000a*). Human factors considerations in the design and evaluation of electronic flight bags (EFBs) Version 1: Basic functions*. (Report No. DOT-VNTSC-FAA-00-22). Cambridge, MA: U.S. DOT Volpe Center.

Chandra, D. C. & Mangold S. J. (2000). *Human factors considerations for the design and evaluation of Electronic Flight Bags*. Proceedings of the 19th Digital Avionics Systems Conference. 10-12 October 2000, Philadelphia, PA.

Chandra, D. C., Yeh, M., Riley, V., & Mangold, S. J. (2003). *Human factors considerations in the design and evaluation of electronic flight bags (EFBs): version 2* (No. DOT-VNTSC-FAA-03-07). United States. Federal Aviation Administration.

Chandra, D. C., & Kendra, A. (2009). *Review of Safety Reports Involving Electronic Flight Bags*. 2009 International Symposium on Aviation Psychology, 407-412. https://corescholar.libraries.wright.edu/isap_2009/48

Chase, S., and Hiltunen, D. (2014). *An Examination of Safety Reports Involving Electronic Flight Bags and Portable Electronic Devices* (DOT-VNTSC-FAA-14-12). Cambridge, MA: U.S. DOT Volpe Center.

Chatman, J.A. & S.E. Spataro. (2005). Using Self-Categorization Theory to Understand Relational Demography – Based Varioations in Peoples Responsivness to Organizational Culture. *Academy Of Management Journal*, 48(2), 321-331.

Chevan A., Sutherland M (1991). "Hierarchical Partitioning." *The American Statistician*, 45, 90–96.

Civil Aviation Safety Authority (CASA) (2013). *Electronic Flight Bags*. Retrieved from https://www.casa.gov.au/aircraft/standard-page/electronic-flight-bag-efb

Deal, T. E. (1982). Corporate cultures. Addison-Wesley.

Department of Transportation, John A. Volpe National Transportation Systems Center. (2003). *Human factors considerations in the design and evaluation of Electronic Flight Bags (EFBs): Version 2*. (DOT Publication No. DOT-VNTSC-FAA-03-07). Retrieved from http://ntl.bts.gov/lib/34000/34200/34292/DOT-VNTSC-FAA-03-07.pdf

Department of Transportation, John A. Volpe National Transportation Systems Center. (2010). *Review of safety reports involving electronic flight bags*. (DOT Publication No. DOT-VNTSC-FAA-10-08). Retrieved from http://ntl.bts.gov/lib/44000/44100/44161/efb042010_1_.pdf

Department of Transportation, John A. Volpe National Transportation Systems Center. (2014). *An examination of safety reports involving electronic flight bags and portable electronic devices.* (DOT Publication No. DOT-VNTSC-FAA-14-12). Retrieved from http://ntl.bts.gov/lib/51000/51900/51968/An_Examination_of_Safety_Reports_Involving_EFBs_and_PEDs.pdf

DiMaggio, P. J., & Powell, W. W. (1991). Introduction. The new institutionalism in organizational analysis. *The New Institutionalism in Organizational Analysis, University of Chicago Press, Chicago, IL*, 1-38.

E. Fix and J. Hodges. Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques, chapter Discriminatory analysis- nonparametric discrimination: Consistency principles. IEEE Computer Society Press, Los Alamitos, CA, 1951. Reprint of original work from 1952.

Edwards, E. (1988). Introductory overview. In *Human factors in aviation* (pp. 3-25). Academic Press.

Endsley, M. R. (1999). Situation awareness in aviation systems. *Handbook of aviation human factors*, 257-276.

European Cockpit Association (2016, February). Every incident or accident is different - The Need for a Just Culture Review Committee. Retrieved from https://www.eurocockpit.be/news/every-incident-or-accident-different-need-just-culture-review-committee

European Cockpit Association (2019, March). Beyond breaking point – is the pilot an 'aircraft part', or a person?. Retrieved from https://www.eurocockpit.be/news/beyond-breaking-point-pilot-aircraft-part-or-person

European Union Aviation Safety Agency (EASA) (2014). *Airworthiness and operational consideration for Electronic Flight Bags (EFBs)*. Retrieved from https://www.easa.europa.eu

Federal Aviation Administration. (2002). *Guidelines for the certification, airworthiness, and operational use of electronic flight bags*. (FAA Publication No. AC 120-76). Retrieved from http://rgl.faa.gov/Regulatory_and_Guidance_Library/rgAdvisoryCircular.nsf/list/AC%20120-76/$FILE/AC120-76.pdf

Federal Aviation Administration. (2007). *Use of Class 1 or Class 2 Electronic Flight Bag (EFB)*. (FAA Publication No. AC 91-78). Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentid/73540

Federal Aviation Administration. (2015). *Safety Management Systems for Aviation Service Providers*. (FAA Publication No. AC 120-92B). Retrieved from https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_120-92B.pdf

Fischer, U., & Orasanu, J. (2000). Do you see what I see? Effects of crew position on interpretation of flight problems. NASA Tech. Mem. 2000-xxxxxx. Moffett Field, CA: NASA Ames Research Center.

Flannery, J. A. (2001). Safety culture and its measurement in aviation. *University of Newcastle, Australia*.

Flight Safety Foundation. (2005, June). Paperless cockpit promises advances in safety, efficiency. *Flight Safety Digest*, 24(6). Retrieved from http://flightsafety.org/fsd/fsd_june05.pdf

Flouris, T., & Yilmaz, A. K. (2009). Change management as a road map for safety management system implementation in aviation operations: focusing on risk management and operational effectiveness. *International Journal of Civil Aviation*, *1*(1), E1.

Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., ... & Heiberger, R. (2012). Package ‘car’. *Vienna: R Foundation for Statistical Computing*.

Fukuyama, F. (1995). Trust: The social virtues and the creation of prosperity.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., & Hothorn, T. (2008). mvtnorm: Multivariate Normal and t Distributions. *R package version 0.9-2, URL http://CRAN. R-project. org/package= mvtnorm*.

Ghosh, A. K. (2006). On optimum choice of k in nearest neighbor classification. *Computational Statistics & Data Analysis*, *50*(11), 3113-3123.

Ghosh, D., & Vogt, A. (2012, July). Outliers: An evaluation of methodologies. In *Joint statistical meetings* (pp. 3455-3460). San Diego, CA: American Statistical Association.

Gibbons, A. M., von Thaden, T. L., & Wiegmann, D. A. (2006). Development and initial validation of a survey for assessing safety culture within commercial flight operations. *The international journal of Aviation Psychology*, *16*(2), 215-238.

Grömping, U. (2006). Relative importance for linear regression in R: the package relaimpo. *Journal of statistical software*, *17*(1), 1-27.

Haddock, K. N., & Beckman, W. S. (2015). The effect of electronic flight bag use on pilot performance during an instrument approach. *Collegiate Aviation Review*, *33*(1), 1-13.

Hamblin, C. J. (2004). Usability of mobile devices in the cockpit. Retrieved from Wichita State University Psychology Department website: http://psychology.wichita.edu/ surl/usabilitynews/61/efb.htm

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. (1986). *Robust statistics* (pp. 29-30). New York: Wiley.

Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999). The evolution of crew resource management training in commercial aviation. *The international journal of aviation psychology*, *9*(1), 19-32.

Helmreich, R. L., & Merritt, A. C. (2017). *Culture at work in aviation and medicine: National, organizational and professional influences*. Routledge.

Higdon, D. (2016, June). Pilot Electronic Flight Bags: Benefits vs Risks . Retrieved from https://www.avbuyer.com

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 65-70.

Huang, P. H., & Wu, H. M. (1994). More order without more law: A theory of social norms and organizational cultures. *JL Econ. & Org.*, *10*, 390.

Hunter, D. R. (2005). Measurement of hazardous attitudes among pilots. *The International Journal of Aviation Psychology*, *15*(1), 23-43.

International Civil Aviation Organization. (2012). *Safety Management Manual (SMM)*. (ICAO Doc. 9859 AN/474). Retrieved from https://www.icao.int/SAM/Documents/RST-SMSSSP-13/

International Civil Aviation Organization. (2019). *SARPs, Standard and recommended Practices*. Retrieved from https://www.icao.int/safety/safetymanagement/pages/sarps.aspx

Jacobsen, D. I., & Thorsvik, J. (2013). Hvordan organisasjoner fungerer, 4 utgave.

Johannessen, A. C., & Christoffersen, L. L. & Tufte, PA (2011). *Forskningsmetode for økonomiskadministrative fag*, *4*.

Johnson, P. E. (2016). GLM with a Gamma-distributed dependent variable.

Jorens, Y., Gillis, D., Valcke, L., De Coninck, J., Devolder, A., & De Coninck, M. (2015). Atypical forms of employment in the aviation sector', European social dialogue, European Commission, 2015. Ghent, Belgium.

Kalton, G., & Schuman, H. (1982). The effect of the question on survey responses: A review. *Journal of the Royal Statistical Society: Series A (General)*, *145*(1), 42-57.

Kang, N. E., & Yoon, W. C. (2008). Age-and experience-related user behavior differences in the use of complicated electronic devices. *International Journal of Human-Computer Studies*, *66*(6), 425-437.

Kaplan, R. S., & Norton, D. P. (2000). Having trouble with your strategy? Then map it. *Focusing Your Organization on Strategy—with the Balanced Scorecard*, *49*.

Kletz, T. (2013). *If you think safety is expensive, try an accident*. Retrieved from http://psc.tamu.edu/trevor-kletz

Komsta, L., & Novomestky, F. (2015). Moments, cumulants, skewness, kurtosis and related tests. *R package version*.

Kotrlik, J. W. K. J. W., & Higgins, C. C. H. C. C. (2001). Organizational research: Determining appropriate sample size in survey research appropriate sample size in survey research. *Information technology, learning, and performance journal*, *19*(1), 43.

Latham, G. P., & Pinder, C. C. (2005). Work motivation theory and research at the dawn of the twenty-first century. *Annu. Rev. Psychol.*, *56*, 485-516.

Lee, J., & See, K. (2004). Trust in automation: designing for appropriate reliance. *Human Factors and Ergonomics Society*, 46(1), 50-80.

Li, G., Baker, S. P., Grabowski, J. G., Qiang, Y., McCarthy, M. L., & Rebok, G. W. (2003). Age, flight experience, and risk of crash involvement in a cohort of professional pilots. *American journal of epidemiology*, *157*(10), 874-880.

Lincoln, J. R., & Kalleberg, A. L. (1992). *Culture, control and commitment: A study of work organization and work attitudes in the United States and Japan*. CUP Archive.

Lindeman RH, Merenda PF, Gold RZ (1980). Introduction to Bivariate and Multivariate Analysis. Scott, Foresman, Glenview, IL.

Lytle, D. (2015). *Pilot perception of Electronic Flight Bags at Part 121 air carriers*. Retrieved from ProQuest Dissertations and Theses, The University of North Dakota.

Martin, W. L., Murray, P. S., Bates, P. R., & Lee, P. S. (2015). Fear-potentiated startle: A review from an aviation perspective. *The International Journal of Aviation Psychology*, *25*(2), 97-107.

Mayo, E. (1945). The Social Problems of An Industrial Civilization (Boston: Harvard University Graduate School of Business Administration).

McKenna, E. (2013, October). Product Focus: Electronic Flight Bags. Retrieved from ttps://www.aviationtoday.com

McQuarrie, A. D., & Tsai, C. L. (1998). *Regression and time series model selection*. World Scientific.

Mitteroecker, P., & Bookstein, F. (2011). Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. *Evolutionary Biology*, *38*(1), 100-114.

O'Reilly, C. (1989). Corporations, culture, and commitment: Motivation and social control in organizations. *California management review*, *31*(4), 9-25.

Ostroff, C., & Schmitt, N. (1993). Configurations of organizational effectiveness and efficiency. *Academy of management Journal*, *36*(6), 1345-1361.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors and Ergonomics Society*, 39(2), 230-253.

Perrow, C. (1986). Complex Organizations: A Critical Essay. Glenview, Ill: Scott Foresmann.

Peters, T.J. & R.H. Waterman jr. (1982). In Search of Exellence. New York: Harper and Row.

Pidgeon, N. and O'Leary, M. (1994). Organisational safety culture: Implications for aviation practice. I: Johnston, N., Fuller, R., McDonald, N. (Eds): Aviation Psychology in Practice. Avebury Technical, Alderschot, s. 21-43.

Powell W.W & P.J. Dimaggio (1991). The New Institutionalism in Organizational Analysis. Chicago : The University of Chicago Press.

Provera, B., Montefusco, A., & Canato, A. (2010). A ‘no blame’approach to organizational learning. *British Journal of Management*, *21*(4), 1057-1074.

Raffalovich, L. E., Deane, G. D., Armstrong, D., & Tsao, H. S. (2008). Model selection procedures in social research: Monte-Carlo simulation results. *Journal of Applied Statistics*, *35*(10), 1093-1114.

Ray, C.A. (1980). Corporate Culture. The Last Frontier of Control. *Journal of Management Studies* 23: 287-297.

Reason, J. (1998). Achieving a safe culture: theory and practice. *Work & Stress*, *12*(3), 293-306.

Reason, J., Hollnagel, E., & Paries, J. (2006). Revisiting the Swiss cheese model of accidents. *Journal of Clinical Engineering*, *27*(4), 110-115.

Reason, J. (2016). *Managing the risks of organizational accidents*. Routledge.

Schlager, S. (2016). Morpho: calculations and visualizations related to geometric morphometrics. Version 2.4. 1.1. *Retrieved October*, *30*, 2017.

Scott, W. R. (1987). Organizations: Rational, Natural and Open Systems (2e édition). Englewood Cliffs (N.J.): Prentice Hall.

Scott, J. E. (2007). An e-transformation study using the technology-organization-environment framework. *BLED 2007 Proceedings*, 55.

Smith, M. J., & Carayon, P. (1995). New technology, automation, and work organization: stress problems and improved technology implementation strategies. *International Journal of Human Factors in Manufacturing*, *5*(1), 99-116.

Stajkovic, A. D., & Luthans, F. (2001). Differential effects of incentive motivators on work performance. *Academy of management journal*, *44*(3), 580-590.

Stolzer, A., Halford, C., & Goglia, J. (2008). Safety Management Systems in aviation. Burlington, VT: Ashgate.

Taylor, J. L., Kennedy, Q., Noda, A., & Yesavage, J. A. (2007). Pilot age and expertise predict flight simulator performance: a 3-year longitudinal study. *Neurology*, *68*(9), 648-654.

Team, R. C. (2017). R: A language and environment for statistical com-puting. *R Foundation for Statistical Computing, Vienna, Austria. URL https://www. R-project. org*.

Thomas, M. J. (2012). *A systematic review of the effectiveness of safety management systems* (No. AR-2011-148). Canberra, Australia: Australian Transport Safety Bureau.

Todnem By, R. (2005). Organisational change management: A critical review. *Journal of change management*, *5*(4), 369-380.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, *5*(2), 99-114.

Tump, R.S., van Es, G.W.H., Karwal, A.K. & Verbeek, M.J. (2014). *Best practices for approval of Performance and MB applications on EFBs*. Retrieved from https://www.easa.europa.eu/sites/default/files/dfu/EASA-Research-Rep-2014-1.pdf

Volpe (2018, April). Volpe EFB Publications and Projects. Retrieved from https://www.volpe.dot.gov/safety-management-and-human-factors/transportation-human-factors/volpe-efb-publications-and-projects

Weakliem, D. L. (2016). *Hypothesis testing and model selection in the social sciences*. Guilford Publications.

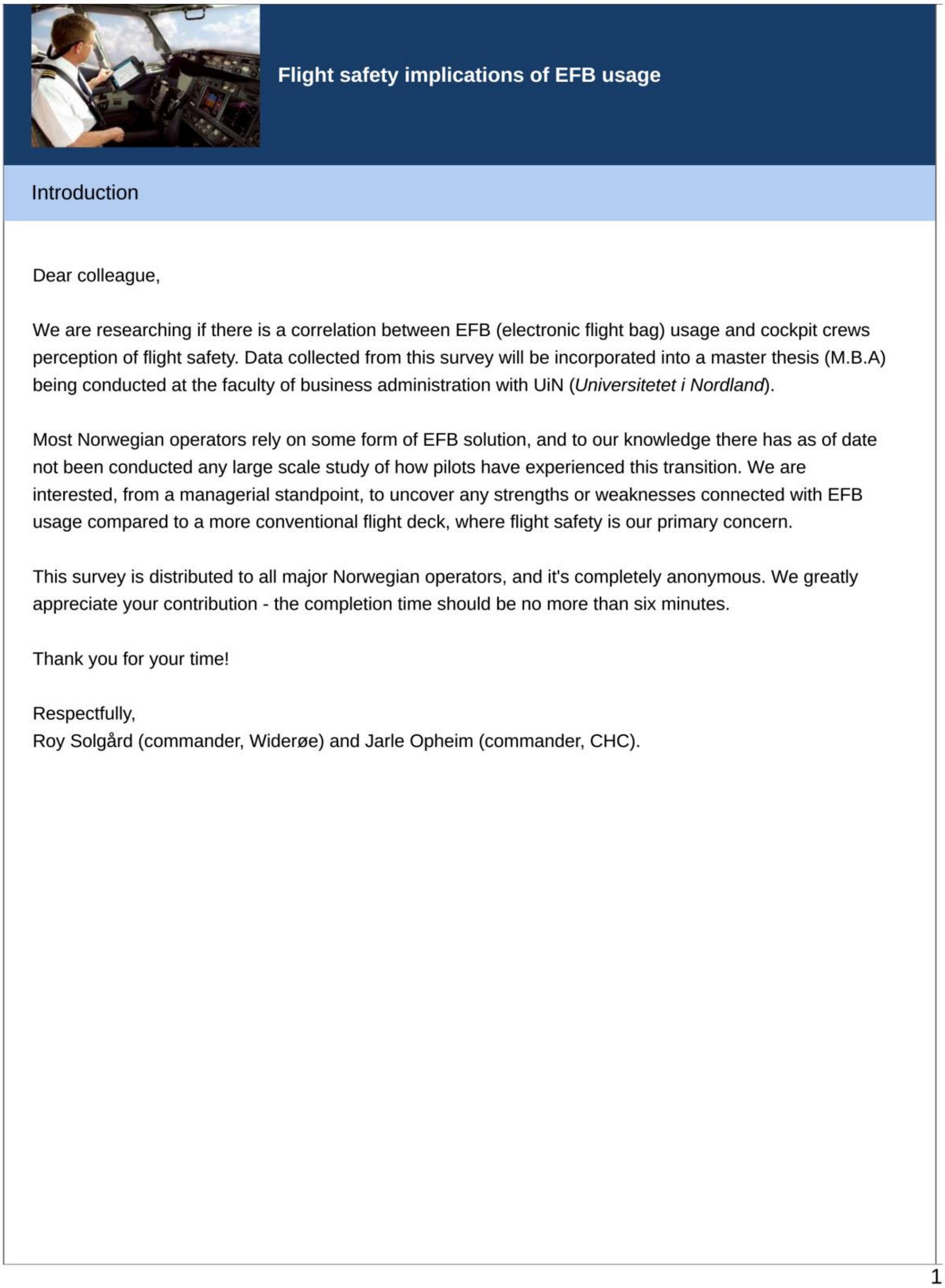
Wiegmann, D. A., & Shappell, S. A. (2001). Human error perspectives in aviation. *The International Journal of Aviation Psychology*, *11*(4), 341-357.

Wong, D. S., Desai, V. M., Madsen, P., Roberts, K. H., & Ciavarelli, A. (2005). Measuring organizational safety and effectiveness at NASA. *Engineering Management Journal*, *17*(4), 59-62.

Woods, D. D., Dekker, S., Cook, R., Johannesen, L., & Sarter, N. (2017). *Behind human error*. CRC Press.

You, X., Ji, M., & Han, H. (2013). The effects of risk perception and flight experience on airline pilots' locus of control with regard to safety operation behaviors. *Accident Analysis & Prevention*, *57*, 131-139.

Appendix I: supplementary figures



Flight safety implications of EFB usage

Introduction

Dear colleague,

We are researching if there is a correlation between EFB (electronic flight bag) usage and cockpit crews perception of flight safety. Data collected from this survey will be incorporated into a master thesis (M.B.A) being conducted at the faculty of business administration with UiN (*Universitetet i Nordland*).

Most Norwegian operators rely on some form of EFB solution, and to our knowledge there has as of date not been conducted any large scale study of how pilots have experienced this transition. We are interested, from a managerial standpoint, to uncover any strengths or weaknesses connected with EFB usage compared to a more conventional flight deck, where flight safety is our primary concern.

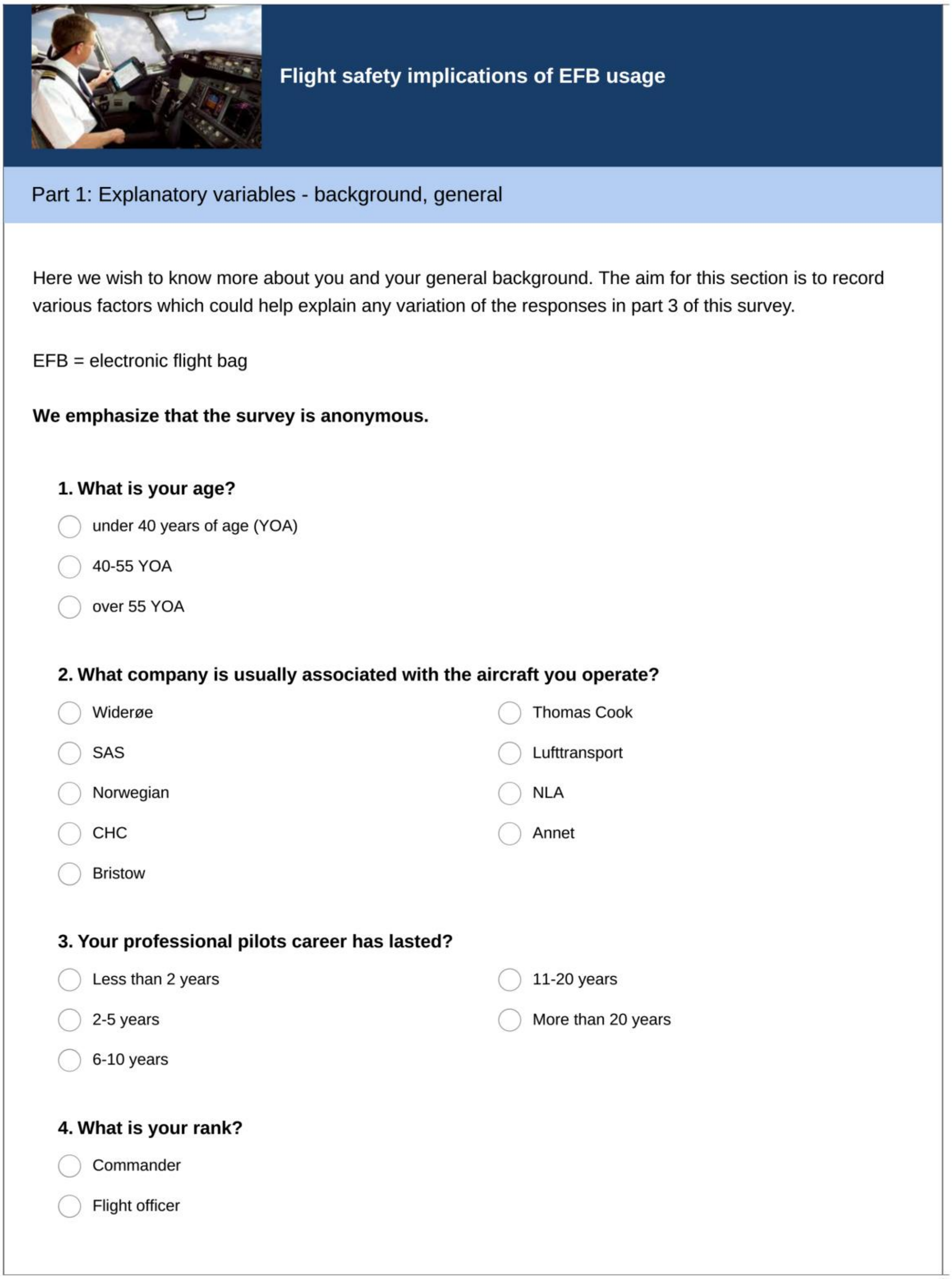
This survey is distributed to all major Norwegian operators, and it's completely anonymous. We greatly appreciate your contribution - the completion time should be no more than six minutes.

Thank you for your time!

Respectfully,
Roy Solgård (commander, Widerøe) and Jarle Opheim (commander, CHC).

1

Cont.



Flight safety implications of EFB usage

Part 1: Explanatory variables - background, general

Here we wish to know more about you and your general background. The aim for this section is to record various factors which could help explain any variation of the responses in part 3 of this survey.

EFB = electronic flight bag

We emphasize that the survey is anonymous.

1. What is your age?

○ under 40 years of age (YOA)
○ 40-55 YOA
○ over 55 YOA

2. What company is usually associated with the aircraft you operate?

○ Widerøe
○ SAS
○ Norwegian
○ CHC
○ Bristow
○ Thomas Cook
○ Lufttransport
○ NLA
○ Annet

3. Your professional pilots career has lasted?

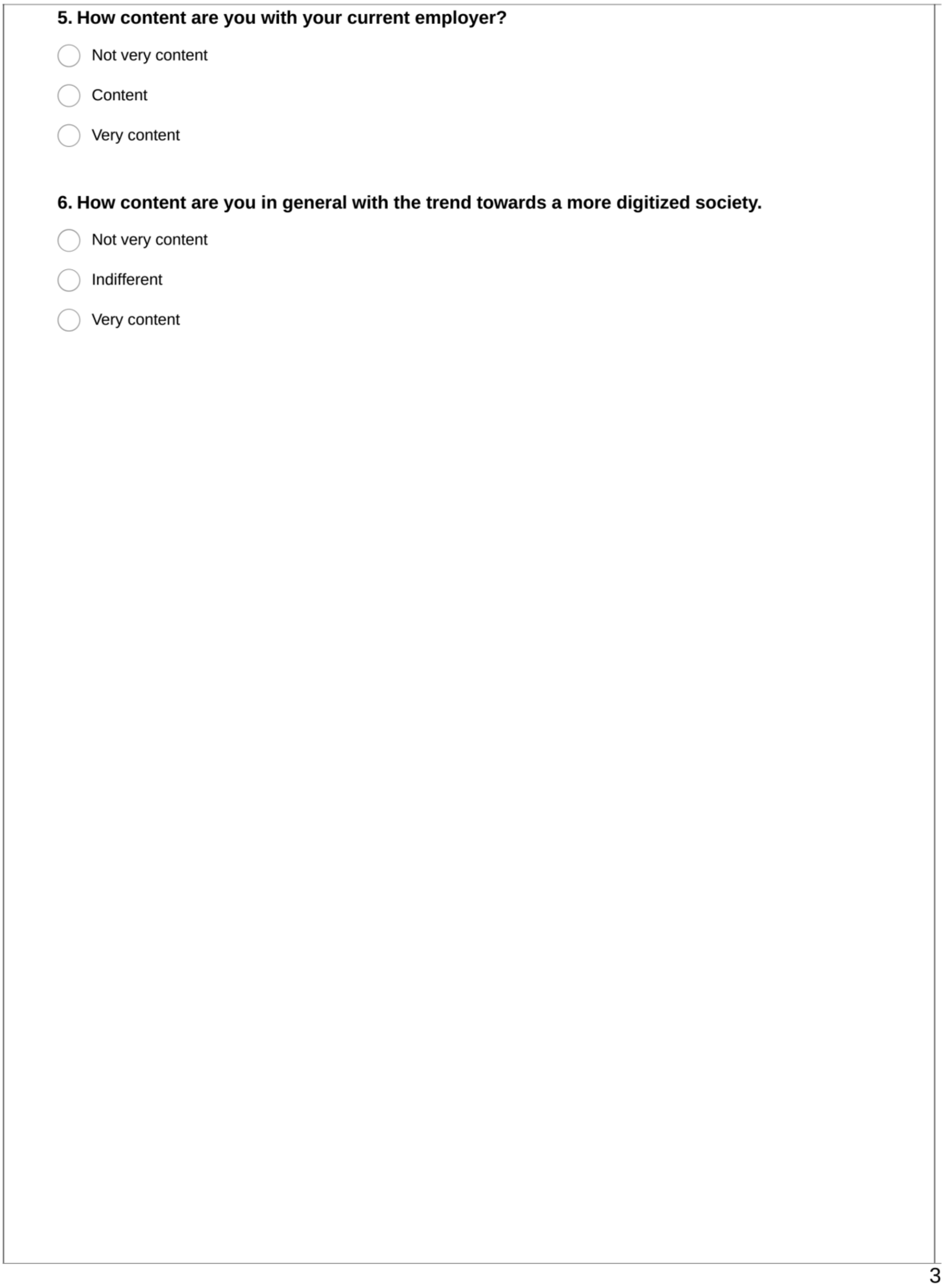
○ Less than 2 years
○ 2-5 years
○ 6-10 years
○ 11-20 years
○ More than 20 years

4. What is your rank?

○ Commander
○ Flight officer

2

Cont



5. How content are you with your current employer?

○ Not very content

○ Content

○ Very content

6. How content are you in general with the trend towards a more digitized society.

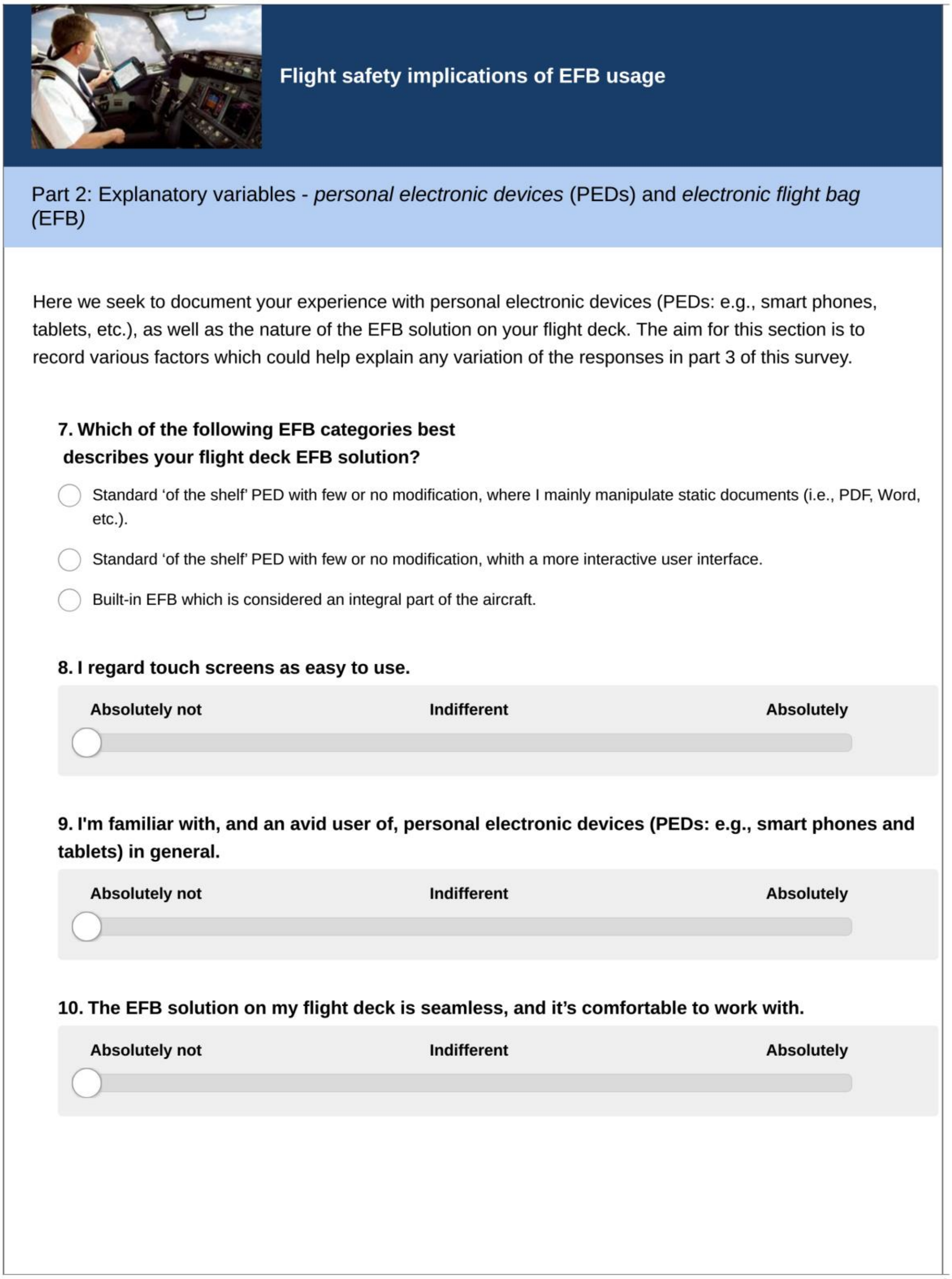
○ Not very content

○ Indifferent

○ Very content

3

Cont.



Flight safety implications of EFB usage

Part 2: Explanatory variables - *personal electronic devices* (PEDs) and *electronic flight bag (*EFB*)*

Here we seek to document your experience with personal electronic devices (PEDs: e.g., smart phones, tablets, etc.), as well as the nature of the EFB solution on your flight deck. The aim for this section is to record various factors which could help explain any variation of the responses in part 3 of this survey.

7. Which of the following EFB categories best describes your flight deck EFB solution?

○ Standard 'of the shelf' PED with few or no modification, where I mainly manipulate static documents (i.e., PDF, Word, etc.).

○ Standard 'of the shelf' PED with few or no modification, whith a more interactive user interface.

○ Built-in EFB which is considered an integral part of the aircraft.

8. I regard touch screens as easy to use.

Absolutely not **Indifferent** **Absolutely**

9. I'm familiar with, and an avid user of, personal electronic devices (PEDs: e.g., smart phones and tablets) in general.

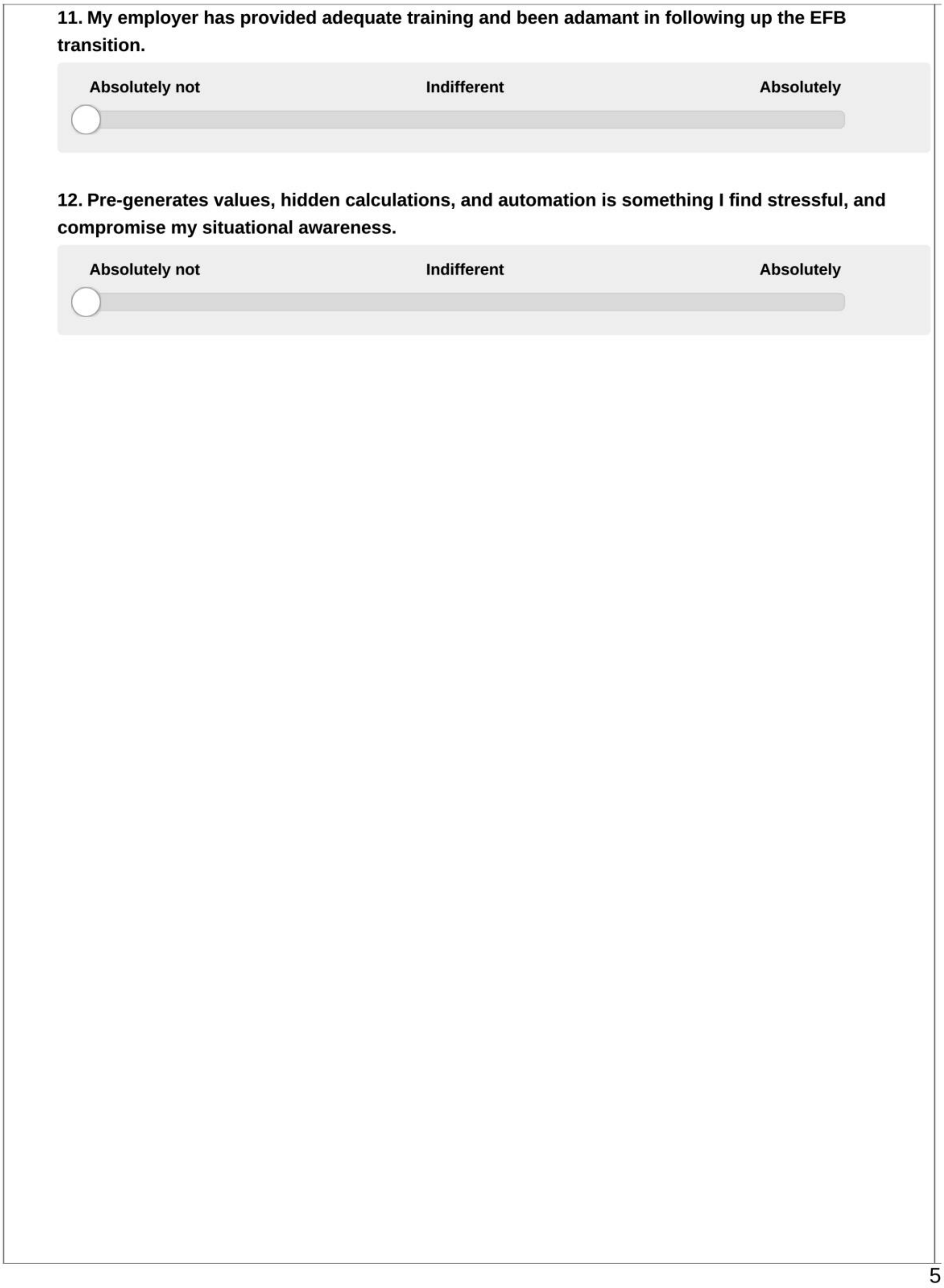
Absolutely not **Indifferent** **Absolutely**

10. The EFB solution on my flight deck is seamless, and it's comfortable to work with.

Absolutely not **Indifferent** **Absolutely**

4

Cont.



11. My employer has provided adequate training and been adamant in following up the EFB transition.

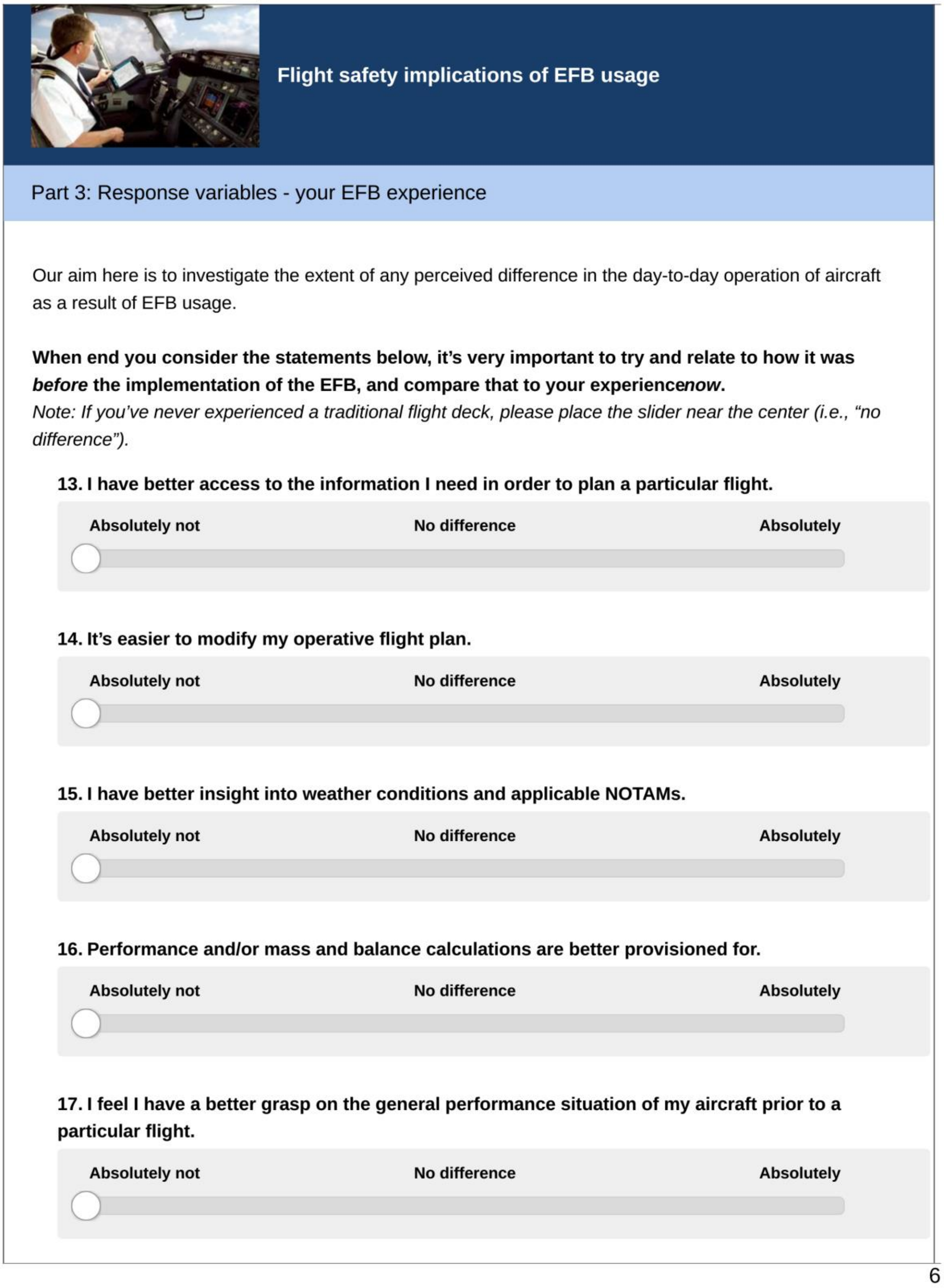
Absolutely not — Indifferent — Absolutely

12. Pre-generates values, hidden calculations, and automation is something I find stressful, and compromise my situational awareness.

Absolutely not — Indifferent — Absolutely

5

Cont.



Flight safety implications of EFB usage

Part 3: Response variables - your EFB experience

Our aim here is to investigate the extent of any perceived difference in the day-to-day operation of aircraft as a result of EFB usage.

When end you consider the statements below, it's very important to try and relate to how it was *before* the implementation of the EFB, and compare that to your experience*now*.
Note: If you've never experienced a traditional flight deck, please place the slider near the center (i.e., "no difference").

13. I have better access to the information I need in order to plan a particular flight.

Absolutely not	**No difference**	**Absolutely**
○		

14. It's easier to modify my operative flight plan.

Absolutely not	**No difference**	**Absolutely**
○		

15. I have better insight into weather conditions and applicable NOTAMs.

Absolutely not	**No difference**	**Absolutely**
○		

16. Performance and/or mass and balance calculations are better provisioned for.

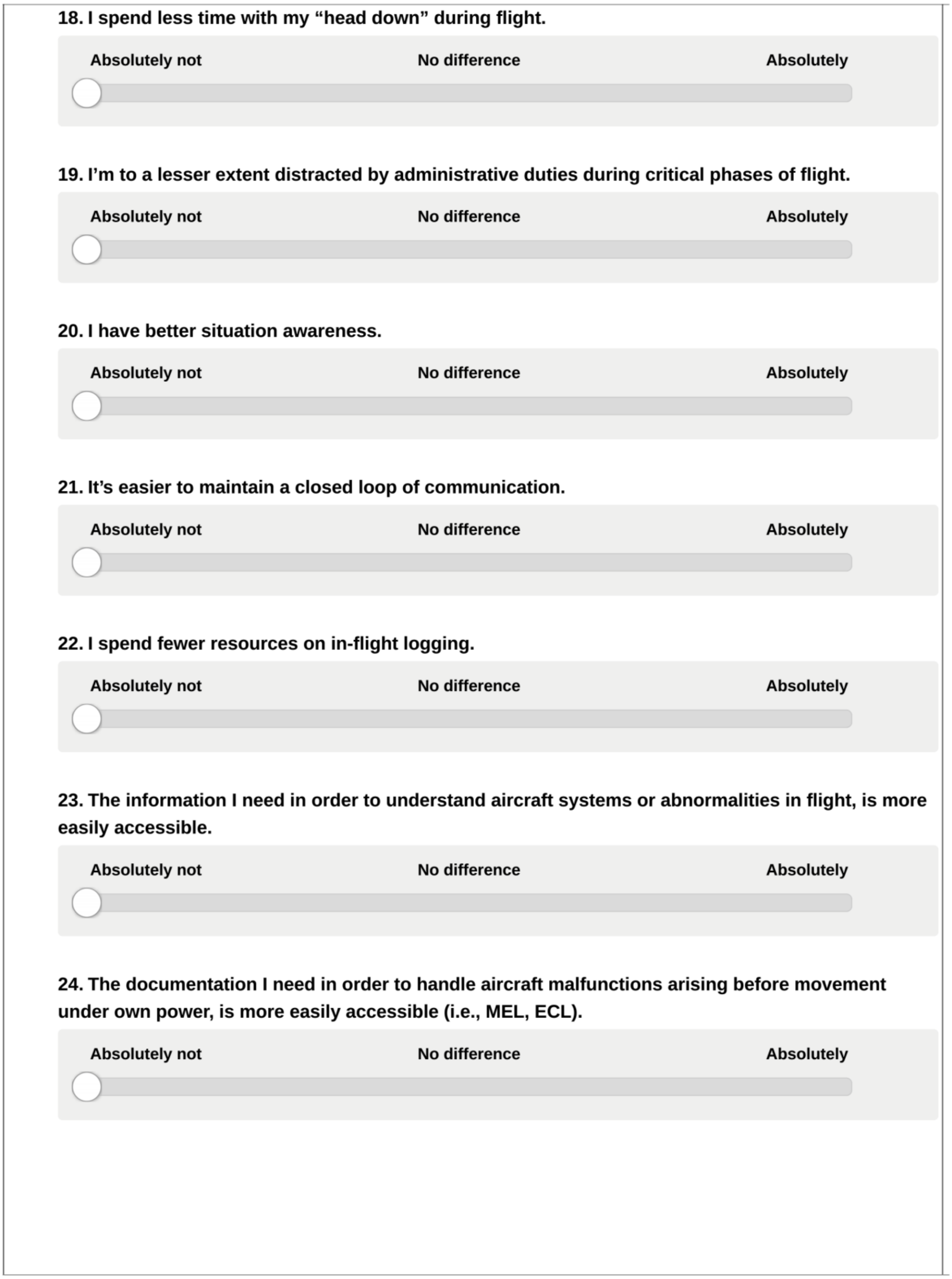
Absolutely not	**No difference**	**Absolutely**
○		

17. I feel I have a better grasp on the general performance situation of my aircraft prior to a particular flight.

Absolutely not	**No difference**	**Absolutely**
○		

6

Cont.



18. I spend less time with my “head down” during flight.

Absolutely not No difference Absolutely

19. I’m to a lesser extent distracted by administrative duties during critical phases of flight.

Absolutely not No difference Absolutely

20. I have better situation awareness.

Absolutely not No difference Absolutely

21. It’s easier to maintain a closed loop of communication.

Absolutely not No difference Absolutely

22. I spend fewer resources on in-flight logging.

Absolutely not No difference Absolutely

23. The information I need in order to understand aircraft systems or abnormalities in flight, is more easily accessible.

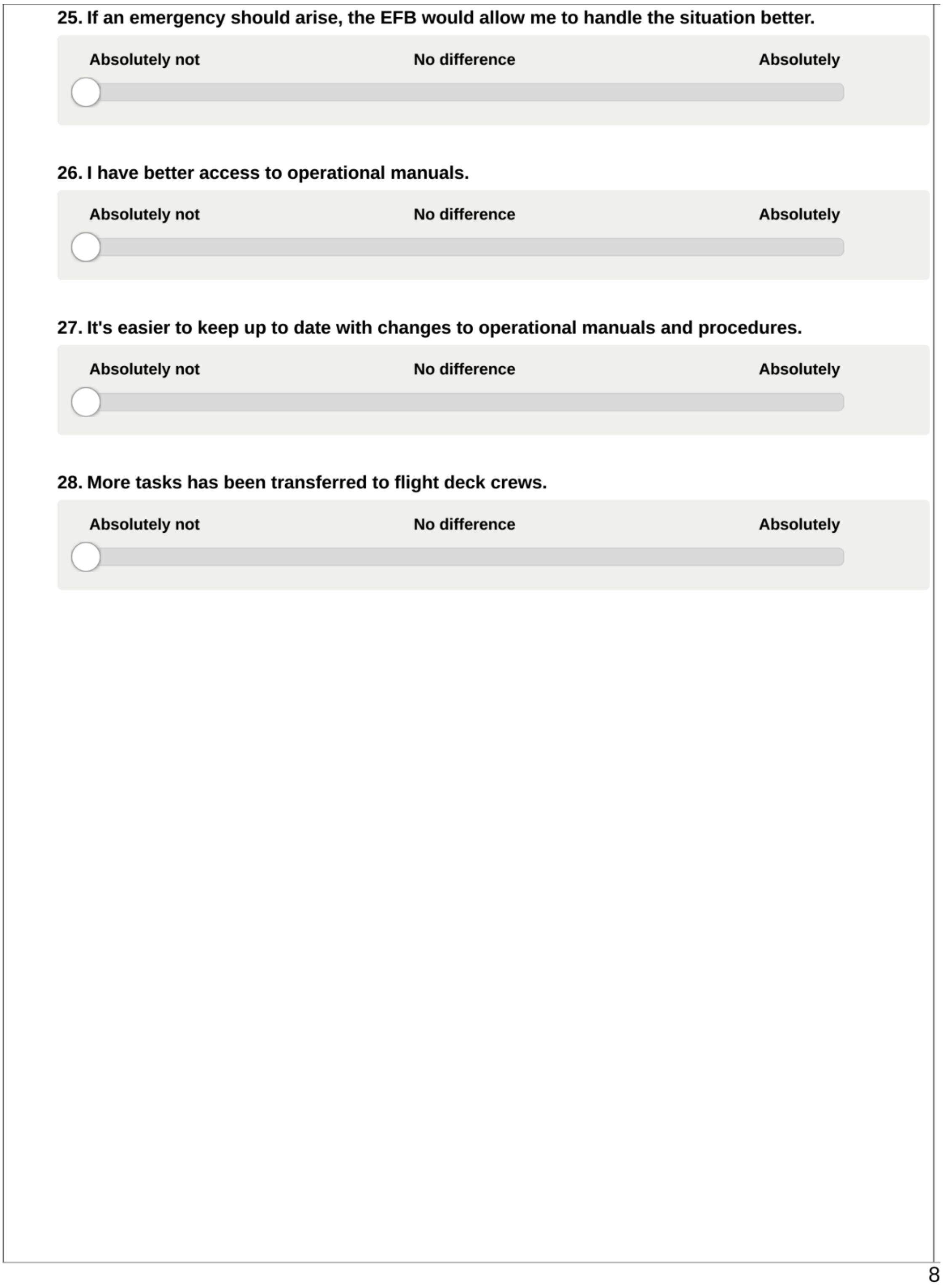
Absolutely not No difference Absolutely

24. The documentation I need in order to handle aircraft malfunctions arising before movement under own power, is more easily accessible (i.e., MEL, ECL).

Absolutely not No difference Absolutely

7

Cont.



25. If an emergency should arise, the EFB would allow me to handle the situation better.

Absolutely not **No difference** **Absolutely**

○

26. I have better access to operational manuals.

Absolutely not **No difference** **Absolutely**

○

27. It's easier to keep up to date with changes to operational manuals and procedures.

Absolutely not **No difference** **Absolutely**

○

28. More tasks has been transferred to flight deck crews.

Absolutely not **No difference** **Absolutely**

○

8

Figure S1. Questionnaire as it appeared for the participant on a non-tablet type PED (last two pages omitted as they only contained non-relevant post-survey feedback information etc.). The survey was designed and distributed using the manufacturer Survey Monkey. The questionnaire is was distributed in the Norwegian language mainly (for that version, please contact the authors (see study cover for more information)).

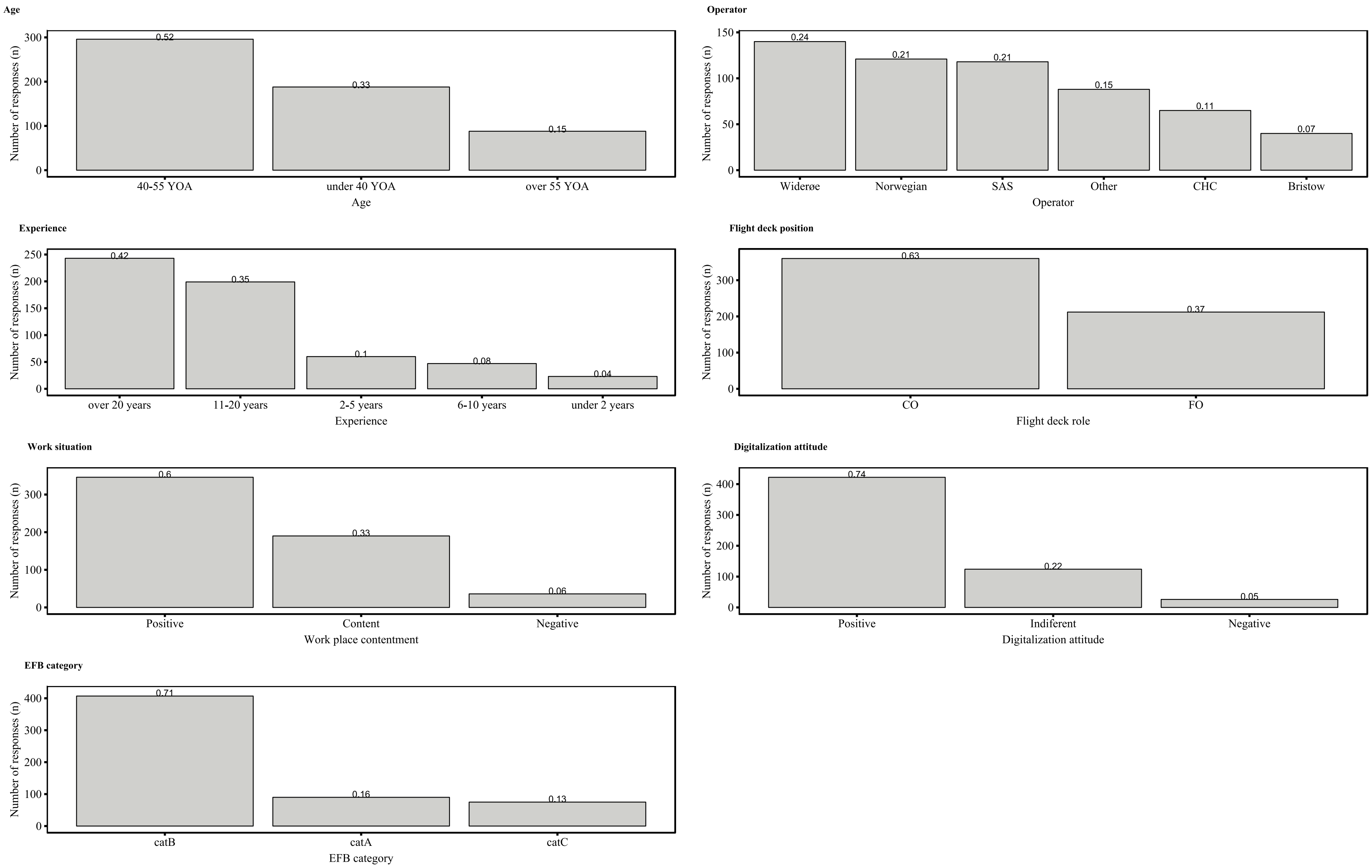


Figure S2. Questionnaire responses for all multiple-choice questions or statements. The bars represent the number of responses within each predictor grouping, with the relative frequency displayed above each bar. The number of responses (n) are derived from the outlier corrected, imputated total response (N = 571).

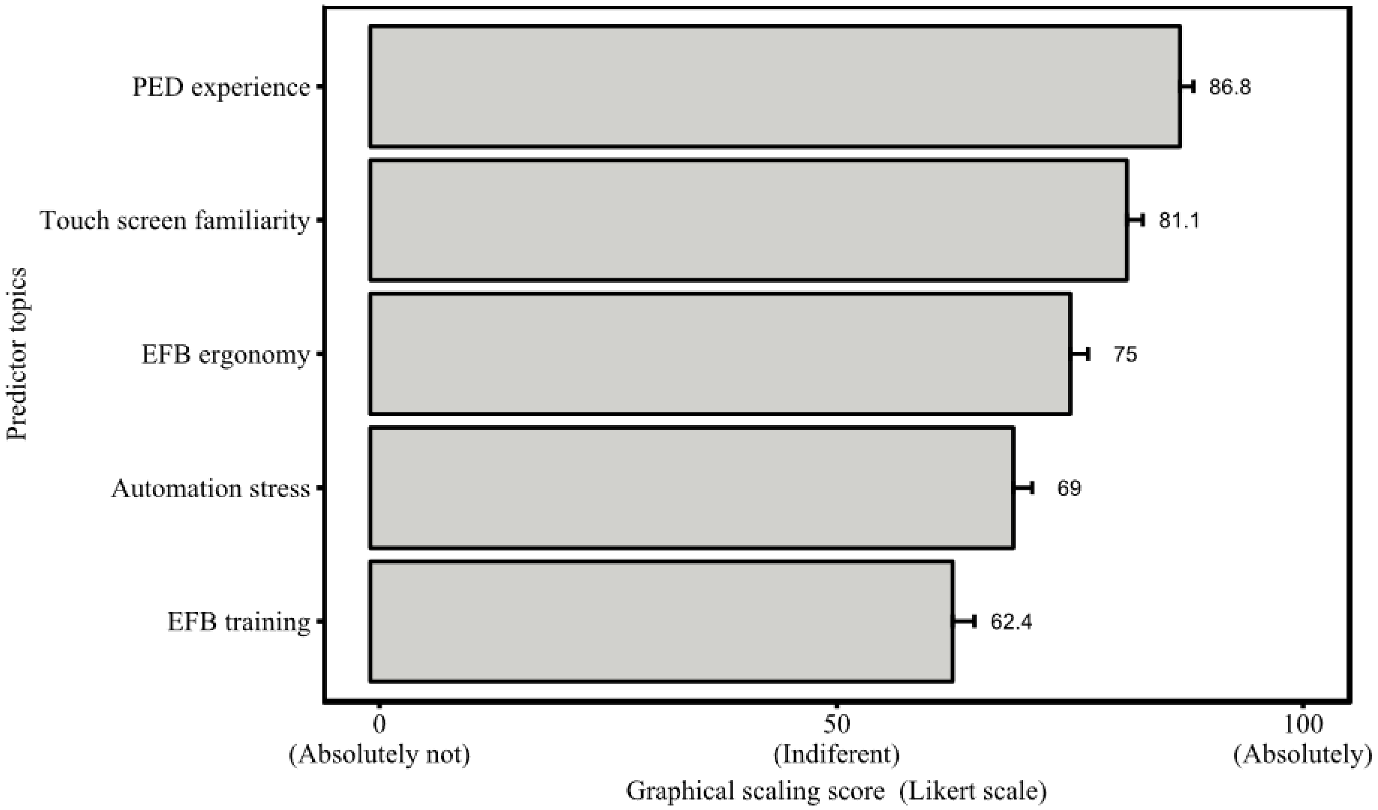


Figure S3. Questionnaire results for all continuous predictor variables. The bars represent the mean score of the graphical scaling score scale. 95% confidence intervals have been added at the end of each bar, along with the mean value. The number of responses (*n*) are derived from the outlier corrected, imputated total response (*N* = 571). Across all predictors, participants on average always to some extend agree with the statements presented to them.

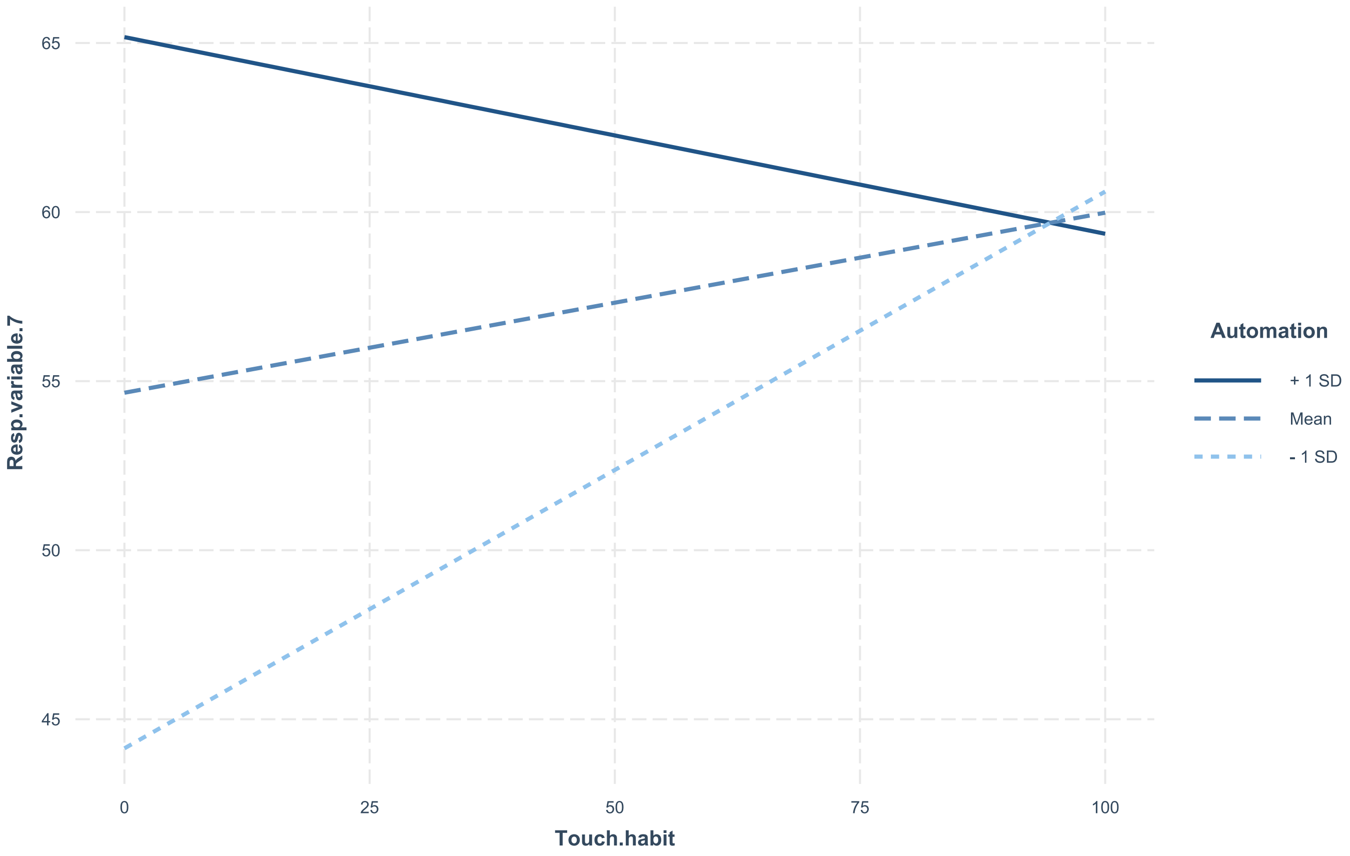


Figure S4. An example of an interaction between touch screen familiarity and one's attitudes towards society becoming more digitalized for statement response 1 (i.e., the access to information). The mean value on the graphical rating scale for the statement response estimated by the model is on the y-axis and the graphical rating scale for touch screen usage in on the x-axis. Depending on one's attitudes towards digitalization, the effect on the mean vary as the graphical rating score for touch screen familiarity vary. Those with a positive attitude do not change how they perceive the access of information to have changed since the implementation of EFBs even as their familiarity with touch screens vary. However, those with a negative or indifferent attitude tend to view an easier access to information as their familiarity with touch screens increase.

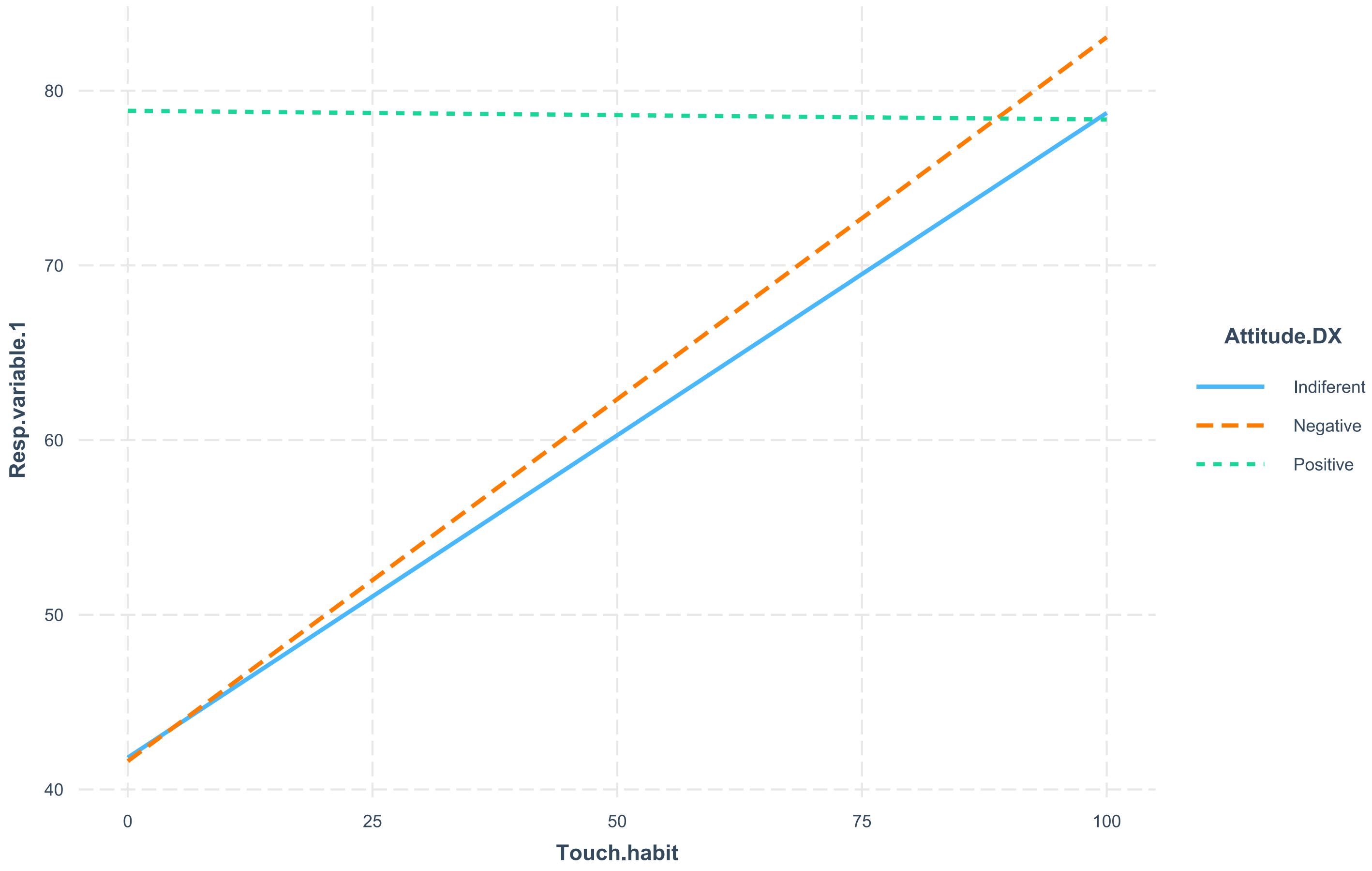


Figure S5. An example of an interaction between touch screen familiarity and the degree to which one becomes stressed by automation for statement response 7 (i.e., in-flight distraction by administrative duties). If one is “averagely” stressed by automation or blow averagely stressed, one feel less distracted by administrative duties in-flight as touch screen familiarity increase. However, if one is above averagely stressed by automation, one feel more distracted by similar duties as touch screen familiarity increase.

Appendix II: supplementary tables

(Intentionally left blank)

Table S1. Normal density distribution comparison table for all statement response variables ($N = 16$). The values are the graphical rating scale, with respective mean, confidence interval, skewness, kurtosis, and test statistics. The pooled, i.e., the total number of responses across all statements, is also included. As the Shapiro-Wilks Test is only appropriate for $n < 5000$, it was inappropriate for the pooled responses. All responses have a positive skewness with highly significant statistical tests. This is indicative of a positive shift in flight safety, as participant answers result in a skew to the right of mean, and with a non-normal distribution of responses.

Statement	**n**	**Mean**	**CI (+ -)†**	**Kurtosis**	**Skewness**	**K-S† With Lilliefors Correction Test**			**Shapiro-Wilk Test**		
						Statistics	**Df†**	***p*-value**	**Statistics**	**Df†**	***p*-value**
Pooled (*)	9152	70.90	0.54	2.80	-0.74	0.34	9152	<0.001*	NA	NA	NA
Information access (1)	572	84.14	1.73	6.09	-1.76	0.51	572	<0.001*	0.76	572	<0.001*
OP modifications (2)	572	70.71	2.24	2.59	-0.68	0.33	572	<0.001*	0.89	572	<0.001*
WX & NOTAMS (3)	572	77.68	2.06	3.69	-1.16	0.39	572	<0.001*	0.83	572	<0.001*
Performance calculations (4)	572	81.26	1.95	4.75	-1.48	0.46	572	<0.001*	0.78	572	<0.001*
Performance insight (5)	572	74.86	1.95	3.31	-0.87	0.39	572	<0.001*	0.88	572	<0.001*
Head down time (6)	572	56.86	2.35	2.22	-0.17	0.19	572	<0.001*	0.95	572	<0.001*
Admin, Duties (7)	572	57.38	2.06	2.73	-0.18	0.25	572	<0.001*	0.96	572	<0.001*
Situational awareness (8)	572	70.88	1.92	3.14	-0.69	0.38	572	<0.001*	0.93	572	<0.001*
Communication (9)	572	59.98	1.71	3.14	0.04	0.34	572	<0.001*	0.94	572	<0.001*
Logging (10)	572	60.30	2.29	2.37	-0.36	0.26	572	<0.001*	0.94	572	<0.001*
In-flight abnorm, (11)	572	77.51	1.98	3.43	-1.04	0.38	572	<0.001*	0.85	572	<0.001*
Ground abnorm, (12)	572	79.08	1.87	4.00	-1.17	0.41	572	<0.001*	0.84	572	<0.001*
Emergency (13)	572	69.44	2.00	3.35	-0.79	0.36	572	<0.001*	0.92	572	<0.001*
Access to OMs (14)	572	84.77	1.66	5.71	-1.68	0.52	572	<0.001*	0.77	572	<0.001*
OM revisions (15)	572	62.35	2.74	1.81	-0.41	0.23	572	<0.001*	0.89	572	<0.001*
Flight deck tasks (16)	572	33.19	1.90	3.07	0.39	0.39	572	<0.001*	0.93	572	<0.001*

*Pooled responses across all 16 statements

†Abbreviations: CI (+ -), confidence interval; K-S, Kolmogorov-Smirnov; Df, degree of freedom

Table S2. Mahalanobis distances (D^2) of between all pairs of operators, calculated from all 16 statement responses. Higher values are indicative of further statistical distance in multivariate space. Paraeneses show statistical reliability derived from 1000 round permutation test. All operator pairs show statistical significance, with participants working for ‘Bristow’ at the greatest distance to all other operators.

	Operator				
	Bristow	CHC	Norwegian	Other	SAS
CHC	2.86 (0.002)*		1.64 (0.002)*	1.50 (0.002)*	1.34 (0.002)*
Norwegian	2.80 (0.002)*	1.64 (0.002)*		1.89 (0.002)*	1.14 (0.002)*
Other	2.14 (0.002)*	1.50 (0.002)*	1.89 (0.002)*		1.71 (0.002)*
SAS	2.85 (0.002)*	1.34 (0.002)*	1.14 (0.002)*	1.71 (0.002)*	
Widerøe	2.80 (0.002)*	1.46 (0.002)*	1.35 (0.002)*	1.60 (0.002)*	0.76 (0.041)*

* *p*-value with α = .05. Significant values are denoted with Asterix

Table S3. MANOVA results for operator performed with all 16 statement responses as dependent variables. For categorical predictors, the test statistics compares within group variation to between group variation. Type-II calculation of sums of squares, and Pillai test-statistics were used. There is at least some response statement where there is a higher variation between operators compared to variation between participants within one or more operators

Predictor	**Df**[1]	**Λ**[2]	**F**[3]	**num. df**[4]	**den. df**[5]	***p***[6]
Operator	5	0.87	7.33	80	2775	<0.001*

1. degrees of freedom
2. Pillai-Bartlett trace
3. approximate F-value
4. numerator degrees of freedom for the F-value calculation
5. denominator degrees of freedom for the F-value calculation
6. *p*-value with α = .05. Significant values are denoted with Asterix (*)

Table S4. ANOVA results with “white adjustment” for each statement response (*N* = 16), with operator as predictor variable. The test statistics compares within group variation to between group variation. Larger F-values are indicative of greater between group variation. Significant ANOVAs indicate that at least some of the operators have different means. Participants relate to most statements differently based on where they work.

Statement response	**SS**[1]	**SS residuals**[1]	**F**[2]	***p***[3]
Information access (1)	8080.5	245333.1	3.7	0.012
OP modifications (2)	58725.3	369496.2	18.0	<0.001*
WX & NOTAMS (3)	16783.1	343479.1	5.5	<0.001*
Performance calculations (4)	31247.6	291847.2	12.1	<0.001*
Performance insight (5)	22877.1	301791.0	8.6	<0.001*
Head down time (6)	10371.9	458915.2	2.6	0.080
Admin, Duties (7)	10360.8	360417.7	3.3	0.026
Situational awareness (8)	1965.6	312143.4	0.7	0.817
Communication (9)	2170.0	242305.2	1.0	0.817
Logging (10)	37521.5	412853.1	10.3	<0.001*
In-flight abnorm. (11)	22579.9	313104.5	8.2	<0.001*
Ground abnorm. (12)	22948.9	276883.1	9.4	<0.001*
Emergency (13)	11423.9	335219.8	3.9	0.011
Access to OMs (14)	11136.7	222684.4	5.7	<0.001*
OM revisions (15)	31530.6	608474.3	5.9	<0.001*
Flight deck tasks (16)	50576.2	258410.7	22.2	<0.001*

1. type-II calculation of sums of squares (SS)

2 calculated from *F*-distribution

3. $\alpha = .01$, with Holms adjustment. Significant values are denoted with Asterix (*)

Table S5. Post-hoc Tukey test for all significant operator ANOVAs. The table shows all operator pairs for each statement response, with their mean values and confidence values, as well as a *p*-value deduced from a student t-test with Holms adjustment. All statement responses have at least some operator pairs with significantly different means

Response statement	**Operator pairs**	**CI-^{1}**	**Mean**	**CI+^{2}**	p^3
OP modifications (2)	CHC-Bristow	25.8	40.5	55.1	<0.001*
	Norwegian-Bristow	19.7	33.0	46.4	<0.001*
	Other-Bristow	1.2	15.1	29.0	0.025*
	SAS-Bristow	16.4	29.7	43.1	<0.001*
	Widerøe-Bristow	14.6	27.7	40.8	<0.001*
	Norwegian-CHC	-18.7	-7.4	3.8	0.411
	Other-CHC	-37.3	-25.3	-13.4	<0.001*
	SAS-CHC	-22.0	-10.7	0.6	0.074
	Widerøe-CHC	-23.7	-12.8	-1.8	0.012*
	Other-Norwegian	-28.2	-17.9	-7.7	<0.001*
	SAS-Norwegian	-12.7	-3.3	6.2	0.920
	Widerøe-Norwegian	-14.4	-5.3	3.7	0.542
	SAS-Other	4.4	14.6	24.9	0.001*
	Widerøe-Other	2.6	12.6	22.5	0.004*
	Widerøe-SAS	-11.2	-2.1	7.1	0.988
WX & NOTAMS (3)	CHC-Bristow	8.1	22.2	36.4	<0.001*
	Norwegian-Bristow	7.0	19.9	32.7	<0.001*
	Other-Bristow	0.7	14.1	27.5	0.033
	SAS-Bristow	7.8	20.6	33.5	<0.001*
	Widerøe-Bristow	6.3	18.9	31.6	<0.001*
	Norwegian-CHC	-13.2	-2.4	8.5	0.989
	Other-CHC	-19.6	-8.1	3.4	0.336
	SAS-CHC	-12.5	-1.6	9.3	0.998
	Widerøe-CHC	-13.9	-3.3	7.3	0.949
	Other-Norwegian	-15.6	-5.7	4.1	0.555
	SAS-Norwegian	-8.3	0.8	9.9	1.000
	Widerøe-Norwegian	-9.7	-0.9	7.8	1.000
	SAS-Other	-3.4	6.5	16.5	0.414
	Widerøe-Other	-4.8	4.8	14.4	0.702
	Widerøe-SAS	-10.5	-1.7	7.1	0.994
Performance calculations (4)	CHC-Bristow	-9.5	3.5	16.6	0.972
	Norwegian-Bristow	1.7	13.6	25.4	0.014
	Other-Bristow	-22.0	-9.6	2.8	0.234
	SAS-Bristow	-2.3	9.6	21.5	0.192
	Widerøe-Bristow	-6.7	5.0	16.6	0.827

Cont.

	Norwegian-CHC	0.1	10.1	20.0	0.047*
	Other-CHC	-23.7	-13.1	-2.5	0.006*
	SAS-CHC	-4.0	6.1	16.1	0.515
	Widerøe-CHC	-8.3	1.4	11.2	0.998
	Other-Norwegian	-32.3	-23.2	-14.1	<0.001*
	SAS-Norwegian	-12.4	-4.0	4.4	0.751
	Widerøe-Norwegian	-16.7	-8.6	-0.6	0.028*
	SAS-Other	10.0	19.2	28.3	<0.001*
	Widerøe-Other	5.7	14.5	23.4	<0.001*
	Widerøe-SAS	-12.7	-4.6	3.5	0.580
Performance insight (5)	CHC-Bristow	-13.5	-0.2	13.0	1.000
	Norwegian-Bristow	-6.7	5.3	17.4	0.805
	Other-Bristow	-23.2	-10.6	2.0	0.155
	SAS-Bristow	-2.5	9.6	21.6	0.211
	Widerøe-Bristow	-8.2	3.6	15.5	0.951
	Norwegian-CHC	-4.6	5.5	15.7	0.624
	Other-CHC	-21.2	-10.4	0.4	0.068
	SAS-CHC	-0.4	9.8	20.0	0.068
	Widerøe-CHC	-6.0	3.9	13.8	0.873
	Other-Norwegian	-25.2	-15.9	-6.7	<0.001*
	SAS-Norwegian	-4.3	4.2	12.8	0.713
	Widerøe-Norwegian	-9.9	-1.7	6.5	0.992
	SAS-Other	10.9	20.2	29.5	<0.001*
	Widerøe-Other	5.3	14.3	23.2	<0.001*
	Widerøe-SAS	-14.2	-5.9	2.3	0.315
Logging (10)	CHC-Bristow	20.9	36.4	52.0	<0.001*
	Norwegian-Bristow	7.6	21.7	35.8	<0.001*
	Other-Bristow	0.4	15.1	29.8	0.041*
	SAS-Bristow	10.2	24.3	38.4	<0.001*
	Widerøe-Bristow	7.5	21.4	35.2	<0.001*
	Norwegian-CHC	-26.6	-14.7	-2.8	0.006*
	Other-CHC	-34.0	-21.3	-8.7	<0.001*
	SAS-CHC	-24.1	-12.1	-0.2	0.044*
	Widerøe-CHC	-26.7	-15.1	-3.5	0.003*
	Other-Norwegian	-17.5	-6.6	4.2	0.495
	SAS-Norwegian	-7.4	2.6	12.6	0.977
	Widerøe-Norwegian	-9.9	-0.4	9.2	1.000
	SAS-Other	-1.7	9.2	20.1	0.150
	Widerøe-Other	-4.2	6.3	16.8	0.525
	Widerøe-SAS	-12.6	-2.9	6.7	0.954
In-flight abnorm, (11)	CHC-Bristow	-10.0	3.6	17.1	0.975
	Norwegian-Bristow	-21.6	-9.3	2.9	0.250
	Other-Bristow	-21.9	-9.1	3.7	0.326
	SAS-Bristow	-10.0	2.3	14.6	0.995
	Widerøe-Bristow	-6.7	5.4	17.4	0.798

Cont.

	Norwegian-CHC	-23.2	-12.9	-2.6	0.005*
	Other-CHC	-23.7	-12.7	-1.7	0.013*
	SAS-CHC	-11.6	-1.2	9.2	0.999
	Widerøe-CHC	-8.3	1.8	11.9	0.995
	Other-Norwegian	-9.2	0.2	9.7	1.000
	SAS-Norwegian	3.0	11.7	20.4	0.002*
	Widerøe-Norwegian	6.4	14.7	23.1	<0.001*
	SAS-Other	2.0	11.4	20.9	0.008*
	Widerøe-Other	5.3	14.5	23.6	<0.001*
	Widerøe-SAS	-5.3	3.1	11.5	0.904
Ground abnorm. (12)	CHC-Bristow	-5.1	7.7	20.4	0.518
	Norwegian-Bristow	-22.0	-10.5	1.0	0.099
	Other-Bristow	-16.1	-4.1	8.0	0.930
	SAS-Bristow	-7.8	3.7	15.3	0.941
	Widerøe-Bristow	-6.9	4.4	15.8	0.874
	Norwegian-CHC	-27.9	-18.1	-8.4	<0.001*
	Other-CHC	-22.1	-11.7	-1.4	0.016*
	SAS-CHC	-13.7	-3.9	5.9	0.861
	Widerøe-CHC	-12.7	-3.2	6.3	0.927
	Other-Norwegian	-2.4	6.4	15.3	0.301
	SAS-Norwegian	6.0	14.2	22.4	<0.001*
	Widerøe-Norwegian	7.1	14.9	22.8	<0.001*
	SAS-Other	-1.1	7.8	16.7	0.126
	Widerøe-Other	-0.1	8.5	17.1	0.056
	Widerøe-SAS	-7.2	0.7	8.6	1.000
Access to OMs (14)	Norwegian-Bristow	-20.0	-9.7	0.6	0.080
	Other-Bristow	-23.0	-12.1	-1.3	0.018*
	SAS-Bristow	-16.5	-6.2	4.2	0.532
	Widerøe-Bristow	-15.0	-4.9	5.3	0.747
	Norwegian-CHC	-20.8	-12.1	-3.3	0.001*
	Other-CHC	-23.8	-14.5	-5.2	<0.001*
	SAS-CHC	-17.3	-8.5	0.2	0.061
	Widerøe-CHC	-15.7	-7.2	1.3	0.149
	Other-Norwegian	-10.4	-2.4	5.5	0.952
	SAS-Norwegian	-3.8	3.5	10.9	0.742
	Widerøe-Norwegian	-2.2	4.8	11.9	0.363
	SAS-Other	-2.0	6.0	14.0	0.270
	Widerøe-Other	-0.4	7.3	15.0	0.077
	Widerøe-SAS	-5.8	1.3	8.4	0.995
OM revisions (15)	CHC-Bristow	-33.1	-14.2	4.6	0.260
	Norwegian-Bristow	-42.9	-25.8	-8.7	<0.001*
	Other-Bristow	-31.5	-13.6	4.3	0.251
	SAS-Bristow	-44.2	-27.0	-9.9	<0.001*
	Widerøe-Bristow	-36.1	-19.3	-2.5	0.014*
	Norwegian-CHC	-26.0	-11.5	2.9	0.200

Cont.

	Other-CHC	-14.7	0.6	15.9	1.000
	SAS-CHC	-27.3	-12.8	1.7	0.117
	Widerøe-CHC	-19.1	-5.1	9.0	0.907
	Other-Norwegian	-1.0	12.2	25.3	0.088
	SAS-Norwegian	-13.4	-1.3	10.9	1.000
	Widerøe-Norwegian	-5.2	6.5	18.1	0.606
	SAS-Other	-26.6	-13.4	-0.2	0.044*
	Widerøe-Other	-18.4	-5.7	7.1	0.798
	Widerøe-SAS	-4.0	7.7	19.5	0.410
Flight deck tasks (16)	CHC-Bristow	-20.4	-8.1	4.2	0.414
	Norwegian-Bristow	-34.7	-23.6	-12.4	<0.001*
	Other-Bristow	-19.7	-8.1	3.6	0.354
	SAS-Bristow	-33.3	-22.1	-11.0	<0.001*
	Widerøe-Bristow	-40.0	-29.1	-18.1	<0.001*
	Norwegian-CHC	-24.9	-15.5	-6.1	<0.001*
	Other-CHC	-10.0	0.0	10.0	1.000
	SAS-CHC	-23.5	-14.0	-4.6	<0.001*
	Widerøe-CHC	-30.2	-21.0	-11.8	<0.001*
	Other-Norwegian	6.9	15.5	24.1	<0.001*
	SAS-Norwegian	-6.5	1.4	9.3	<0.001*
	Widerøe-Norwegian	-13.1	-5.5	2.1	0.302
	SAS-Other	-22.7	-14.1	-5.4	<0.001*
	Widerøe-Other	-29.3	-21.0	-12.7	<0.001*
	Widerøe-SAS	-14.6	-6.9	0.7	0.098

1, 2. Lower and upper 95% confidence values
3. α = .05, with Holms adjustment. Significant values are denoted with Asterix (*)

Table S6. Model confidence interval for the top multivariate models following a stepwise model selection scheme. All models within an AIC$_c$ of 2 from the best model is shown. Each row represents one model, with plus signs marking the predictors included. The green band marks the best model.

Predictor											
Age	Operator	Touch screen familiarity	Digitalization attitude	EFB ergonomy	Automation stress	EFB training	Flight deck position	Touch[…]:Digitalization[…]	Touch[…]:Automation[…]	EFB erg[…]:Automation[…]	AIC$_c$
+	+	+	+	+	+	+	+	+	+		79295.5
+	+	+	+	+	+	+	+	+	+	+	79296.9

Table S7. An approximated F-test using Pillai-Bartlett trace statistics between the best multivariate model against an intercept only model. The best model is significantly different from the intercept only model, with a loss of 17 degrees of freedom.

Model	**Df**[1]	**ΔDf**[2]	**Λ**[3]	**F**[4]	**Df**[5]	**Df**[6]	***p***[7]
Intercept	571	-	-	-	-	-	-
Best	554	-17	1.73	3.96	272	8864	<0.001*

1. degrees of freedom
2. difference in Df between the two models
3. Pillai-Bartlett trace
4. approximate F-value
5. numerator degrees of freedom for the F-value calculation
6. denominator degrees of freedom for the F-value calculation
7. *p*-value with $\alpha = .05$. Significant values are denoted with Asterix (*)

Table S8. MANOVA results with “white adjustment” on each predictor variable in the best multivariate model using all 16 responses as dependent variables. For categorical predictors, the test statistics compares within group variation to between group variation. For continuous variables, the test statistics checks if the regression slope is significantly different from zero. The *p*-values are calculated using type-III sums of squares calculation with Pillai-test statistics. The two interactions are shown in the last two rows of the table, where the respective predictor has been shorthanded with ‘[…]’. The interaction term is denoted with colon (**:**).

Predictor	**Df**[1]	**Λ**[2]	**F**[3]	**num. df**[4]	**den. df**[5]	***p***[6]
Age*	2	0.12	2.05	32	1080	<0.001*
Operator*	5	0.87	7.13	80	2715	<0.001*
EFB ergonomy*	1	0.07	2.51	16	539	<0.001*
EFB training *	1	0.09	3.16	16	539	<0.001*
Flight deck position*	1	0.06	2.26	16	539	0.003*
Touch screen familiarity*	1	0.10	3.57	16	539	<0.001*
Digitalization attitude*	2	0.14	2.61	32	1080	<0.001*
Automation stress*	1	0.08	2.88	16	539	<0.001*
Touch[…] **:** Digitalization[…]*	2	0.12	2.23	32	1080	<0.001*
Touch[…] **:** Automation[…]*	1	0.07	2.51	16	539	0.001*

1. degrees of freedom
2. Pillai-Bartlett trace
3. approximate F-value
4. numerator degrees of freedom for the F-value calculation
5. denominator degrees of freedom for the F-value calculation
6. *p*-value with $\alpha = .05$. Significant values are denoted with Asterix (*)

Table S9. Model confidence interval for the top univariate models following a stepwise model selection scheme. All models within an AIC_c of 2 from the best model is shown. Each row represents one model, with plus signs marking the predictors included. The green band marks the best model.

Statement response	**Predictor**										**adjusted R^2**	**AICc**
	Age	Digitalization attitude	Automation stress	EFB ergonomy	Flight deck position	Operator	Touch screen familiarity	EFB training	Touch[…]:Digitalization[…]	Touch[…]:Automation[…]		
1	**-**	**+**	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**0.31**	**4926.31**
1	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**0.31**	**4927.38**
1	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**0.31**	**4927.44**
2	-	+	+	+	+	+	-	+	-	-	0.26	5257.69
2	-	+	+	+	+	+	+	+	-	-	0.26	5257.91
2	-	+	-	+	+	+	+	+	-	-	0.26	5258.10
2	-	+	-	+	+	+	+	+	+	-	0.27	5258.24
2	-	+	+	+	+	+	+	+	+	-	0.27	5258.32
2	**-**	**+**	**-**	**+**	**+**	**+**	**-**	**+**	**-**	**-**	**0.26**	**5258.42**
2	-	+	+	+	-	+	-	+	-	-	0.26	5258.97
2	-	+	+	+	-	+	+	+	-	-	0.26	5259.22
2	-	+	+	+	+	+	+	+	-	+	0.26	5259.37
2	+	+	+	+	+	+	-	+	-	-	0.26	5259.46
2	-	+	-	+	-	+	+	+	-	-	0.26	5259.50
2	+	+	+	+	+	+	+	+	-	-	0.27	5259.61
2	+	+	-	+	+	+	+	+	-	-	0.26	5259.66
3	+	+	+	+	+	+	+	+	+	-	0.27	5169.01
3	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**0.26**	**5169.58**
4	**-**	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**+**	**0.21**	**5142.57**
4	+	-	+	+	+	+	+	+	-	+	0.21	5143.84
5	-	+	+	+	+	+	+	-	+	+	0.22	5138.36
5	-	+	+	+	+	+	+	+	+	+	0.22	5139.38
5	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**+**	**-**	**0.22**	**5139.52**
6	+	-	+	+	-	+	-	+	-	-	0.10	5421.03
6	+	+	+	+	-	+	-	+	-	-	0.11	5421.23
6	+	+	+	+	-	+	+	+	+	-	0.12	5421.48
6	-	+	+	+	-	+	-	+	-	-	0.10	5421.91
6	-	+	+	+	-	+	+	+	+	-	0.11	5421.95
6	**-**	**-**	**+**	**+**	**-**	**+**	**-**	**+**	**-**	**-**	**0.10**	**5421.96**

Cont.

6	+	-	+	+	+	+	-	+	-	-	0.10	5422.67
6	+	-	+	+	-	+	+	+	-	-	0.10	5422.87
6	+	+	+	+	+	+	-	+	-	-	0.11	5422.94
7	+	-	+	+	-	+	+	+	-	+	0.16	5235.13
7	+	+	+	+	-	+	+	+	-	+	0.17	5235.16
7	+	-	+	+	+	+	+	+	-	+	0.16	5236.85
7	-	+	+	+	-	+	+	+	-	+	0.16	5236.92
7	+	+	+	+	+	+	+	+	-	+	0.17	5236.95
7	**-**	**-**	**+**	**+**	**-**	**+**	**+**	**+**	**-**	**+**	**0.15**	**5237.01**
8	+	-	+	+	+	-	+	+	-	+	0.24	5089.00
8	+	+	+	+	+	-	+	+	-	+	0.25	5089.09
8	+	+	+	+	+	-	+	-	-	+	0.25	5089.15
8	+	+	+	+	+	-	+	+	+	+	0.25	5089.65
8	+	-	+	+	+	-	+	-	-	+	0.24	5089.69
8	+	+	+	+	+	-	+	-	+	+	0.25	5089.79
8	-	+	+	+	+	-	+	+	-	+	0.24	5090.76
8	**-**	**-**	**+**	**+**	**+**	**-**	**+**	**+**	**-**	**+**	**0.24**	**5090.99**
9	**-**	**+**	**+**	**+**	**-**	**+**	**+**	**+**	**-**	**+**	**0.14**	**5035.66**
9	-	+	+	+	-	+	+	+	+	+	0.15	5036.55
9	+	+	+	+	-	+	+	+	-	+	0.15	5036.70
9	-	+	+	+	+	+	+	+	-	+	0.14	5037.24
9	+	+	+	+	-	+	+	+	+	+	0.15	5037.49
10	**-**	**-**	**+**	**+**	**-**	**+**	**+**	**-**	**-**	**-**	**0.20**	**5319.66**
10	-	-	+	+	-	+	+	+	-	-	0.21	5320.10
10	-	-	+	+	+	+	+	-	-	-	0.20	5321.64
11	-	+	-	+	+	+	+	+	+	-	0.22	5152.13
11	-	+	-	+	-	+	+	+	+	-	0.22	5152.88
11	**-**	**+**	**-**	**+**	**+**	**+**	**-**	**+**	**-**	**-**	**0.21**	**5153.78**
11	-	+	+	+	+	+	+	+	+	-	0.22	5154.08
12	**-**	**+**	**-**	**+**	**-**	**+**	**+**	**+**	**+**	**-**	**0.26**	**5053.89**
12	-	+	+	+	-	+	+	+	+	-	0.26	5055.58
12	-	+	-	+	+	+	+	+	+	-	0.26	5055.86
13	**-**	**+**	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**0.20**	**5176.97**
13	-	+	+	+	+	+	+	+	+	-	0.21	5177.23
13	-	+	+	+	+	+	+	+	+	+	0.21	5177.34
13	+	+	-	+	+	+	+	+	+	-	0.21	5177.42
13	+	+	+	+	+	+	+	+	+	-	0.21	5178.03
13	+	+	+	+	+	+	+	+	+	+	0.21	5178.14
14	**-**	**+**	**-**	**+**	**+**	**+**	**+**	**+**	**+**	**-**	**0.22**	**4946.67**
15	-	+	-	+	-	+	+	+	+	-	0.17	5556.21
15	**-**	**+**	**-**	**-**	**-**	**+**	**+**	**+**	**+**	**-**	**0.16**	**5557.19**
15	-	+	+	+	-	+	+	+	+	-	0.17	5557.31
15	+	+	-	+	-	+	+	+	+	-	0.17	5557.60
15	-	+	-	+	+	+	+	+	+	-	0.17	5557.90
15	-	+	+	-	-	+	+	+	+	-	0.17	5557.96
16	+	-	+	-	+	+	+	+	-	+	0.24	5091.22
16	+	-	+	+	+	+	+	+	-	+	0.24	5092.58
16	**+**	**-**	**+**	**-**	**-**	**+**	**+**	**+**	**-**	**+**	**0.23**	**5092.94**

Table S10. A null-model test using F-test statistics between the best univariate model derived from the multivariate model space, against a respective intercept only model. Models for all 16 statement responses are listed. All models show significant deviations from an intercept only model, indicating that the predictors are contributing to explaining the variation.

Statement response	Model	ΔDf[1]†	SS† residuals	ΔSS[2]	F[3]	*p*[4]
Flight planning (1)	Intercept		174111			
	Best	-13	253414	-79303	19.55	<0.001*
Operative flight plan (2)	Intercept		319644			
	Best	-9	428221	-108577	21.21	<0.001*
WX & NOTAMs (3)	Intercept		265913			
	Best	-14	360262	-94349	14.12	<0.001*
Performance calculations (4)	Intercept		255970			
	Best	-11	323095	-67124	13.35	<0.001*
Performance insight (5)	Intercept		256421			
	Best	-12	324668	-68247	12.40	<0.001*
Head-down time (6)	Intercept		423079			
	Best	-8	469287	-46208	7.69	<0.001*
In-flight distraction (7)	Intercept		317469			
	Best	-10	370778	-53310	9.42	<0.001*
Situational awareness (8)	Intercept		237601			
	Best	-7	314109	-76508	25.94	<0.001*
Closed-loop com. (9)	Intercept		210417			
	Best	-11	244475	-34059	8.24	<0.001*
In-flight logging (10)	Intercept		364239			
	Best	-8	450375	-86136	16.64	<0.001*
In-flight abnormalities (11)	Intercept		263354			
	Best	-9	335684	-72331	17.15	<0.001*
Pre-flight abnormalities (12)	Intercept		219613			
	Best	-12	299832	-80219	17.02	<0.001*
Emergency (13)	Intercept		276491			
	Best	-13	346644	-70153	10.89	<0.001*
OM† access (14)	Intercept		181118			
	Best	-13	233821	-52703	12.49	<0.001*
OM† revisions (15)	Intercept		533383			
	Best	-11	640005	-106621	10.18	<0.001*
Flight deck workload (16)	Intercept		238664			
	Best	-11	308987	-70323	15.00	<0.001*

† Abbreviations: Df, degree of freedom; SS, sums of squares; OM, operations manual

1. difference in degrees of freedom between intercept model and best model
2. difference in sums of squares between intercept model and best model
3. calculated from F-distribution
4. *p*-value with α = .01. Significant values are denoted with Asterix (*)

Table S11. ANOVA results with “white adjustment” for each predictor over all the best univariate models (*N* = 16). For categorical predictors, the test statistics compares within group variation to between group variation. For continuous variables, the test statistics checks if the regression slope is significantly different from zero. Type-III sums of squares calculation was used. The two interactions have the respective predictor shorthanded with ‘[…]’. The interaction term is denoted with ‘**:**’.

Response	**Predictor**	**Df**†	**F**¹	***p***²
Situational awareness (8)	Age	2	3.6	0.029
Flight deck workload (16)		2	5.4	0.009*
Flight planning (1)	Digitalization Attitude	2	10.3	<0.001*
Operative flight plan (2)		2	7.2	0.007*
WX & NOTAMs (3)		2	6.2	0.011
Performance insight (5)		2	3.2	0.055
Closed-loop com. (9)		2	3.6	0.055
In-flight abnormalities (11)		2	5.2	0.017
Pre-flight abnormalities (12)		2	10.1	<0.001*
Emergency (13)		2	6.7	0.011
OM† access (14)		2	5.9	0.012
OM† revisions (15)		2	6.6	0.011
Flight planning (1)	Digitalization[…] **:** Touch[…]	2	8.1	0.002*
WX & NOTAMs (3)		2	4.1	0.042
Pre-flight abnormalities (12)		2	6.0	0.013
Emergency (13)		2	4.3	0.042
OM† access (14)		2	5.4	0.019
OM† revisions (15)		2	3.6	0.042
WX & NOTAMs (3)	Automation stress	1	4.0	0.089
Performance calculations (4)		1	7.6	0.031
Performance insight (5)		1	11.9	0.005*
Head-down time (6)		1	4.1	0.089
In-flight distraction (7)		1	7.5	0.031
Situational awareness (8)		1	17.6	<0.001*
Closed-loop com. (9)		1	9.4	0.016
In-flight logging (10)		1	7.9	0.031
Flight deck workload (16)		1	7.2	0.031
Performance calculations (4)	Automation[…] **:** Touch[…]	1	5.8	0.044
Performance insight (5)		1	6.8	0.044
In-flight distraction (7)		1	6.2	0.044
Situational awareness (8)		1	12.9	0.002*
Closed-loop com. (9)		1	6.9	0.044
Flight deck workload (16)		1	4.8	0.044
Flight planning (1)	EFB ergonomy	1	8.5	0.034
Operative flight plan (2)		1	13.1	0.004*
WX & NOTAMs (3)		1	16.1	<0.001*
Performance calculations (4)		1	5.4	0.041
Performance insight (5)		1	7.2	0.041
Head-down time (6)		1	6.9	0.041
In-flight distraction (7)		1	6.2	0.041

Cont.

Situational awareness (8)		1	25.4	<0.001*
Closed-loop com. (9)		1	10.5	0.014
In-flight logging (10)		1	9.2	0.026
In-flight abnormalities (11)		1	8.4	0.034
Pre-flight abnormalities (12)		1	3.3	0.071
Emergency (13)		1	7.8	0.039
OM† access (14)		1	7.4	0.041
Flight planning (1)	Flight deck position	1	12.2	0.003*
WX & NOTAMs (3)		1	16.0	<0.001*
Performance calculations (4)		1	17.9	<0.001*
Performance insight (5)		1	9.4	0.007*
Situational awareness (8)		1	8.8	0.007*
Emergency (13)		1	10.3	0.006*
OM† access (14)		1	5.6	0.018
Flight planning (1)	Operator	5	2.2	0.085
Operative flight plan (2)		5	19.9	<0.001*
WX & NOTAMs (3)		5	5.6	<0.001*
Performance calculations (4)		5	14.3	<0.001*
Performance insight (5)		5	7.9	<0.001*
Head-down time (6)		5	5.0	<0.001*
In-flight distraction (7)		5	6.0	<0.001*
Closed-loop com. (9)		5	2.3	0.085
In-flight logging (10)		5	10.6	<0.001*
In-flight abnormalities (11)		5	3.7	0.010*
Pre-flight abnormalities (12)		5	6.3	<0.001*
Emergency (13)		5	3.6	0.010*
OM† access (14)		5	7.0	<0.001*
OM† revisions (15)		5	8.1	<0.001*
Flight deck workload (16)		5	26.2	<0.001*
Flight planning (1)	Touch screen familiarity	1	16.1	<0.001*
WX & NOTAMs (3)		1	.7	0.518
Performance calculations (4)		1	3.7	0.268
Performance insight (5)		1	6.7	0.087
In-flight distraction (7)		1	7.7	0.063
Situational awareness (8)		1	16.7	<0.001*
Closed-loop com. (9)		1	6.7	0.087
In-flight logging (10)		1	3.7	0.268
Pre-flight abnormalities (12)		1	3.1	0.268
Emergency (13)		1	4.7	0.211
OM† access (14)		1	1.3	0.518
OM† revisions (15)		1	4.6	0.211
Flight deck work load (16)		1	7.5	0.064
Flight planning (1)	EFB training	1	21.9	<0.001*
Operative flight plan (2)		1	7.9	0.031
WX & NOTAMs (3)		1	5.2	0.091
Performance calculations (4)		1	5.9	0.077
Head-down time (6)		1	4.7	0.091
In-flight distraction (7)		1	5.1	0.091

Cont.

In-flight abnormalities (11)	1	22.8	<0.001*
Pre-flight abnormalities (12)	1	17.1	<0.001*
Emergency (13)	1	9.9	0.014
OM† access (14)	1	9.2	0.018
OM† revisions (15)	1	22.8	<0.001*
Flight deck workload (16)	1	4.7	0.091

† Abbreviations: Df, degree of freedom; OM, operations manual

1. calculated from F-distribution.

2. Holms adjusted *p*-values with $\alpha = .01$. Significant values are denoted with Asterix (*)

Table S12. Summary of all the best univariate models (*N* = 16) with an emphasis on the importance of standardized regression coefficients (β). The regression coefficient of continuous variables is standardized by two times their standard deviation, so the value is equivalent of going from one extreme to another. This makes for easier inference when comparing with categorical predictors

Model	**R^2[1]**	**Predictor**	**Group**	**β[2]**	***p*[3]**
Flight planning (1)	0.31	(Intercept)†	NA	75.3	<0.001*
		Digitalization attitude	Negative	3.5	0.983
			Positive	6.7	<0.001*
		EFB ergonomy	NA	7.6	<0.001*
		Flight deck position	Commander	5.4	<0.001*
		Operator	Bristow	- 0.6	0.849
			CHC	7.4	0.013*
			Norwegian	7.1	0.005*
			SAS	6.9	0.009*
			Widerøe	4.2	0.090*
		Touch screen familiarity	NA	15.2	<0.001*
		EFB training	NA	9.	<0.001*
		Digitalization[…]:Touch[…]	Negative	1.9	0.739
			Positive	-15.4	<0.001*
Operative flight plan (2)	0.26	(Intercept)†	NA	52.5	<0.001*
		Digitalization attitude	Negative	7.8	0.137
			Positive	10.6	<0.001*
		EFB ergonomy	NA	10.3	<0.001*
		Operator	Bristow	-17.1	<0.001*
			CHC	23.	<0.001*
			Norwegian	19.	<0.001*
			SAS	13.8	<0.001*
			Widerøe	7.5	0.025*
		EFB training	NA	7.1	0.005*
WX & NOTAMs (3)	0.26	(Intercept)†	NA	70.6	<0.001*
		Digitalization attitude	Negative	4.7	0.162
			Positive	9.	<0.001*
		Automation stress	NA	4.7	0.019*
		EFB ergonomy	NA	12.9	<0.001*
		Flight deck position	Commander	7.6	<0.001*
		Operator	Bristow	-17.8	<0.001*
			CHC	3.6	0.328
			Norwegian	5.7	0.070
			SAS	4.5	0.165
			Widerøe	- 0.9	0.760
		Touch screen familiarity	NA	4.1	0.271
		EFB training	NA	5.8	0.014*
		Digitalization[…]:Touch[…]	Negative	10.9	0.116

Cont.

			Positive	- 9.7	0.024*
Performance calculations (4)	0.21	(Intercept)†	NA	68.6	0.089
		Automation stress	NA	4.9	<0.001*
		EFB ergonomy	NA	7.2	0.006*
		Flight deck position	Commander	7.9	<0.001*
		Operator	Bristow	8.	0.052
			CHC	10.7	0.003*
			Norwegian	23.9	<0.001*
			SAS	18.6	<0.001*
			Widerøe	10.9	<0.001*
		Touch screen familiarity	NA	- 1.6	0.009*
		EFB training	NA	5.8	0.012*
		Automation[…]:Touch[…]	NA	- 9.6	0.002*
Performance insight (5)	0.21	(Intercept)†	NA	60.7	0.559
		Digitalization attitude	Negative	- 0.8	0.865
			Positive	5.4	0.019*
		Automation stress	NA	7.2	<0.001*
		EFB ergonomy	NA	7.9	0.001*
		Flight deck position	Commander	5.4	0.003*
		Operator	Bristow	8.5	0.039*
			CHC	7.0	0.049*
			Norwegian	15.2	<0.001*
			SAS	17.3	<0.001*
			Widerøe	10.6	<0.001*
		Touch screen familiarity	NA	1.6	0.002*
		Automation[…]:Touch[…]	NA	- 9.3	0.003*
Head-down time (6)	0.10	(Intercept)†	NA	57.2	<0.001*
		Automation stress	NA	5.4	0.029*
		EFB ergonomy	NA	8.8	0.002*
		Operator	Bristow	7.8	0.137
			CHC	1.8	0.691
			Norwegian	6.9	0.074
			SAS	- 3.0	0.451
			Widerøe	- 8.4	0.02*
		EFB training	NA	7.2	0.015*
In-flight distraction (7)	0. 14	(Intercept)†	NA	59.0	0.666
		Automation stress	NA	3.0	0.004*
		EFB ergonomy	NA	8.6	0.003*
		Operator	Bristow	8.6	0.060
			CHC	5.7	0.152
			Norwegian	2.8	0.402
			SAS	- 1.5	0.666
			Widerøe	- 8.8	0.008*
		Touch screen familiarity	NA	2.2	0.003*
		EFB training	NA	6.3	0.014*
		Automation[…]:Touch[…]	NA	- 9.2	0.009*

Cont.

Situational awareness (8)	0.24	(Intercept)†	NA	71.1	0.692
		Age	over 55 YOA	- 3.9	0.122
			under 40 YOA	3.9	0.069
		Automation stress	NA	6.3	<0.001*
		EFB ergonomy	NA	14.7	<0.001*
		Flight deck position	Commander	6.0	0.003*
		Touch screen familiarity	NA	2.6	<0.001*
		Automation[…]:Touch[…]	NA	-11.6	<0.001*
Closed-loop com. (9)	0.14	(Intercept)†	NA	56.1	0.209
		Digitalization attitude	Negative	6.8	0.109
			Positive	5.3	0.011*
		Automation stress	NA	4.1	<0.001*
		EFB ergonomy	NA	9.1	<0.001*
		Operator	Bristow	2.0	0.592
			CHC	0.7	0.835
			Norwegian	4.6	0.094
			SAS	- 1.3	0.643
			Widerøe	- 3.3	0.214
		Touch screen familiarity	NA	0.4	0.002*
		Automation[…]:Touch[…]	NA	- 8.9	0.002*
In-flight logging (10)	0.19	(Intercept)†	NA	57.7	<0.001*
		Automation stress	NA	7.3	0.001*
		EFB ergonomy	NA	10.2	<0.001*
		Operator	Bristow	-17.0	<0.001*
			CHC	17.1	<0.001*
			Norwegian	5.8	0.108
			SAS	4.8	0.183
			Widerøe	1.7	0.623
		Touch screen familiarity	NA	6.3	0.030*
In-flight abnormalities (11)	0.22	(Intercept)†	NA	67.9	<0.001*
		Digitalization attitude	Negative	- 8.9	0.061
			Positive	5.1	0.027*
		EFB ergonomy	NA	7.5	<0.001*
		Operator	Bristow	6.1	0.142
			CHC	10.3	0.004*
			Norwegian	1.4	0.643
			SAS	11.1	<0.001*
			Widerøe	8.0	0.008*
		EFB training	NA	11.6	<0.001*
Pre-flight abnormalities (12)	0.27	(Intercept)†	NA	75.1	<0.001*
		Digitalization attitude	Negative	- 4.2	0.004*
			Positive	4.9	0.020*
		EFB ergonomy	NA	5.5	0.025*
		Operator	Bristow	0.0	0.990
			CHC	8.3	0.013*

Cont.

			Norwegian	- 6.3	0.027*
			SAS	5.7	0.052
			Widerøe	1.3	0.629
		Touch screen familiarity	NA	7.4	0.030*
		EFB training	NA	9.5	<0.001*
		Digitalization[…]:Touch[…]	Negative	13.9	0.029*
			Positive	- 6.4	0.098
Emergency (13)	0.20	(Intercept)†	NA	58.4	0.144
		Digitalization attitude	Negative	- 7.1	0.812
			Positive	5.8	<0.001*
		EFB ergonomy	NA	9.3	<0.001*
		Flight deck position	Commander	6.0	0.002*
		Operator	Bristow	6.7	0.117
			CHC	9.3	0.013*
			Norwegian	13.1	<0.001*
			SAS	11.7	<0.001*
			Widerøe	6.2	0.049*
		Touch screen familiarity	NA	10.4	0.006*
		EFB training	NA	7.5	0.002*
		Digitalization[…]:Touch[…]	Negative	- 2.1	0.763
			Positive	-14.7	<0.001*
OM† access (14)	0.23	(Intercept)†	NA	78.5	<0.001*
		Digitalization attitude	Negative	5.0	0.002*
			Positive	4.6	0.021*
		EFB ergonomy	NA	7.5	<0.001*
		Flight deck position	Commander	3.7	0.017*
		Operator	Bristow	8.8	0.011*
			CHC	10.7	<0.001*
			Norwegian	2.0	0.434
			SAS	3.8	0.157
			Widerøe	1.9	0.456
		Touch screen familiarity	NA	4.9	0.114
		EFB training	NA	6.1	0.002*
		Digitalization[…]:Touch[…]	Negative	18.4	0.002*
			Positive	- 5.7	0.103
OM† revisions (15)	0.17	(Intercept)†	NA	67.0	0.006*
		Digitalization attitude	Negative	- 2.1	0.119
			Positive	5.9	0.032*
		Operator	Bristow	9.7	0.102
			CHC	- 2.6	0.617
			Norwegian	-10.9	0.014*
			SAS	-13.4	0.003*
			Widerøe	-14.2	0.001*
		Touch screen familiarity	NA	11.5	0.020*
		EFB training	NA	16.0	<0.001*
		Digitalization[…]:Touch[…]	Negative	12.4	0.207

Cont.

			Positive		0.103
Flight deck workload (16)	0.23	(Intercept)[†]	NA	43.0	0.589
		Age	over 55 YOA	7.0	0.001*
			under 40 YOA	5.2	0.008*
		Automation stress	NA	4.8	0.001*
		Operator	Bristow	7.0	0.079
			CHC	- 0.6	0.862
			Norwegian	-15.1	<0.001*
			SAS	-15.8	<0.001*
			Widerøe	-24.1	<0.001*
		Touch screen familiarity	NA	1.3	0.002*
		EFB training	NA	5.0	0.018*
		Automation[…]:Touch[…]	NA	- 8.5	0.005*

† reference level: *Age*; "40-55 YOA", *Operator*; "Other", *Digitalization attitude*; "Indifferent", *Flight deck position*; "FO"

1. adjusted R^2 model fit

2. standardized regression coefficients

3. Holms adjusted and calculated from students t-distribution, with α = .05. Significance is denoted by Asterix (*)

Table S13. Relative importance for each predictor across all 16 univariate models. The relative importance is shown as absolute values of the total variance explained by the respective model.

	Statement response models (n = 16)															
Predictor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Age								0.010								0.020
Digitalization attitude	0.05	0.03	0.04		0.02				0.02		0.03	0.05	0.04	0.03	0.03	
Operator	0.02	0.14	0.05	0.10	0.06	0.03	0.04		0.01	0.07	0.04	0.05	0.03	0.03	0.05	0.18
Touch[…] **:** Digitalization[…]	0.03		0.02									0.02	0.02	0.03	0.01	
Automation stress			0.02	0.02	0.04	0.02	0.01	0.04	0.02	0.03						0.01
EFB ergonomy	0.07	0.06	0.06	0.03	0.04	0.03	0.04		0.04	0.05	0.06	0.05	0.05	0.05		
Flight deck position	0.01		0.02	0.02	0.01			0.10					0.01	0.01		
Touch screen familiarity	0.08		0.03	0.02	0.04		0.03	0.07	0.03	0.04		0.05	0.03	0.04	0.02	0.00
EFB training	0.06	0.02	0.02	0.01		0.02	0.02				0.07	0.06	0.03	0.03	0.06	0.01
Touch[…] **:** Automation[…]				0.01	0.01		0.01	0.02	0.01							0.01
adj. R^2:	0.31	0.25	0.26	0.21	0.21	0.10	0.14	0.24	0.14	0.19	0.22	0.27	0.20	0.23	0.17	0.23

(Intentionally left blank)