Volume 11 / Number 1 / 2021

Aviation Psychology and Applied Human Factors

hogrefe

Aviation Psychology and Applied Human Factors

Volume 11/Number 1/2021

Official Organ of the European Association for Aviation Psychology (EAAP)
and the Australian Aviation Psychology Association (AAvPA)

Publisher
Hogrefe Publishing, Merkelstr. 3, D-37085 Göttingen, Germany, Tel. +49 551 99950-0,
Fax +49 551 99950-425, E-mail publishing@hogrefe.com, Web http://www.hogrefe.com
North America: Hogrefe Publishing, 361 Newbury Street, 5th Floor, Boston, MA 02115, USA,
Phone (866) 823-4726, Fax (617) 354-6875, E-mail publishing@hogrefe.com

Production
Regina Pinks-Freybott, Hogrefe Publishing, Merkelstr. 3, D-37085 Göttingen, Germany,
Tel. +49 551 99950-428, Fax +49 551 99950-425, E-mail production@hogrefe.com

Subscriptions
Hogrefe Publishing, Herbert-Quandt-Str. 4, D-37081 Göttingen, Germany,
Tel. +49 551 99950-900, Fax +49 551 99950-998, E-mail zeitschriftenvertrieb@hogrefe.de

Advertising/Inserts
Marketing, Hogrefe Publishing, Merkelstr. 3, D-37085 Göttingen, Germany,
Tel. +49 551 99950-423, Fax +49 551 99950-425, E-mail marketing@hogrefe.com

ISSN
ISSN-L 2192-0923, ISSN-Print 2192-0923, ISSN-Online 2192-0931

Publication
Published in 2 issues per annual volume.

Subscription Prices
Calendar year subscriptions only. Rates for 2021: Institutions – from US $283.00/€219.00 (print only; pricing for online access can be found in the journals catalog at hgf.io/journalscatalog);
Individuals – US $148.00/€106.00 (print & online). Shipping costs: Germany €6.00,
other countries US $9.00/€8.00. Single copies – US $149.00/€117.50 + postage and handling.

Payment
Payment may be made by check, international money order, or credit card, to Hogrefe Publishing, Merkelstr. 3, D-37085 Göttingen, Germany, or, for North American customers, to Hogrefe Publishing, 361 Newbury Street, 5th Floor, Boston, MA 02115, USA.

Electronic Full Text
The full text of *Aviation Psychology and Applied Human Factors* is available online at http://econtent.hogrefe.com and in PsycARTICLES.

Abstracting/Indexing Services
Abstracted/indexed in PsycINFO, PSYNDEX, and Academic Index.

Contents

Original Article

A Qualitative Study Exploring Well-Being and the Potential Impact of Work-Related Stress Among Commercial Airline Pilots

Paul Cullen, Joan Cahill, and Keith Gaynor

School of Psychology, Trinity College Dublin, Ireland

Abstract: Increasing evidence suggests that commercial airline pilots can experience physical, mental, and social health difficulties. Qualitative interviews with commercial airline pilots explored the relationship between work-related stress and well-being. Participatory workshops involving pilots were conducted. The methodology of this action-based research involved a blend of person-centered design approaches; specifically, "stakeholder evaluation" and "participatory design." The findings further support the hypothesis that pilot well-being is being negatively affected by the nature of their work. The biopsychosocial model of the lived experience of a pilot, as presented in this paper, provides a useful structure to examine pilot well-being, and to identify and scope potential coping strategies to self-manage health and well-being issues associated with the job of being a pilot.

Keywords: pilot well-being, biopsychosocial, work-related stress, risk management, coping strategies

Despite exceptionally low risks posed to the travelling public (International Civil Aviation Organization, 2009; Flight Safety Foundation, 2019), there is mounting evidence to suggest that commercial aviation poses long-term health risks to pilots. Pilots can be considered as both "shift workers" and "remote workers," with a wealth of studies examining well-being issues related to these types of work. Numerous studies indicate these types of work can be detrimental to one's well-being, and demonstrate that such duties are shown to increase the risk of:

- Anxiety, depression, increased neuroticism, and impaired cognitive function (Eldevik et al., 2013; Proctor et al., 1996);
- Reduction in quality and quantity of sleep (Caruso, 2014; Reis et al., 2016);
- Widespread complaints of fatigue (Lee & Kim, 2018; Park et al., 2001);
- Increased risk of adverse cardiovascular effects (Brown et al., 2009; Hermansson et al., 2007; Pimenta et al., 2011);
- Possible increased risk of certain types of cancer (Anjum et al., 2012; Sancar et al. 2015);
- Increased risk of type 2 diabetes (Axelsson & Puttonen, 2012; Knutson & Kempe, 2014) and metabolic syndrome (Wang et al., 2014);
- Possible increase in gastrointestinal effects (Nojkov et al., 2010);
- Marital strain, family dysfunction, and social marginalization (Muurlink et al., 2014); and
- Increased risk of fertility difficulties (Goldstein & Smith, 2016).

Following the 2015 GermanWings tragedy, the psychological well-being of airline pilots came into sharp focus. The European Aviation Safety Agency (EASA) went to great lengths to strengthen medical requirements for pilots. In July 2018, new safety rules were published, including improved provisions to better support the mental fitness of pilots. Commission Regulation EU 2018/1042 mandates the introduction of pilot support programs, alcohol/drug testing, and pre-employment psychological assessment. The EASA's response, and the subsequent recommendations, focus on assessment and identification of current risk, and presently contain no preventative elements.

Evidence suggests that pilot mental health might be under threat from sources of work-related stress (WRS). Overall, 40% of airline pilots (n = 1,147) are reported as experiencing levels of high burnout (Demerouti et al., 2018). Separately, 12.6% of airline pilots sampled (n = 1,848) were reported to have met the threshold for clinical depression or displayed major depressive disorder

https://doi.org/10.1027/2192-0923/a000199

symptoms (Wu et al., 2016), while 4.1% of these pilots were reported as having had suicidal thoughts within the same period. In 2016, 17% of pilots sampled agreed that their company cared about their well-being (n = 7,239), while 21% felt that fatigue was taken seriously within their organization (Reader et al., 2016). In 2012 a correlation between measured levels of common mental disorders (CMD) among pilots and workload was reported (Feijó et al., 2012). Encouragingly, this research demonstrated that regular physical exercise was associated with a lower risk of CMD. Little focus to date has been placed on the relatively higher proportion of pilots who are not experiencing mental health difficulties.

This research highlights the general well-being of pilots as an important area of focus allowing significant preventative work to be done. Specifically, work-related factors that impinge on pilot well-being are not well understood and/or managed, and an opportunity has been missed to potentially correct this.

Biopsychosocial Models of Well-Being

Despite extensive literature detailing the impact of shift work, remote work, and high cognitive demands on other professions, there is a lack of information concerning these factors and commercial airline pilots. Equally, while fatigue in pilots, for instance, is widely studied (e.g., Johansson & Melin, 2018), few studies explore broader conceptualizations of pilot well-being and even fewer give voice to pilot experience. This paper is part of an overall project to develop a broader perspective of the relationships between WRS, pilot well-being, pilot performance, and flight safety (The Pilot Lived Experience Project).

Research has focused on developing a biopsychosocial model of pilot lived experience (Cahill et al., 2018; Cullen et al., 2016, 2017). According to biopsychosocial models of health and well-being (Engel, 1977; Havelka et al., 2009), the cause, manifestation, and outcome of wellness and disease are determined by a dynamic interaction between biological, psychological, and social factors. None of these factors in isolation are sufficient to lead definitively to wellness or illness. Instead, the interrelationships between all three pillars result in a given outcome (Engel, 1977; Havelka et al., 2009; Santrock, 2007). The biopsychosocial model has been utilized to develop interventions to lower stress and/or improve people's ability to cope with stressors, aiming toward mental and physical health (Johnson & Acabchuk, 2018). A health psychology perspective is fundamentally behavioral, in that the majority of chronic illness can be avoided or reduced through healthy lifestyles.

Previous qualitative research involved informal semistructured scoping interviews with active commercial airline pilots (n = 103) in which their lived experience was explored. This involved identifying sources of WRS and discussing how these impacted physically, mentally, and socially on the individual pilots and their families. This was further supported by a broad literature review of a range of factors that have been documented as affecting well-being, leading to the development of a preliminary biopsychosocial model of pilot well-being (Cullen et al., 2017). This current paper describes a subsequent qualitative study of 33 pilots aged between 25 and 60 years who attended participatory workshops exploring the impact of WRS on well-being.

Method

Research Aim

The workshops had two objectives. First, to examine WRS and the specific work factors impacting on pilot well-being. The workshop endeavored to validate a proposed biopsychosocial model of the lived experience of being a pilot. Secondly, the workshops attempted to map the relationship between WRS, pilot well-being, pilot performance, and flight safety. This paper focuses on the first objective. Subsequent papers will examine the impact of WRS and potential intervention tools.

Research Design

This was an action research study involving a series of three sequential workshops with commercial pilots (n = 33). The methodology involved a blend of person-centered design approaches; specifically, "stakeholder evaluation" (Cousins et al., 2013) and "participatory design" (Bødker & Buur, 2002). The workshops were led by the second author, an experienced qualitative researcher.

This research was premised on two relevant theoretical frameworks: (1) biopsychosocial models of well-being (Engel, 1977; Havelka et al., 2009), and (2) phenomenological approaches to eliciting information about "lived experience" (Lindseth & Norberg, 2004; Van Maanen, 1988). The research effort focused on understanding the context and meaning of airline pilots experience (phenomenological approach).

Ethics approval for this research was obtained from the School of Psychology, Trinity College Dublin (TCD), Ireland.

Workshop Structure

Briefing information was provided to participants 7 days prior to each workshop, including information about the preliminary biopsychosocial model of pilot lived experience. Participants were briefed about confidentiality issues. All provided written consent and agreement to maintain confidentiality in relation to anything discussed/disclosed by workshop attendees.

The biopsychosocial model of health was further explained. Taking the form of a "round the table" discussion, participants were invited to review the model. Efforts were made to ensure that participants were not asked leading questions. A group discussion followed, concerning relationships between WRS, pilot well-being, pilot performance, and flight safety. As part of this, participants reviewed a series of performance/safety impact scenarios. These are reported in a different paper (Cahill et al., 2018). After each session, participants were debriefed and the need for confidentiality re-emphasized. Participants were also invited to undertake an optional review exercise from home.

Sampling Method

The study used a mix of (1) quota sampling (selection of participants for interviews) and (2) opportunity sampling. The workshops were advertised on social media. Inclusion criteria for participants were: being in an age range of 25–65 years; possessing a commercial and/or air transport pilot license. We also sought a mixture of males/females and first officer and captains.

Participants

Three workshops were conducted at Dublin Airport between March and April 2018, involving 33 active commercial pilots. Workshops 1, 2, and 3 were attended by 12, 10, and 11 pilots, respectively, spanning three Irish registered airlines. Participants had on average 9,178 hr of flying experience, and included 20 captains and 13 first officers. Of them, seven were female and 26 were male. Overall, 25 were on full-time contracts, with eight working part-time; four flew regional/short-haul operations, 22 flew medium-haul, and seven flew long-haul.

Data Analysis

Workshop transcripts were written up by one member of the research team, and reviewed by the other team member. Further, participants reviewed transcripts after each session.

Overall findings of all three workshops pertaining to (1) sources/causes of WRS, (2) manifestation/symptoms, and (3) health outcomes were analyzed. Participant lived experience data were organized into a series of 16 themes. Specific themes were linked to each of the three pillars of well-being. These are reported in the results section. Specific findings are organized into a series of infographics for each pillar (see next section).

Results

Participants gave a wide variety of feedback, some of which has been reported in a separate paper (Cahill et al., 2019) that examined the impact of WRS on flight safety. However, the data reported in this paper look deeper at the relationships between WRS and pilot well-being, with the intention of validating the preliminary biopsychosocial model of pilot well-being.

The Relationship Between WRS and Pilot Well-Being

Three major super-ordinate constructs emerged from the workshops, largely mapping onto the constructs of biological, psychological, and social factors. Within these, 16 themes were evident and these are outlined in Tables 1, 2, and 3. Within the participants' feedback, there were strong inter-relationships between these themes. This feedback was used to validate and further develop the biopsychosocial model of pilot well-being, which is depicted in Figures 1, 2, and 3.

Theme 1: Biological Issues, Consequences and Causes

Participants highlighted a range of biological issues that they identified as impacting them as a result of work and WRS.

Workshop participants reported frequently feeling tired and sleep-deprived, due to working long, irregular, antisocial hours with frequent time-zone changes and associated sleep displacement. This was normalized as a routine aspect of life as a pilot. Many considered themselves to be either "early birds or night owls," feeling that they often worked against their body clock and that their sleep was disrupted.

Furthermore, they reported diet as a concern, and either consumed crew meals provided by their airline or brought their own food, with many critical of the portion sizes of crew meals and suspicious of the nutritional content and processed nature. Those who self-catered reported frequent stress due to issues with airport security. Participants reported regular snacking between meals, with meals normally consumed when opportunity allowed (e.g., during quieter periods of their duty), rather than when hungry. Many reported gastrointestinal (GI) issues, such as irritable bowel syndrome (IBS), abdominal bloating, and

Table 1. The relationship between WRS and pilot biological well-being

Super-ordinate theme	Themes	Participant feedback (direct quotes)
Biological issues, Consequences & Causes	• Fatigue	"…not enough time between shifts to recover" "I'm a night owl, and when on early duties, stay awake early in the morning using coffee. Still can't get to sleep until late, and end up only getting about 4 hours sleep…" "Wear & tear & fatigue hits you when you stop… sick during days off & annual leave" "I never had IBS until I started work as a pilot…"
	• Diet, hydration, and bowel movements	"…lack of breaks, including bathroom breaks. Postponed defecation is a big stress" "…can't drink sufficient water, otherwise running to the bathroom"
	• Back pain/musculoskeletal issues • Low levels of physical exercise	"…different environment from the 90's… health issues emerging now, not seen as often in the past, or not at all… young captains needing back surgery…" "…restricted movement due to locked cockpit door… can't rotate in the chair… can't stretch out or stand up straight… muscle cramps and stiffness" "…not simply sedentary, but effectively impaled into the seat… need to get permission to use the bathroom"

Note. WRS = work-related stress. IBS = irritable bowel syndrome.

Table 2. The relationship between WRS and pilot psychological well-being

Super-ordinate theme	Themes	Participant feedback (direct quotes)
Psychological issues, consequences & causes	• Not feeling valued	"…feel dehumanised by management, referred to as a 'fulltime equivalent'…just a staff number… feel objectified…" "Management know the cost of everything, but the value of nothing. Experience doesn't count anymore…"
	• Diminished authority	"Many pilots are high achievers, and their abilities are not being tapped into" "No top cover provided by management… they don't have our backs…"
	• Social isolation	"Get to ops, go to aircraft, get things moving, 30-minute turn-around, do it all over again. Don't get 5 minutes to myself…" "The vagueness of some rules is a source of stress…" "Battle with management a constant source of stress and anxiety… they are bonus chasers and don't experience the impact of their decisions… they're not on the same team as the pilots"
	• Employment practices	"We're always understaffed…" "After years of this treatment, you become weary and disengaged" "…no personal space while at work" "Pilots show up to work and tick all the boxes. Things don't give until the end…" "We're task-orientated, tend to keep going, get on with things to achieve the task. We keep pushing on…" "…pilot might be struggling… something on the day pushes them over the edge, something gives on a particular day…" "I'm not proud of my company…" "Flying 900 hours is the new norm, not a boundary or limit. It's the target…" "If the company want to engage, start by stop kicking us in the head…" "Our EAP [employee assistance program] has had its resources slashed, now at a time when its needed more than ever…"
	• Aeromedical requirements	"Better off saying nothing. Germanwings pushed mental health issues underground…"

Note. WRS = work-related stress.

Table 3. The relationship between WRS and pilot social well-being

Super-ordinate theme	Themes	Participant feedback (direct quotes)
Social issues, consequences & causes	• Family strain	"…fixed-pattern roster has benefits, predictability, but can be very inflexible… missing major family and social events…" "…come home and your body clock is out of sync with the rest of the family"
	• Marriage/Spousal relationship	"We're not special or unique, but our job is unique…"
	• Lack of understanding from others	"We're 5 miles up, hanging onto a pair of wings. It's unnatural but we've normalised this… can't just step out for a minute…"
	• Loneliness	"…dragging your life behind you in a suitcase"
	• Home life Isolation	"As a foreigner, it's hard to make new friends, in a new country… cultural differences… hard to fit in"
	• Work life isolation	"…can be stuck in close confines with someone you don't like…" "Due to lack of familiarity with colleagues, each day in work is like a first date" "Things can snowball quickly. You're fine at the start of the week, and suddenly by the end of the week, you're not coping… we're no good at seeing this in ourselves"
	• "Macho" culture	"I off-loaded myself… considered a wimp by management… conversations with managers reflected macho culture and stigma around mental health…"

Note. WRS = work-related stress.

hemorrhoids, and some attributed this to diet, disrupted toilet habits, and sitting for long periods within the confines of the cockpit.

Due to a reluctance to urinate regularly, dehydration was also reported as an issue. Numerous pilots reported a reluctance to have a bowel movement while on an aircraft, delaying defecation until after their shift, with reasons given such as the disruption to the cabin crew, a lack of privacy, and also the perception of being monitored by their passengers.

Reports of lower back pain were common among participants, as were complaints of poor ergonomics within the cockpit.

Owing to the long working days and sedentary nature of the job, many pilots reported taking little or no physical activity during their working week. Antisocial and irregular hours made regular exercise as part of teams/clubs difficult.

Theme 2: Psychological Issues, Consequences and Causes

Participants highlighted a range of psychological stresses associated with their work.

Many participants, particularly more experienced crew, reported that levels of stress and responsibility have increased during the past 15 years, highlighting that new responsibilities have appeared, due to commercial pressures.

Some felt that the commander's authority and autonomy were diminishing. Participants reported working longer duties with less rest time and for less remuneration. Overall, there was a sense that pilots were not valued as much as previously by their employers. The term "glorified bus drivers" was frequently used.

Many pilots described a disconnection between their own values and those of their line managers, feeling that their employer cared little about their welfare. Some pilots reported feeling that occasionally safety, but more often staff well-being, was compromised in favor of commercial requirements. Participants reported not feeling psychologically safe in raising well-being-related concerns with their managers. Fatigue risk management and work–life balance were considered to be company philosophies that had not made it into practice. Some pilots reported feeling psychologically drained, having spent years of working long, irregular, and antisocial hours in an environment where they felt undervalued. A number of pilots reported morale as a safety concern, with some saying that they found it difficult to remain motivated in maintaining their professionalism.

Due to changing employment practices, many pilots are hired indirectly via agencies, with some not receiving payment if unable to fly due to sickness. Increasing numbers of pilots have substantial financial debt owing to the high initial training costs. This was reported as a deterrent in reporting sick, and a cause of distress due to professional conflict. Some captains also reported this as a source of concern when flying with such pilots. In some cases, pilots reported feeling that "calling in sick" was not considered acceptable by their line managers or base captains.

The stringent aeromedical requirements were considered by many to be a deterrent in openly discussing well-being issues, particularly those related to mental health. Due to the perception of pilots possessing "the right stuff" and having been "cut from the same cloth as astronauts," pilots

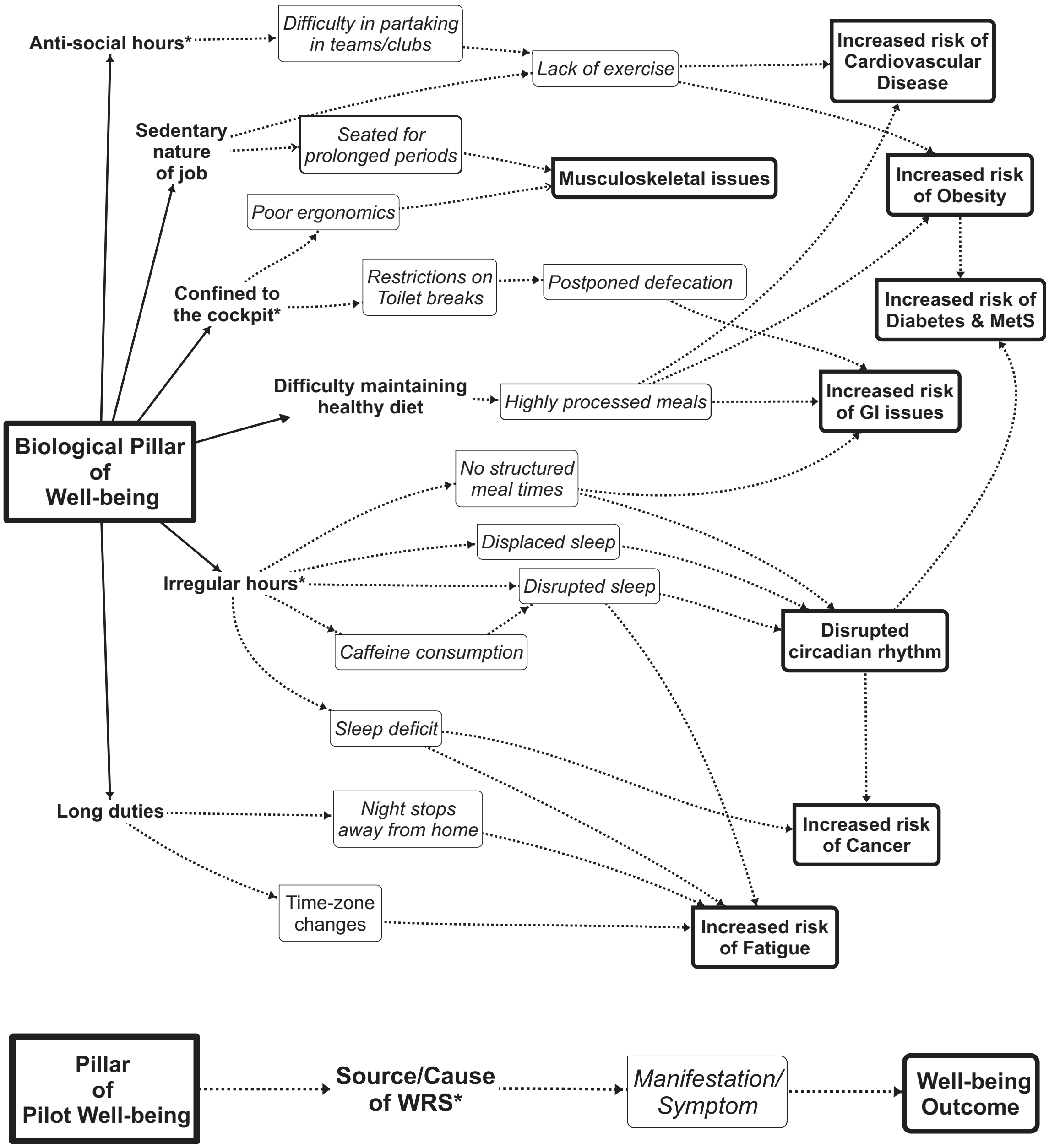


Figure 1. Biological pillar of pilot well-being. MetS = metabolic syndrome. WRS = work-related stress.

reported that if they developed mental health issues, they would most likely not discuss it with colleagues. Other pilots reported that owing to the perception of "living the dream" and a lack of support from those outside the profession, they would be unlikely to disclose such an issue to "an outsider."

Theme 3: Social Issues, Consequences and Causes

The relationship between work practices, the impact of work, and social outcomes was highlighted widely by participants.

A strained home life, loneliness, and poor social networks were reported by many participants, and they associated these with working irregular and antisocial hours. Managing and navigating the home–work interface was perceived as challenging, and spousal relationships were sometimes reported as strained.

Despite generous annual leave entitlements, many participants reported difficulties getting leave during peak times such as school holidays, and reported that important family events were frequently missed because of inflexible rosters. Delays were seen as a cause of strain, with spouses

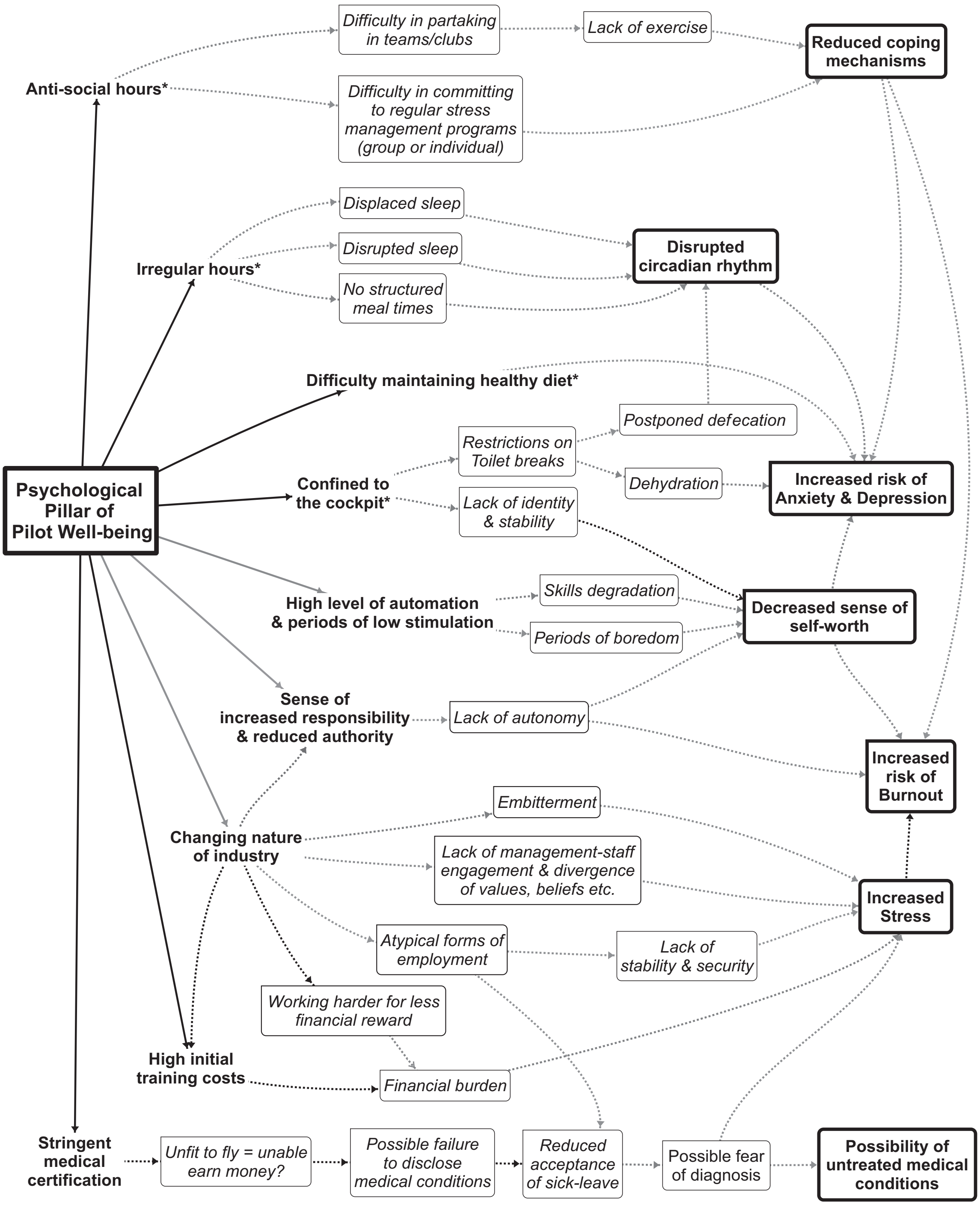


Figure 2. Psychological pillar of pilot well-being.

not confident that their pilot-partners would be off-duty as scheduled and able take up their share of domestic/parental responsibilities.

Lack of control over rosters and frequent working at weekends/holidays were highlighted as a source of family strain. It was felt, particularly in families where both parents were working, that domestic responsibilities were unequally shared. This was further exacerbated by the fact that pilots frequently spend nights away from base, while spouses manage the family alone.

Two distinct types of loneliness were reported, and appeared to be linked to the length of time spent working as a pilot and to the ages of their children. Those with younger children, and perhaps those who were younger themselves, reported loneliness borne out of not being where they wanted to be, that is, with loved ones. The

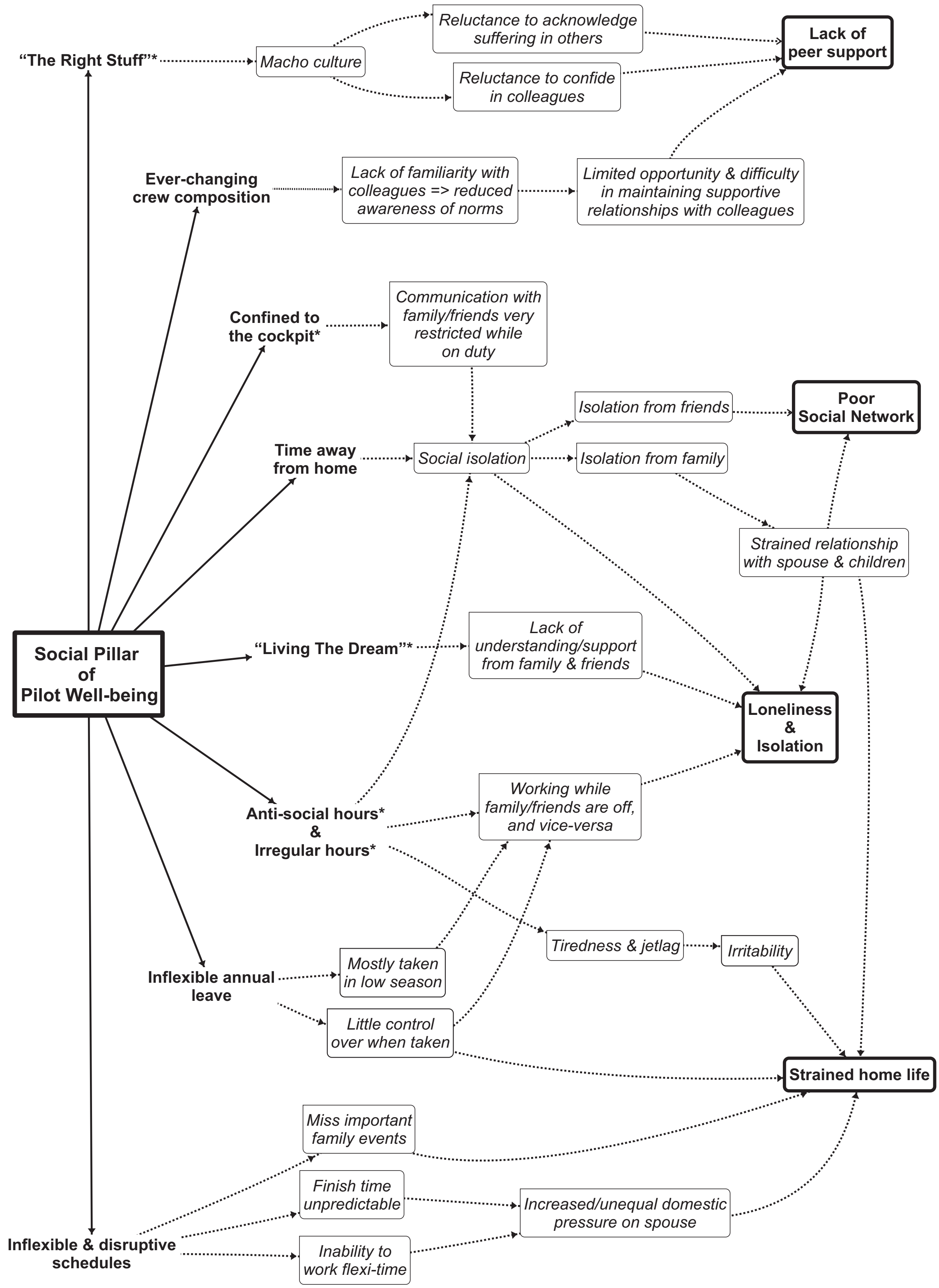


Figure 3. Social pillar of pilot well-being.

second group were older than the first, had worked as pilots for a longer period, typically had children in the teenage years or older. These pilots reported feeling distant from their families, with some feeling that they were "just paying the bills," and were somewhat removed or isolated from their families. They felt this was due to their repeated absences over many years. Some pilots who were approaching retirement also expressed concern over how they would "fit back into their family" after retirement.

Most pilots reported that their families and friends held unbalanced views of the challenges facing pilots, not fully appreciating the negative aspects of being a pilot, and instead only focused on the positives. As such, many felt they did not receive appropriate understanding and support in dealing with well-being issues, such as loneliness, stress, or fatigue. With many pilots fulfilling a childhood ambition, many reported that in the eyes of others they were "living the dream."

Due to irregular and antisocial working hours, many reported difficulties in maintaining regular social routines and connections. Some spoke of social isolation and loneliness. Many pilots spoke of hobbies that they enjoyed, but no longer did since becoming a pilot. Regular commitment to team sports or societies was seen as very difficult. Many reported physical exercise was often solitary, for example, going to the gym, swimming, running, cycling. Pilots spoke of having to make a disproportionate effort in arranging events with friends, as the friends usually assumed they would be unavailable because of work commitments.

Due to ever-changing crew composition and lack of familiarity with fellow crew members, participants reported difficulties in forging close bonds with colleagues. This prevented them from knowing whether colleagues were having good or bad days and was seen as detrimental to building supportive work environments. Some captains described a sense of social isolation, due to the often-sizeable age difference between themselves and other crew members. This was exacerbated on duties involving night stops. Some participants (both captains and copilots) described a lack of peer support on a routine basis.

Some pilots described the culture in which they worked as male-dominated and "macho," that is, stable extroverts who are mentally resilient and calm under pressure. This was seen as feeding a culture in which pilots were considered to be immune from experiencing mental health issues. Peer support programs operate with a high degree of confidentiality, and consequently participants expressed a view that if a pilot approaches them, he/she is most likely unaware of how frequently the service is utilized, and therefore may fear that they are the only one ever to have approached the group. If the pilot is already feeling overwhelmed, this perception may further increase feelings of isolation.

Validation of the Biopsychosocial Model

Although there was naturally a significant crossover between themes, participants' concerns largely fell within the broad framework of biological, psychological, and social factors. When participants were presented with the model, they endorsed it as a whole, highlighting specific aspects, particular to their individual experiences. Firstly, these data validate the concept of pilots' experiences being understood using a biopsychosocial model rather than a more specific focus, such as fatigue. Secondly, these factors largely mapped onto the preliminary biopsychosocial model of pilot lived experience as proposed by Cullen and colleagues (2017). Thirdly, these data validated many of the individual factors identified within that model. Fourthly, these factors are in line with literature reviews of WRS and pilots, the area of WRS and well-being in general.

Discussion

The aim of this paper was to (1) map the relationships between WRS and pilot well-being and (2) to validate the preliminary biopsychosocial model of pilot lived experience.

The workshops with the pilots gave rich and varied data concerning issues of WRS and pilot well-being. It was noticeable how many of the pilots identified with issues of WRS and were concerned about their own long-term well-being. The participants' data were organized into three super-ordinate themes (biological, psychological, and social). Each of these are associated with a series of subthemes (of which there are 16 in total). We attempted to capture these in Figures 1–3. Although pilot well-being often focuses on specific issues such as fatigue or rosters, the pilots have highlighted a broad range of issues that affect them and that should be addressed in order to maintain pilot health.

The breadth of the issues discussed within the biopsychosocial model of pilot lived experience stands in stark contrast to a limited focus of pilot fatigue, or suicidality for instance. The model presented in this paper highlights the need for a multidimensional approach to pilot well-being. The breadth of this conceptual framework highlights the range of different ways that pilot well-being is affected by their role. It offers a broad range of opportunities for intervention. This conceptualization is much broader than often discussed and gives a range of areas to be explored in further research.

Despite being proposed over 40 years ago, the biopsychosocial model remains influential today. However, the model has been criticized in the literature for being too vague and for not providing enough detail as to how the individual pillars interact and contribute to wellness and illness (Benning, 2015; Farre & Rapley, 2017). Although literature exists on how biological, psychological, and social factors are separately associated with health, causal links between these factors have not been clarified. Recent studies demonstrate how interrelationships among these factors can be investigated. Karunamuni et al. (2020) propose an updated theoretical model: the biopsychosocial pathways model, which considers potential pathways between the individual pillars (biological, psychological, and social), and attempts to explain how these pathways can contribute to subjective well-being and objective physical health outcomes.

Limitations

Limitations of this study include the qualitative nature of small-scale workshops, potential bias due to the self-selected sample, and the fact that the workshops took place over three points in time. The sample composition is made up of commercial pilots flying for airlines based in Ireland, and the study results need to be replicated in other cohorts of pilots and with large-scale quantitative research.

Areas for Further Analysis

An online health questionnaire has been launched, incorporating the standard instruments to measure levels of distress used in previous studies (Demerouti et al., 2018; Wu et al., 2016). This detailed questionnaire also examines lifestyle factors that are commonly accepted to impact on physical, mental, and social well-being. Within the questionnaire, additional quantitative and qualitative research will be undertaken to further validate and develop the biopsychosocial model of pilot lived experience.

It is hoped that the findings of this questionnaire will assist in furthering the understanding of what contributes to some pilots being susceptible to distress while others remain resilient. This will build on the findings previously reported (Cahill et al., 2019), regarding effective coping strategies that can be utilized by pilots.

Following this, it is planned to engage in a consultation process with all stakeholders to identify how the well-being/mental health issues identified in the survey might be addressed at different levels (i.e., pilot self-management, airline, regulator etc.). Although our research to date has focused solely on commercial airline pilots, many of our results are likely to apply to other pilot groups such as cargo operations, military, and search and rescue, and perhaps other staff groups within aviation. Some of our findings were recently used by the Flight Safety Foundation (2020) to assist aviation professionals in dealing with the stress associated with the Covid-19 pandemic.

Until recently, absence of evidence suggesting that pilots were suffering was broadly taken as evidence of absence of suffering. However, substantial evidence now demonstrates that pilots are under stress and experiencing well-being problems. Furthermore, evidence suggests aspects of their job are contributing to these health problems. Pilots are potentially more at risk of developing well-being issues than has previously been considered, despite the perception that pilots are naturally mentally resilient. Despite the extremely low number of lives lost due to pilot suicide, there are possibly a significantly large number of pilots flying today with untreated mental health issues, such as depression, anxiety, and suicidal thoughts. Potentially, burn-out and embitterment with work/work practices can lead to disengagement/loss of motivation. Disengagement/loss of motivation can have an impact on task performance and professionalism (i.e., use of procedures, attitudes to change, willingness/interest in quality/safety processes, e.g., voluntary reporting). This in turn has an impact on flight safety.

Conclusion

If the well-being of pilots is being negatively affected by the nature of their work, this needs to be identified and measured, and the associated risks managed accordingly. The biopsychosocial model of the lived experience of a pilot, as presented in this paper, provides a useful starting point for this research, and perhaps could be utilized in training pilots for (1) identification of risky behavior and (2) development of coping strategies.

Both pilots and airlines are responsible for managing well-being issues. A first step is the identification of the challenges faced by pilots. If the true picture of pilot well-being (including the causes for well-being problems) were to emerge, this may very well help reduce or even remove the current stigmatization of mental health issues among pilots, thus enabling open disclosure and increased support. In time, perhaps airline management might reconsider their duty of care to their employees and regulators may implement measures that not only protect the safety of the travelling public, but also the well-being of the crew operating the aircraft.

Potentially, these findings can be used to identity and scope possible coping strategies for use by pilots, to self-manage health and well-being issues associated with the job of being a pilot.

References

Anjum, B., Singh, R. B., Verma, N., Singh, R., Mahdi, A. A., Singh, R. K., De Meester, F., Wilczynska, A., Takahashi, T., Dharwadkar, S., & Wilson, D. W. (2012). Associations of circadian disruption of sleep and nutritional factors with risk of cancer. *The Open Nutraceuticals Journal, 5*, 124–135. https://doi.org/10.2174/1876396001205010124

Axelsson, J., & Puttonen, S. (2012). Night shift work increases the risk for type 2 diabetes. *Evidence-Based Medicine, 17*(6), 193–194. https://doi.org/10.1136/ebmed-2012-100649

Benning, T. B. (2015). Limitations of the biopsychosocial model in psychiatry. *Advances in Medical Education and Practice, 6*, 347–352. https://doi.org/10.2147/AMEP.S82937

Bødker, S., & Buur, J. (2002). The design collaboratorium. *A place for usability design. ACM Transactions on Compututer-Human Interaction, 9*(2), 152–169. https://doi.org/10.1145/513665.513670

Brown, D. L., Feskanich, D., Sanchez, B. N., Rexrode, K. M., Schernhammer, E. S., & Lisabeth, L. D. (2009). Rotating night shift work and the risk of ischemic stroke. *American Journal of Epidemiology, 169*(11), 1370–1377. https://doi.org/10.1093/aje/kwp056

Cahill, J., Cullen, P., & Gaynor, K. (2018). *Estimating the impact of work-related stress on pilot wellbeing and flight safety*. Paper presented at the 2nd International Symposium on Human Mental Workload: Models and Applications (H-WORKLOAD 2018), Amsterdam (September 20–21, 2018). Netherlands Aerospace Centre (NLR), The Netherlands.

Cahill, J., Cullen, P., & Gaynor, K. (2019). Interventions to support the management of work-related stress (WRS) and wellbeing/mental health issues for commercial pilots. *Cognition, Technology, and Work, 22*, 517–547. https://doi.org/10.1007/s10111-019-00586-z

Caruso, C. C. (2014). Negative impacts of shiftwork and long work hours. *Rehabilitation Nursing, 39*(1), 16–25. https://doi.org/10.1002/rnj.107

Cousins, J. B., Whitmore, E., & Shulha, L. (2013). Arguments for a common set of principles for collaborative inquiry in evaluation. *American Journal of Evaluation, 34*(1), 7–22. https://doi.org/10.1177/1098214012464037

Cullen, P., Cahill, J., & Gaynor, K. (2016, November 8). *Managing mental health issues in pilots*. Paper presented at the Human Factors in Aviation Safety Conference, CIEHF, London.

Cullen, P., Cahill, J., & Gaynor, K. (2017, October 23–25). *Understanding pilot wellbeing and its impact on flight safety: Identifying self-management strategies to promote the biopsychosocial pillars of wellbeing*. Paper presented at the Flight Safety Foundation International Air Safety Summit, Dublin, Ireland.

Demerouti, E., Veldhuis, W., Coombes, C., & Hunter, R. (2018). Burnout among pilots: Psychosocial factors related to happiness and score on simulator training. *Ergonomics, 62*, 233–245. https://doi.org/10.1080/00140139.2018.1464667

Eldevik, M. F., Flo, E., Moen, B. E., Pallesen, S., & Bjorvatn, B. (2013). Insomnia, excessive sleepiness, excessive fatigue, anxiety, depression and shift work disorder in nurses having less than 11 hours in-between shifts. *PLoS One, 8*(8), e70882. https://doi.org/10.1371/journal.pone.0070882

Engel, G. (1977). The need for a new medical model: A challenge for biomedical science. *Science, 196*, 126–129.

Farre, A., & Rapley, T. (2017). The new old (and old new) medical model: Four decades navigating the biomedical and psychosocial understandings of health and illness. *Healthcare, 5*(4), 88. https://doi.org/10.3390/healthcare5040088

Feijó, D., Luiz, R. R., & Camara, V. M. (2012). Common mental disorders among civil aviation pilots. *Aviation, Space and Environmental Medicine, 83*, 509–513. https://doi.org/10.3357/ASEM.3185.2012

Flight Safety Foundation. (2019, January). *Aviation Safety Network releases 2018 airliner accident statistics*. https://news.aviation-safety.net/2019/01/01/aviation-safety-network-releases-2018-airliner-accident-statistics/

Flight Safety Foundation. (2020). *An aviation professional's guide to wellbeing*. Flight Safety Foundation.

Goldstein, C. A., & Smith, Y. R. (2016). Sleep, circadian rhythms, and fertility. *Current Sleep Medicine Reports, 2*, 206–217. https://doi.org/10.1007/s40675-016-0057-9

Havelka, M., Lucanin, J. D., & Lucanin, D. (2009). Biopsychosocial model – the integrated approach to health and disease. *Collegium Antropologicum, 33*(1), 303–310.

Hermansson, J., Gådin, K. G., Karlsson, B., Lindahl, B., Stegmayr, B., & Knutsson, A. (2007). Ischemic stroke and shift work. *Scandiavian Journal of Work, Environment and Health, 33*(6), 435–439. https://doi.org/10.5271/sjweh.1165

International Civil Aviation Organization. (2009). *Safety management manual (Doc 9859)* (2nd ed.). ICAO.

Johnson, B. T., & Acabchuk, R. L. (2018). What are the keys to a longer, happier life? Answers from five decades of health psychology research. *Social Science & Medicine, 196*, 218–226. https://doi.org/10.1016/j.socscimed.2017.11.001

Johansson, F., & Melin, M. (2018). Fit for flight? Inappropriate presenteeism among Swedish commercial airline pilots and its threats to flight safety. *The International Journal of Aerospace Psychology, 28*(3–4), 84–97. https://doi.org/10.1080/24721840.2018.1553567

Karunamuni, N., Imayama, I., & Goonetilleke, D. (2020). Pathways to well-being: Untangling the causal relationships among biopsychosocial variables. *Social Science & Medicine*. https://doi.org/10.106/j.socscimed.2020.112846

Knutson, A., & Kempe, A. (2014). Shift work and diabetes – a systematic review. *Chronobiology International, 31*(10), 1146–1151. https://doi.org/10.3109/07420528.2014.957308

Lee, S., & Kim, J. K. (2018). Factors contributing to the risk of airline pilot fatigue. *Journal of Air Transport Management, 67*, 197–207. https://doi.org/10.1016/j.jairtraman.2017.12.009

Lindseth, A., & Norberg, A. (2004). A phenomenological hermeneutical method for researching lived experience. *Scandinavian Journal of Caring Sciences, 18*, 145–153. https://doi.org/10.1111/j.1471-6712.2004.00258.x

Muurlink, O., Peetz, D., & Murray, G. (2014). Work-related influences on marital satisfaction amongst shiftworkers and their partners: A large, matched-pairs study. *Community, Work & Family, 17*(3), 288–307. https://doi.org/10.1080/13668803.2014.933775

Nojkov, B., Rubenstein, J. H., Chey, W. D., & Hoogerwerf, W. A. (2010). The impact of rotating shift work on the prevalence of irritable bowel syndrome in nurses. *American Journal of Gastroenterology, 105*, 842–847. https://doi.org/10.1038/ajg.2010.48

Park, J., Kim, Y., Chung, H. K., & Hisanga, N. (2001). Long working hours and subjective fatigue symptoms. *Industrial Health, 39*, 250–254. https://doi.org/10.2486/indhealth.39.250

Pimenta, A. M., Kac, G., De Souza, R. R. C., Ferreira, L. M. B. A., & Silqueira, S. M. F. (2011). Night-shift work and cardiovascular risk among employees of a public university. *Revista da Associação Médica Brasileira, 58*(2), 168–177. https://doi.org/10.1016/S2255-4823(12)70177-7

Proctor, S. P., White, R. F., Robins, T. G., Echevierra, D., & Rocskay, A. Z. (1996). Effect of overtime work on cognitive function in automotive workers. *Scandinavian Journal of Work, Environment and Health, 22*, 124–132.

Reader, T. W., Parand, A., & Kirwan, B. (2016). *European pilots' perceptions of safety culture in European aviation*. Technical deliverable, by the Project P5: Resolving the organisational accident of future sky safety. Reference ID FSS_P5_LSE_D5.4. https://safeorg.eu/beta/wp-content/uploads/2018/07/5_Safety-Culture.pdf

Reis, C., Mestre, C., Canhão, H., Gradwell, D., & Paiva, T. (2016). Sleep complaints and fatigue of airline pilots. *Sleep Science, 9*(2), 73–77. https://doi.org/10.1016/j.slsci.2016.05.003

Sancar, A., Lindsey-Boltz, L. A., Gaddameedhi, S., Selby, C. P., Ye, R., Chiou, Y. Y., Kemp, M. G., Hu, J., & Ozturk, N. (2015). Circadian clock, cancer, and chemotherapy. *Biochemistry, 54*(2), 110–123. https://doi.org/10.1021/bi5007354

Santrock, J. W. (2007). *A topical approach to human life-span development* (3rd ed.). McGraw-Hill.
Van Maanen, J. (1988). *Tales of the field: On writing ethnography*. University of Chicago Press.
Wang, F., Zhang, L., Zhang, Y., Zhang, B., He, Y., Xie, S., Li, M., Miao, X., Chan, E. Y. Y., Tang, J. L., Wong, M. C. S., Li, Z., Yu, I. T. S., & Tse, L. A. (2014). Meta-analysis on night shift work and risk of metabolic syndrome. *Obesity Reviews, 15*(9), 709–720. https://doi.org/10.1111/obr.12194
Wu, A. C., Donnelly-McLay, D., Weisskopf, M. G., McNeely, E., Betancourt, T. S., & Allen, J. G. (2016). Airplane pilot mental health and suicidal thoughts: A cross-sectional descriptive study via anonymous web-based survey. *Environmental Health, 15*, 121. https://doi.org/10.1186/s12940-016-0200-6

History
Received January 29, 2020
Revision received June 7, 2020
Accepted August 20, 2020
Published online February 4, 2021

Publication Ethics
Ethics approval for this research was obtained from the School of Psychology, Trinity College Dublin (TCD), Ireland.

ORCID
Paul Cullen
https://orcid.org/0000-0002-0600-7268

Paul Cullen
School of Psychology
Trinity College Dublin
Dublin
Ireland
cullenp4@tcd.ie

Captain Paul Cullen is an Airbus pilot, IFALPA accredited accident investigator and research associate at the Centre for Innovation in Human Systems (CIHS), School of Psychology, Trinity College Dublin, Ireland. Paul is keen to understand how the Safety-II philosophy can be incorporated into the management of well-being within aviation.

Joan Cahill (PhD) is a Research Fellow and Principal Investigator at the Centre for Innovation in Human Systems (CIHS), at the School of Psychology, Trinity College Dublin, Ireland. Dr Cahill's research spans three fields – Human Factors, Ethics & Behaviour Science.

Keith Gaynor is Assistant Professor in Clinical Psychology in the School of Psychology, University College Dublin, and Senior Clinical Psychologist in St. John of Gods Hospital, Dublin. Keith specializes in cognitive behavioral therapy (CBT) and mental health. His current research interests include emotional regulation, psychosis, CBT, pilot mental health and loneliness.

Original Article

Testing the Compliance Behavior Model in General Aviation

A Pilot Study

Anthony A. Stanton[1], Sidney W. Dekker[1], Patrick S. Murray[2], and Gui Lohmann[3]

[1]Safety Science Innovation Laboratory, Griffith University, Brisbane, QLD Australia
[2]Aviation and Logistics, University of Southern Queensland, Australia
[3]School of Engineering and Built Environment, Griffith University, Brisbane, QLD, Australia

Abstract: Australian general aviation accident data show pilots who conduct operations into adverse weather, when against the rules, remain as a significant cause of fatal accidents. This paper presents the background, methodology, and results of a theory of planned behavior (TPB) elicitation study, which extracted key psychological beliefs of aircraft pilots in such circumstances. The present study established a psychometric survey instrument with items that are valid and reliable, to then further explore the TPB psychological constructs concerning the intentions of pilots when presented with adverse weather. Given the principled deliberations associated with rule-related behavior, the project explores an extension of the TPB by investigating the addition of two psychological constructs – personal norms and anticipated affect and their power to provide a discrete contribution and improved explanation of variance.

Keywords: pilot violations, theory of planned behavior, rule-related behavior, general aviation compliance, personal normative influences

Australia aviation accident data between 2008 and 2017 show that pilots conducting operations into adverse weather, when against the rules, resulted in 109 reported safety occurrences (Australian Transport Safety Bureau, 2018). Safety data show the lethality of venturing into instrument meteorological conditions (IMC) for those pilots who are limited to operations under the visual flight rules (VFR). Reviewing US and Canadian data, Batt and O'Hare (2005) have shown that pilots operating VFR into IMC, when against the rules, is around four times more likely to prove fatal than any other general aviation safety occurrence. The causation is often resolved to general statements around a pilot's desire to just press on, or what is often called "*get-home-itis*." Is this behavior really that simple? Or are there much more complex influences affecting the pilot's behavior? While some researchers have attempted to understand the problem, sometimes examined in the context of *plan continuation error*, the research particularly within general aviation is limited. Much of the research has focused on occurrence statistics and related contexts, rather than an attempt to uncover an understanding of people and the latent psychological factors that influence their safety-related behaviors (decision-making). The present study explores the latent psychological beliefs of general aviation pilots when faced with adverse weather, by adapting an expectancy–value psychosocial behavior theory – the theory of planned behavior (TPB) – as a theoretical framework.

The objective of this article is to outline the theoretical framework of the research project, including a conceptual extension of the TPB and to present the results of an elicitation study. The paper begins with a description of the TPB, its sufficiency assumptions, and evidence of its appropriate application here. A discussion of a conceptual compliance behavior model then follows, which is put forward as an extension of the TPB for empirical testing in an attempt to improve the explained variance in the present rule-related behavior context. The paper then articulates the methods and results of the elicitation study, which has identified modal salient beliefs (i.e., those commonly held among the target population) and provided formative research for a following study.

The TPB is an expectancy–value behavior model that theorizes an explanation for the formation of behavioral intentions (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 1975). As shown in Figure 1, the TPB posits that a particular behavior follows reasonably and consistently from an individual's salient psychological beliefs associated with performing that behavior. Further, it is suggested that only a limited number of these beliefs dominate and hence predict an individual's intention to perform the behavior. More specifically, the theory posits that a person's behavioral

https://doi.org/10.1027/2192-0923/a000200

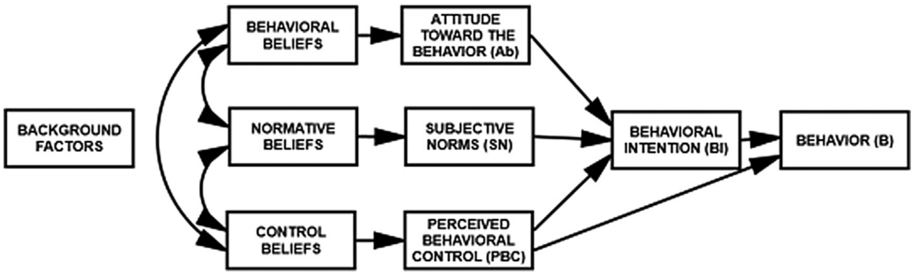


Figure 1. The theory of planned behavior (adapted from Ajzen, 2006).

intention (BI) is the most immediate antecedent of performing the behavior (B) and, further, these intentions are formulated as a result of three key psychological determinants:

(1) An individual's attitude toward the specified behavior (A_b);
(2) An individual's perceived social pressures in relation to performing or not performing the behavior (SN); and
(3) Perceptions of behavioral control (PBC) or self-efficacy.

The relationship between each of the model's direct psychological constructs (B, BI, A_b, SN, and PBC) is illustrated by the equation

$$B \sim BI = A_b\,(\beta_1) + SN\,(\beta_2) + PBC\,(\beta_3). \qquad (1)$$

The equation suggests that each of the direct constructs is separated by a beta weight (β) that reflects the influence of that particular construct (e.g., A_b) toward the formation of behavioral intentions. Different behaviors have been shown empirically to be influenced to varying extents by each of these constructs. The TPB allows this to be explored through a well-established methodology. As a result of identifying these influences and importantly their respective dominance on the specified behavior, intervention strategies can be developed that target the latent influences that have the greatest leverage on the behavior. For example, if a particular behavior, in a given context, was identified to be most influenced by subjective norms (social influence), a behavior change program may focus its efforts on injunctive (what is perceived others might expect) and descriptive (what others are seen to be doing) normative beliefs, rather than distributing information about the disadvantages or disadvantages of performing the particular behavior (attitude toward the behavior).

By further review of Figure 1, we can see that each of these direct constructs are said to be formulated from underlying latent beliefs, that is, behavioral beliefs, normative beliefs, and control beliefs. As an expectancy–value model, each of these beliefs is characterized by a two-factor composite form, which is shown empirically to correlate with the corresponding direct construct. These composite forms are often referred to as "indirect measures." The behavioral beliefs composite is expressed by $\sum_{i=1}^{n} b_i e_i$ where b represents the subjective probability that outcome i exists when performing the behavior and e represents an evaluation of that outcome i. Concerning outcome i, these two composites are then multiplied (i.e., expectancy–value). The total set of the salient belief composites is then summed. Likewise, the normative beliefs are expressed by $\sum_{i=1}^{n} n_i m_i$ where n represents the normative belief about referent i and m represents the motivation to comply with referent i. Finally, control beliefs are expressed by $\sum_{i=1}^{n} c_i p_I$ where c represents the subjective probability that the control factor i will be present when performing the behavior and p represents the perceived power of that factor to make performing the behavior easier or more difficult. Empirically, each aspect of these equations is explored by asking respondents a series of questions that explore respective subjective probabilities and evaluative aspects using 5- or 7-point scales.

Reason and coworkers (1990) have shown the likely heritage of violation or rule-related behavior to be within social (social norms) and motivational (intention) foundations, as opposed to errors, which have origins within human information-processing limitations. Because of this, and as argued by Fogarty and Shaw (2010), the TPB is therefore an ideal theoretical framework from which to examine rule-related behavior, since the TPB encompasses these influences and others. In their research, Fogarty and Shaw (2010) specifically explored the usefulness of the TPB to understand violation behavior within an aircraft maintenance setting. Their study identified the TPB was highly successful in explaining the variance in behavior and resulted in the development of a behavior model appropriate for that setting. The present study applies a similar approach, although within a very different context with quite different influences and it also explores the sufficiency of the model.

While the TPB has been used extensively as a conceptual framework for behavioral science investigations

(Barber, 2011; Conner & Armitage, 1998; Rivis et al., 2009), some researchers have explored the sufficiency of the model (Conner & Armitage, 1998; Parker et al., 1995) to find additional gains in variance explanation, through inclusion and adaption of psychological constructs. The TPB sufficiency assumption states that "additional variables should not improve prediction of either intention or behavior" (Fishbein & Ajzen, 2010, p. 281), although Ajzen (1991, p. 199) has stated that the TPB is "in principle, open to the inclusion of additional predictors if it can be shown that they capture a significant proportion of the variance in intention or behavior after the theory's current variables have been taken into account." The present research explores this sufficiency assumption. That is, the addition of other variables to improve the explained variance in rule-related behavior intention.

The subjective norm construct has been cited as the TPB weakest predictor (Armitage & Conner, 2001; Sheppard et al., 1988). As discussed earlier, this construct considers the influence of significant others. Specifically, the subjective norm construct evaluates perceived expectations of what ought to be done (injunctive social norms) and whether significant others are themselves performing the behavior (descriptive social norms). Injunctive social norms influence behavior by reflecting the patterns of the collective group, enticing reward or threatening sanction by the group for acquiescence (Cialdini et al., 1991). Descriptive norms influence differently, where observations of effective and adaptive action provide information-processing advantages (Cialdini et al., 1991). Neither of these two normative components cogitates self-expectations, being internalized personal values, which it is argued may independently influence rule-related decision-making.

The TPB does not explicitly include consideration of personal normative influence; instead, such influences are considered to be embedded within behavioral beliefs. A review of the theory of propositional control (Dulany, 1961, 1968), from which the TPB evolved, and the early adaptions by Fishbein (1967), show a personal normative component in those models. For example, the personal normative component can be identified clearly in the theory of propositional control equation $[(NB_p)(MC_p)]w_1 + [(NB_s)(MC_s)]w_2$ where subscript p represents beliefs of a personal nature, and subscript s represents the beliefs of a social nature (Fishbein, 1967).

Personal norms have been investigated extensively by Schwartz (1973, 1977) in relation to altruism, resulting in the norm activation model. In this model, pro-social environmental behavior is hypothesized to result from three determinants: awareness of consequences, the ascription of responsibility, and personal norms. Schwartz (1977, p. 227) defines personal norms as self-expectations or an internalized sense of duty constructed from general norms and personal values. Schwartz (1977) hypothesizes that conformity, or otherwise, with personal norms are experienced in a state of subjective self-awareness (Duval & Wicklund, 1972), as opposed to more conscious intellectual judgments of right and wrong. Schwartz (1977) suggests individuals experience personal norms as feelings of moral obligation, such as pride and guilt.

In the domain of illegal, antisocial, and dishonest behaviors, Manstead (2000) provides an extensive review of researchers who have sought to improve the TPB through the inclusion of a personal norm or moral norm construct. Gorsuch and Ortberg (1983) included a single item, a direct measure of perceived moral obligation, and demonstrated within moral settings that the construct provided an additional 20% explanation of the variation in behavioral intentions. The moral obligation construct provided an independent contribution and was more highly correlated with behavioral intentions than either the attitude or subjective norm constructs.

In a TPB investigation of driving violations, Parker et al. (1995) included items representing a personal normative construct: a single moral norm item and two anticipated regret items. The additional measures resulted in an improved variance explanation of behavioral intentions of up to 15%. In the Parker et al. (1995) study, the personal norm constructs contributed more to the variance than did the extant constructs. This suggests, in relevant contexts, that the personal norm construct has a considerable influence on behavior. Akin to the investigation of driving violations by Parker et al. (1995), the situational violation behavior of general aviation pilots is one likely to be influenced by internalized self-expectations, and, therefore, the personal norm construct is expected to improve the explanation of behavioral intentions. As identified in the Huntzinger (1997) study, it is plausible that when forming a behavioral intention to commit a situational violation, a general aviation pilot may hold beliefs of anticipated affect, activated as a result of their own internalized values or personal norms and the resulting moral dilemma (Seligman et al., 2002). This anticipated negative affect could include regret, apprehension, anxiety, shame, guilt, anger, or fear and influence behavioral intention (Moan et al., 2005).

While researchers have, to some extent, tested the inclusion of anticipated affect and personal norm constructs to the TPB, this has never been in the context of pilot rule-related behavior. Each domain and each behavior have unique influences, hence a need for empirical validation in each. The compliance behavior model (CBM) shown in Figure 2 is a conceptual model that is unique since it concedes anticipated affect to be a behavioral belief construct and postulates, based on the work of Schwartz (1977), that anticipated affect is an experiential (affective) attitude associated with a personal normative construct. The model,

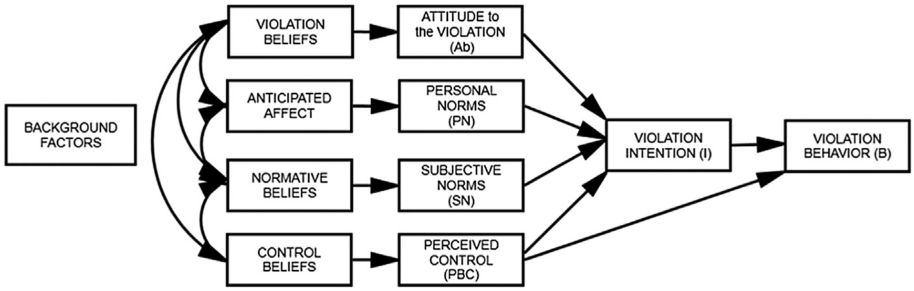


Figure 2. The compliance behavior model is a conceptual model, adding two additional psychological constructs to the TPB and labeling the behavior.

if proven, would enable a methodology for uncovering and evaluating personal normative influences on rule-related behavior in this domain. This separation of the constructs facilitates practical distinctions and interventional interests between the influences. Noting an observation by Ajzen and Sheikh (2013), unlike other researchers, the compliance behavior model would measure anticipated affect to performing the behavior itself, consistent with other measures (i.e., not the alternate course of action). In Figure 2 the compliance behavior model shows the addition of the personal norm and anticipated affect, constructs including the expected covariance. The compliance behavior model relabels extant TPB constructs more relevantly to the present application. It is hypothesized that: The compliance behavior model will predict the intentions of general aviation pilots to conduct a situational violation associated with adverse weather, and that the addition of the personal norms and anticipated affect (from the TPB) will improve the explained variance of violation intentions after existing TPB measures have been considered.

Method

TPB methodology requires that researchers conduct two separate studies. Initially, researchers conduct an elicitation or pilot study, where the primary objective is to elicit from the target population's latent beliefs in relation to the behavior. In the second (main) study, researchers leverage the elicitation study results in order to explore potential associations between the TPB constructs and the behavior. The following sections articulate the method and results of the elicitation study.

The elicitation study also has two secondary objectives: to formulate and then test the internal reliability of items for the measurement of each of the TPB direct constructs, and to evaluate the suitability of a set of TPB background measures. The methodological and analytical frameworks for TPB studies are well documented elsewhere (Fishbein & Ajzen, 2010; Francis et al., 2004). Approval to commence the elicitation and main studies was obtained from a university Human Research Ethics Committee after demonstrating compliance to specified conditions and guidelines.

The elicitation study survey was published by using the Survey Monkey web tool with invitations to participate primarily generated by social media posts from several general aviation flying organizations, such as flying clubs and flying schools. The survey was open for 2 days with the average respondent taking 6 min to complete all items. Of the 42 respondents, 30 answered all questions providing a completed response rate of 71%. A total of 47 items were employed as part of the survey. Prior to publishing the Survey Monkey questionnaire, five respondents were used to construct content for the questions that elicited beliefs in the Survey Monkey web tool. This smaller group was asked open-ended questions exploring belief themes. Each belief theme was then included in the survey.

Participants

The target population was defined as any licensed aircraft pilot, or student pilot, who is currently operating, or has ever previously operated, as a pilot within the general aviation sector. The target population was not limited to respondents within Australia, nor pilots who are operating within the general aviation sector, since any licensed pilot or trainee is reasonably able to contemplate their influences and reactions to the behavior under examination and then provide a considered response. Responses would be influenced by their particular past and present background factors.

A total of 42 participants provided responses, which is a representative sample of the target population as recommended by Fishbein and Ajzen (2010) for this formative research stage. To validate respondents were within the target population, the elicitation study asked two validation questions (1) regarding the level of pilot license held and (2) the country in which they conducted the majority of their general aviation flying. Each question had a response option allowing them to identify themselves as being outside the defined target population. Nil respondent exclusions were required.

Table 1. Notable respondent characteristics

Characteristic	Value
Respondents (*n*)	42
Male	90.48%
Female	9.52%
Median age	31–40
Min. age range	Under 20
Max. age range	71–80
Country of flying (Australia)	Australia
Flying hours (experience) < 500 hr	50%
Flying hours (experience) > 500 hr	50%
Employed as pilot (current or past)	50%
Own skill rating – same as average	47.62%
Own skill rating – slightly better than average	45.24%
License level – < Student, recreational, private	45.24%
License level – > Commercial, airline	44.24%

Background Factors

The elicitation study included seven items as trial background measures. The items related to age, gender, flying hours (experience), country, past or present employment as a pilot, self-rated skill compared with others with the same experience, and the level of pilot license held. Table 1 summarizes the key characteristics of the respondents' background information. Interestingly, the majority of respondents (92.86%) indicated that they rated their flying skill as either the same or better than other pilots of the same experience level. These seven background measures were shown to provide meaningful information. The two questions relating to license level and country in which general aviation flying took place provided a simple test that the participant was within the target population. It is plausible the background measures may have correlations with other measures. For example, the self-rated skill response is likely to be shown as correlated with PBC in a study with higher statistical power.

The Behavioral Criterion

The TPB requires the behavior that is under research examination (the behavioral criterion) to be clearly defined by four specific elements. These elements are (1) the action, (2) the target, (3) the context, and (4) time. For the elicitation study, the behavioral criterion is broadly considered – a situational violation. Such a violation is where a person operates at a rule-based level of cognition (Rasmussen, 1983), applying predefined action to preconceived situations. The behavioral criterion was described to respondents by way of a detailed scenario that incorporated each of these four elements. The scenario included a photograph taken from the perspective of the pilot's seat, illustrating the imagined weather conditions to support the text description and to ensure all respondents had the same perspective of the environmental context. The scenario depicted a hypothetical private, recreational flight in which the respondent operated an aircraft with five passengers in deteriorating weather conditions and, in doing so, committed a situational violation. The scenario depicts the respondent encountering adverse weather for which they are unable to lawfully operate within, 10 min from the destination having flown the aircraft for 50 min at that point toward the destination. This situational violation is known by pilots as operating *visual flight rules (VFR) into instrument meteorological conditions (IMC)*.

Results

The IBM SPSS software package was used for statistical analysis of respondent data. Primarily, reliability analysis and descriptive statistical reporting were used. The results and statistical analysis are discussed in this section by way of indirect measures of background factors as well as direct and indirect measures of the model.

Elicited Behavioral, Normative and Control Beliefs (Indirect Measures)

As stated earlier, the typical TPB methodology requires the pilot of beliefs from the sample population through free text responses and subsequent content analysis. Modal salient beliefs are then identified and used in the principal study for the calculation of indirect constructs according to the belief equations mentioned earlier. As an alternative, before the elicitation study, a small group of participants were asked a series of open-ended questions to obtain lists of potential salient behavioral, normative, and control beliefs associated with performing the behavior. All of these responses were then included in the elicitation study for respondents to select those that readily and spontaneously came to mind (readily accessible beliefs). To supplement this, elicitation study respondents were also provided with a free text box to include additional salient beliefs if they were not identified in the available list. For example, to obtain behavioral beliefs, respondents were asked to list the advantages and disadvantages of performing the behavior. To obtain normative referents, respondents were asked who might approve or disapprove of you performing the behavior. To obtain control factors, respondents were asked what factors make performing the behavior easier or more difficult.

To construct a set of modal salient beliefs for the principal study, TPB methodology applies a 75% rule (Francis et al., 2004). That is, the salient beliefs that reflect at least 75% of the elicitation study respondents are considered modal and are adopted in the principal study. In other words, these are likely to be the majority of the readily accessible beliefs for the sample population. Table 2 shows the modal salient beliefs and their cumulative account.

Elicited Anticipated Affect (Conceptually an Indirect Measure of Personal Norms)

Typically, in TPB methodology, instrumental attitude (affect) is obtained during the elicitation study by asking respondents about the *advantages and disadvantages* of performing the particular behavior as a behavioral outcome (behavioral belief). The conceptual compliance behavior model advocates that such a methodology is unlikely to yield affective responses associated with performing the behavior. Rather, such pilot questioning usually directs a respondent to consider behavioral outcomes that are experiential or cognitive and hence affective responses are not exposed as outcomes associated with the behavior. As such, in this elicitation study, respondents were directed to an extensive list of affects that was potentially related to performing the behavior. This list was constructed by asking a small focus group earlier to select affects that could be associated with performing the behavior. Nine items were tested as measures of the indirect construct – anticipated affect.

Table 3 presents the statistical analysis that was conducted. The analysis identified two items that scored highly on the frequency of the mid-point score (i.e., the percentage of respondents who selected a mid-point *neither* score), indicating that a high proportion (40.5% and 81.3%) of the respondents did not associate these two particular types of affect (dull vs. exciting and fun vs. boring) with the behavioral criterion. As a result, these two items were dropped. Additionally, two other items were dropped to reduce the overall number of items; unpleasant–pleasant, and worried–unconcerned. Internal reliability is not a requirement of the indirect measures since different accessible beliefs may be inconsistent with each other (Fishbein & Ajzen, 2010).

A Test of Direct Measures

As shown in Table 4, a total of 21 items were formulated for testing as measures of the direct construct scales; attitude towards behavior (A_b), personal norms (PN), social norms

Table 2. Indirect measures – other than anticipated affect

	% of elicited responses	Cumulative % of responses
Elicited behavioral outcomes (modal)		
Loss of control	19.7	–
Flight into terrain	19.7	39.4
Valued time and money invested	18.7	58.1
Keeps on schedule as committed	17.2	75.3
Spatial disorientation	1.5	76.8
Penalty from regulator	1.3	78.1
Avoids inconvenience to others	1.0	79.1
Elicited normative referents (modal)		
Other pilots like me	22.0	–
Flight instructors	13.6	35.6
Passengers on board	13.3	48.9
Pilots much more senior than me	12.9	61.8
Regulator	8.7	70.5
My employer	5.9	76.4
Elicited control factors (modal)		
Local area knowledge	20.2	–
Much longer distance flown so far	18.5	38.7
Safe terrain	17.1	55.8
More flying hours	14.5	70.3
Shorter distance remaining	5.7	76.0

Note. Elicited modal salient beliefs with cumulative responses greater than 75% for each construct.

(SN), perceived behavioral control (PBC), and violation intention (I). Each item consisted of a question or statement stem (e.g., continuing would be against my principles) and a corresponding 7-point bipolar adjective scale, with positive and negative endpoints (e.g., agree vs. disagree, bad vs. good). Positive and negative endpoints were mixed from left to right to reduce "response set," as recommended by Francis et al. (2004). Respondents were asked to select the score that best represented their opinion concerning the question stem. Each scale consisted of multiple items. Items were re-coded in SPSS to reflect a high rating as being a positive attitude toward performing the behavior (i.e., that they would perform the situational violation).

Reliability analysis was conducted to ascertain the level of internal consistency between items of the same scale (i.e., for each direct construct). The first of a series of reliability tests for the direct measures is shown in Table 4. Values for Cronbach's α and the corrected item total correlation are reported. For each scale, items were removed in successive reliability tests until a Cronbach α of 0.70 was exceeded for the scale and the correlated item total correlations were above 0.50. Table 5 shows the final reliability results for the direct constructs with six items having been removed from the original.

Table 3. Indirect measures (anticipated affect constructs)

Scale	Abbreviated stem description	*N*	Cronbach α if item deleted	Corrected item total correlation	Freq. mid-point score
a	Dull/exciting	32	0.760	0.257	40.5%
a	Stressful/relaxing	32	0.699	0.397	0.0%
a	Restfulness/tension	32	0.704	0.378	0.0%
a	Unpleasant/pleasant	32	0.687	0.569	0.0%
a	Anxious/calm	32	0.696	0.431	5.6%
a	Self-respect/guilt	32	0.700	0.398	16.7%
a	Worried/unconcerned	32	0.676	0.602	3.0%
a	Fun/boring	32	0.703	0.372	81.3%
a	Regret/satisfied	32	0.653	0.596	18.2%

Note. Shaded items were dropped to improve the suitability of the anticipated affect scale.

Table 4. Direct measures – first reliability test

Scale	Abbreviated stem description	*N*	Cronbach α if item deleted	Corrected item total correlation	*SD*
A_b	Bad/good	36	0.423	0.535	0.64488
A_b	Wise/foolish	36	0.387	0.533	0.72320
A_b	Harmful/beneficial	36	0.800	0.426	1.83852
PN	Against my principles	36	0.513	0.456	0.80277
PN	Would be morally wrong	36	0.841	0.410	2.02122
PN	Would be irresponsible	36	0.314	0.661	0.84092
SN	Valued others would do	33	0.645	0.410	1.81586
SN	People important think safe	33	0.625	0.605	0.86930
SN	People valued would approve	33	0.618	0.564	1.05886
SN	I would feel under pressure	33	0.701	0.181	2.32004
SN	People important think safe	33	0.608	0.702	1.00284
SN	Would be expected of me	33	0.671	0.172	1.95305
PBC	Safe for me	31	0.733	0.441	1.07663
PBC	Easy for me	31	0.655	0.676	1.85959
PBC	Up to me	31	0.774	0.205	1.23393
PBC	I am confident I could	31	0.676	0.617	1.77194
PBC	Difficult for me	31	0.754	0.356	1.79904
PBC	I have the ability	31	0.655	0.673	1.99731
I	I would	32	0.688	0.733	1.13192
I	I would not	32	0.820	0.620	1.54502
I	Similar circumstances I intend	32	0.731	0.675	1.20775

Note. Shaded items were dropped to improve the scale reliability.

Discussion

The primary objective for this study was to uncover the salient beliefs associated with each of the model's constructs and to then construct a set of modal salient beliefs for the target population. Fishbein and Ajzen (2010) have contended that salient beliefs, which are those that are readily accessible in memory and activated spontaneously with limited cognitive effort, are the primary determinants of a person's attitude toward performing the behavior. Further, Fishbein and Ajzen (2010) argue that there are only five to nine of these beliefs that are the dominant influences of a given psychological construct. To identify these limited determinants of attitude toward the behavior, Fishbein and Ajzen (2010) advocate eliciting from respondents the advantages and disadvantages of performing the behavior (i.e., specify the behavioral outcomes of performing the behavior). The behavioral outcomes that are most commonly elicited from a representative sample are then considered as a modal set of salient beliefs for the target

Table 5. Direct measures (A_b, PN, SN, PBC, and I) – final reliability test

Scale	Stem abbreviated description	Scale Cronbach α	Mean	Corrected item total correlation
A_b	Bad/good	0.802	1.3784	0.673
	Wise/foolish			0.673
PN	Against my principles	0.841		0.727
	Would be irresponsible			0.727
SN	Valued others would do	0.790		0.546
	People important think safe			0.675
	People valued would approve			0.662
	People important think safe			0.748
PBC	Safe for me	0.799		0.537
	Easy for me			0.652
	I am confident I could			0.689
	I have the ability			0.645
I	Similar circumstances I would	0.822		0.700
	Similar circumstances I intend			0.700

population. A similar concept applies to the beliefs associated with the other direct constructs of the model. The elicitation study identified salient beliefs for each of the TPB direct constructs.

The modal salient behavioral beliefs were shown to reside within seven beliefs. These limited beliefs accounted for 79.1% of all behavioral outcomes obtained from respondents. The modal behavioral beliefs were: loss of control, spatial disorientation, regulatory penalty, flight into terrain, influences of time and money pressure, keeping the flight on schedule as committed to others, and avoiding an inconvenience to others. The last three beliefs provide for an enhanced description of what has been referred to as "get-home-itis." The last two beliefs are of particular interest, as they may be more latent influences and hence less expected, although they are consistent with the explanation provided by Reason (2008) for situational violation motivations. Modal anticipated affect was shown to reside within five sensations: stressfulness, tension, anxiousness, regret, and guilt. The association of these sensations with the behavior is consistent with the findings of Causse and coworkers (2013). Using a neuroergonomics approach, these researchers demonstrated a temporary impairment of decision-making when some pilots were faced with an adverse weather-related decision. Likewise, modal social influencers were identified as: other pilots like "me," flight instructors, the passengers onboard the aircraft, pilots who are considered more senior, the regulator, and the pilot's employer. Again, the influence of a person's employer provides additional context to get-home-itis and the perceived pressure to meet work commitments. Modal control beliefs were identified as: local area knowledge, the distance flown so far, whether the terrain was considered safe, experience in the form of flying hours, and the distance remaining to the destination. The most interesting in this set of beliefs are references to the distance remaining and the distance from the departure. These two themes were evident in research by Batt and O'Hare (2005), where their analysis of 491 adverse weather-related events identified the majority of weather-related occurrences took place in the second half of the flight. Such decision-making may be associated with themes of sunk cost (Arkes & Blumer, 1985) and self-justification and escalating commitment (Staw, 1976).

The secondary objectives of the study were to: (1) formulate and test scales for the measurement of the direct constructs of the model and (2) to evaluate potential background measures. The study appraised 20 scale items, of which six were dropped to achieve an acceptable level of internal consistency for the respective constructs. The sample size within the present study is insufficient to have adequate statistical power to evaluate correlations among the model constructs. Such investigations of correlations and β are intended for a subsequent study, which leverages the findings here. The present study tested seven items as background measures and identified each of these were suitable for use in a principal study. The most interesting results here are related to the item that asked respondents to rate their own skill in comparison with other pilots of the same experience. Interestingly, 45.2% of respondents suggested their own self-rated skill was *slightly better than* a pilot with the very same level of experience. Comparatively, only 7.1% of respondents suggested their self-rated skill was *slightly less than* a pilot with the same level of experience. Such a statistic alludes to the target population having high perceptions of perceived behavioral control.

In summary, having identified the latent modal beliefs of general aviation pilots in relation to conducting VFR flight into IMC, it is these themes in particular that should be central in any intervention program that attempts to change pilots' attitudes toward the behavior, their perceptions of what important others expect and do, and also a pilot's own self-efficacy. For example, a persuasive safety education program is likely to be more effective if it leverages the social influencers listed in Table 2, who deliver messages that provide new information that underpins the beliefs also listed in that table. That is, providing pilots with new information that might vary these identified beliefs, leveraging those who provide social influence, or highlighting how the perceived control factors have influenced historical tragedy.

In conclusion, the elicitation study has achieved each of the set objectives and in doing so has provided aviation safety education experts lists by which to theme their messaging on the topic. A key limitation is that this study has not identified which of these beliefs, or even which of the direct constructs themselves, *most* strongly influence the intention to perform the behavior. Such an assessment requires statistical techniques such as path analysis, for which there is insufficient statistical power here. It is this more advanced analysis with a broader scale of respondents that takes place in the next stage of research, building on the essential findings here.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2006). *Constructing a TPB questionnaire: Conceptual and methodological considerations*. University of Massachusetts. https://people.umass.edu/aizen/pdf/tpb.measurement.pdf

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice Hall.

Ajzen, I., & Sheikh, S. (2013). Action versus inaction: Anticipated affect in the theory of planned behavior. *Journal of Applied Social Psychology, 43*(1), 155–162. https://doi.org/10.1111/j.1559-1816.2012.00989.x

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35*(1), 124–140. https://doi.org/10.1016/0749-5978(85)90049-4

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *The British Journal of Social Psychology/The British Psychological Society, 40*(Pt 4), 471.

Australian Transport Safety Bureau. (2018). *Aviation occurrence statistics 2008 to 2017*. Australian Transport Safety Bureau.

Barber, J. S. (2011). The theory of planned behaviour: Considering drives, proximity and dynamics. *Vienna Yearbook of Population Research, 9*, 31–35.

Batt, R., & O'Hare, D. (2005). Pilot behaviors in the face of adverse weather: A new look at an old problem. *Aviation, Space, and Environmental Medicine, 76*(6), 552–559.

Causse, M., Dehais, F., Péran, P., Sabatini, U., & Pastor, J. (2013). The effects of emotion on pilot decision-making: A neu roergonomic approach to aviation safety. *Transportation Research Part C, 33*, 272–281. https://doi.org/10.1016/j.trc.2012.04.005

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 201–234). Academic Press. https://doi.org/10.1016/S0065-2601(08)60330-5 (http://www.sciencedirect.com/science/article/pii/S0065260108603305)

Conner, M., & Armitage, C. J. (1998). Extending the theory of planned behavior: A review and avenues for further research. *Journal of Applied Social Psychology, 28*(15), 1429–1464. https://doi.org/10.1111/j.1559-1816.1998.tb01685.x

Dulany, D. E. (1961). Hypotheses and habits in verbal "operant conditioning". *The Journal of Abnormal and Social Psychology, 63*(2), 251–263. https://doi.org/10.1037/h0047703

Dulany, D. E. (1968). Awareness, rules, and propositional control: A confrontation with S-R behavior theory. In T. Dixon & D. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 340–387). Prentice-Hall.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self awareness*. Academic Press.

Fishbein, M. (1967). Attitude and the prediction of behavior. In M. Fishbein (Ed.), *Readings in attitude theory and measurement* (pp. 477–492). John Wiley & Sons.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior*. Addison-Wesley Publishing.

Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. Taylor & Francis Group.

Fogarty, G. J., & Shaw, A. (2010). Safety climate and the theory of planned behavior: Towards the prediction of unsafe behavior. *Accident Analysis and Prevention, 42*(5), 1455–1459. https://doi.org/10.1016/j.aap.2009.08.008

Francis, J., Johnston, M., Eccles, M., Walker, A., Grimshaw, J. M., Foy, R., Kaner, E. F. S., Smith, L., & Bonetti, D. (2004). *Constructing questionnaires based on the theory of planned behaviour: A manual for Health Services Researchers. Quality of life and management of living resources*. Centre for Health Services Research. http://openaccess.city.ac.uk/id/eprint/1735

Gorsuch, R. L., & Ortberg, J. (1983). Moral obligation and attitudes: Their relation to behavioral intentions. *Journal of Personality and Social Psychology, 44*(5), 1025–1028. https://doi.org/10.1037/0022-3514.44.5.1025

Huntzinger, D. (1997). Pilots behaving badly. *Flight Safety Australia*. (September 1997).

Manstead, A. S. R. (2000). The role of moral norm in the attitude-behavior relation. In D. J. Terry & M. A. Hogg (Eds.), *Attitudes, behavior, and social context: The role of norms and group membership* (pp. 14–25). Psychology Press.

Moan, I. S., Rise, J., & Andersen, M. (2005). Predicting parents' intentions not to smoke indoors in the presence of their children using an extended version of the theory of planned behaviour. *Psychology & Health, 20*(3), 353–371. https://doi.org/10.1080/08870440512331317706

Parker, D., Manstead, A. S. R., & Stradling, S. G. (1995). Extending the theory of planned behaviour: The role of personal norm. *British Journal of Social Psychology, 34*, 127–137.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13, 13*(3), 257–266. https://doi.org/10.1109/TSMC.1983.6313160

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

Reason, J. T., Manstead, A., Stradling, S., Baxter, J., & Campbell, K. (1990). Errors and violations on the road: A real distinction? *Ergonomics, 33*(10/11), 1315–1332.

Rivis, A., Sheeran, P., & Armitage, C. J. (2009). Expanding the affective and normative components of the theory of planned behavior: A meta-analysis of anticipated affect and moral norms. *Journal of Applied Social Psychology, 39*(12), 2985. https://doi.org/10.1111/j.1559-1816.2009.00558.x

Schwartz, S. H. (1973). Normative explanations of helping behavior: A critique, proposal, and empirical test. *Journal of Experimental Social Psychology, 9*(4), 349–364. https://doi.org/10.1016/0022-1031(73)90071-1

Schwartz, S. H. (1977). Normative Influences on altruism. In B. Leonard (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221–279). Academic Press.

Seligman, C., Olson, J. M., & Zanna, M. P. (2002). *The psychology of values: The Ontario* (Vol. 8). Psychology Press.

Sheppard, B. H., Hartwick, J., & Warshaw, P. R. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15*(3), 325–343. https://doi.org/10.1086/209170

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16*(1), 27–44. https://doi.org/10.1016/0030-5073(76)90005-2

History
Received February 29, 2020
Revision received June 13, 2020
Accepted August 20, 2020
Published online December 2, 2020

Publication Ethics
The views expressed by and attributable to this author are his own and do not necessarily reflect those of the Civil Aviation Safety Authority or the Australian Government.

ORCID
Anthony A. Stanton
https://orcid.org/0000-0003-0206-1483

Anthony A. Stanton
Safety Science Innovation Laboratory
Griffith University
37 Sandpiper Avenue
Brisbane, QLD 4509
Australia
tony.stanton@griffith.edu.au

Anthony Stanton is a PhD candidate at the Griffith University Safety Science Innovation Laboratory, Brisbane, QLD, Australia, with a research interest in the psychology of compliance behavior. Tony has held leadership roles such as CEO, Board Member, Chief Pilot and Flight Examiner. Today he holds a senior manager position with the Australian aviation regulator.

Sidney Dekker is Professor at Griffith University, Brisbane, QLD, Australia. He is best-selling author of, most recently, *The Safety Anarchist* (2018), *The End of Heaven* (2017), *Just Culture* (2016), *Safety Differently* (2015), *The Field Guide to Understanding 'Human Error'* (2014), *Second Victim* (2013), *Drift Into Failure* (2012), and *Patient Safety* (2011).

Patrick Murray is Professor of Aviation and Logistics at the University of Southern Queensland, Australia. Previously holding senior positions in the military, a major airline, and the Australian aviation regulator, he is an active flying instructor and researches in LOSA, evidence-based training, and airline safety.

Gui Lohmann is Professor in Aviation and Deputy Head of School of Engineering and Built Environment at Griffith University, Brisbane, QLD, Australia. His research expertise includes air transport studies, particularly the interface with tourism, airline business models, transport geography, the regulatory regime of Australian airports, as well as airport passengers' travel patterns and behaviors.

Original Article

Ranking Pictorial Cues in Simulated Landing Flares

Danny Benbassat[1], Abigail Konopasky[2], and Ting Dong[2]

[1]Department of Physical Medicine & Rehabilitation, School of Medicine, Uniformed Services University of the Health Sciences (USUHS), Bethesda, MD, USA

[2]The Henry M. Jackson Foundation for Military Medicine/Center for Health Professions Education, School of Medicine, USUHS, Bethesda, MD, USA

Abstract: Two-dimensional pictorial cues provide depth perception information that help pilots initiate the landing flare 10–20 ft from the ground. Although prior studies established the importance of three specific cues, they failed to rank-order their importance. This exploratory paper presents two studies, with different methodologies, that examine the effect of these pictorial cues on depth perception. In both studies, participants experienced simulated scenarios and attempted to initiate the landing flare 10–20 ft above ground level. Study 1 included 121, and Study 2 included 141, naïve participants with no prior flight experience. Combined, the findings suggest that flight instructors, training literature, and airport architects should emphasize the runway above all other pictorial cues.

Keywords: landing flare, roundout, leveloff, pictorial cues, monocular cues, depth perception, general aviation

The National Transportation Safety Board (NTSB) defines the landing phase as the beginning of the landing flare until the aircraft comes to a stop, exits the runway, or initiates a touch-and-go (NTSB, 2014). The Airplane Flying Handbook (Federal Aviation Administration [FAA], 2016) explains that in order to prevent crashing the airplane into the ground, general aviation pilots flare the aircraft 10–20 ft (3–6 m) above ground level (AGL) by gradually raising the nose of the aircraft until a smooth main-wheels touchdown. After touchdown, the pilot continues to hold the nose up until the nose wheel touches down and the after-landing roll speed is reduced to normal taxi speed. This paper focuses on the point at which the pilot initially arrests the descent 10–20 ft AGL. Some refer to this transitory, and momentary, stage as the "leveloff" (Benbassat et al., 2005), but this paper uses the term "landing flare" for simplicity.

According to the NTSB, the landing phase of operations accounted for most personal and instructional flying accidents between 2007 and 2009 (NTSB, 2011) and in 2011 (NTSB, 2014). One of the principal findings was that hard landings, or other abnormal runway contact, were among the primary defining events for instructional and personal accidents. Nevertheless, the NTSB made no distinction among the various landing phase maneuvers. Benbassat and Abramson (2002a) analyzed 6,676 NTSB accident reports and found that the landing flare accounted for 18.33% of all landing accidents in 1995, 1996, and 1997. An additional analysis of 6,655 reports revealed that the trend did not change in 1998, 1999, and 2000 (Benbassat et al., 2005).

Whereas NTSB reports provided indirect evidence to the difficulty of the landing flare, pilot reports provided direct evidence. Olson and Austin (2006) measured the landing performance of 28 novice flight students. The authors examined 12 landing dimensions and found that both flight instructors and students found the landing flare to be most difficult. Benbassat and Abramson (2002a) found that 134 pilots believed the flare to be more difficult than nine other general aviation flight maneuvers. Lastly, aviation authors provided anecdotal evidence to the difficulty of the landing flare (e.g., Langewiesche, 1972; Love, 1995). Bramson (1982) stated that the landing flare "can cause the majority of student pilots to question why they took up flying (and make their instructors wish they had stuck to golf)..." (p. 44).

It is possible that the maneuver is difficult because it is ambiguous. The *Airplane Flying Handbook* teaches that the landing flare is started in what "appears to be" 10–20 ft from the ground (FAA, 2016). Since onboard general aviation altimeters may be off by as much as 75 ft (22.8 m), pilots must visually determine their altitude AGL before gradually increasing the pitch and angle of attack. Failure to accurately determine what appears to be 10–20 ft AGL may result in hard landings (wheelbarrowing), stalls, and porpoise bouncing, which may translate into nosewheel damage and increased wear on the nosewheel shimmy

https://doi.org/10.1027/2192-0923/a000205

dampener and tire (Butcher, 1996; Christy, 1991; Jorgensen & Schley, 1990; Kershner, 1998; Love, 1995).

Currently, there are three schools of thought that attempt to explain how pilots determine what “appears to be” 10–20 ft from the ground. The first approach, *time-to-contact* (TTC), evolved from Gibson’s work on optical flow in the 1950s (Entzinger & Suzuki, 2008; Gibson et al., 1955). It suggests that pilots time the landing flare based on perceived angular change known as *Tau* (θ). The second approach, *critical runway angle*, suggests that pilots time the landing flare based on a perceived absolute angle known as Psi (Ψ) (Entzinger & Suzuki, 2008). The third approach, *pictorial cues*, is used to determine absolute distance from the ground and is based on depth perception cues (Entzinger & Suzuki, 2008).

Pictorial cues provide pilots with depth perception information from a two-dimensional picture (Benbassat et al., 2002; Goldstein, 2007). These cues have been studied for at least five decades (Riordan, 1974) but it was unclear which cues are most important. Benbassat and Abramson attempted to answer this question by conducting a series of survey studies. In one survey study, 134 pilots rated the horizon and end of runway, shape of runway or runway markings, and familiar objects as most important (Benbassat & Abramson, 2002a). In another study, 92 pilots, in two flight schools, used the shape of the runway or runway markings, altimeter, and horizon or runway end to time the landing flare. Whereas Benbassat and Abramson succeeded in naming familiar objects/familiar size (F), horizon/end of runway (H), and runway shape/runway markings (R) as most prominent, they failed to rank-order these cues. In one study, pilots from one university selected H as most important, but pilots from another university selected R (Benbassat & Abramson, 2002a). In another, whereas most expert- and intermediate-level pilots used R, most novice pilots used the altimeter in the cockpit (Benbassat, 2005).

The purpose of this exploratory paper is to rank-order F, H, and R using two independent simulated scenarios. There was no attempt to investigate other pictorial cues or the mechanisms by which pictorial cues are used.

Method

Design Overview

The design of this study is grounded in associative learning (Benbassat & Abramson, 2002b; Qian et al., 2017). Student pilots seem to implicitly learn to associate a critical angle, dynamic angle, or pictorial cue with optimal landing flare altitude through trial and error (Benbassat & Abramson, 2002b). Attempting to translate the optical flow approaches to flight training is tricky at best. The use of dynamic and critical angles may confuse the most mathematically inclined student pilots. It is plausible to program automated systems to flare the aircraft at θ or Ψ. It is implausible to use these mathematical models to teach student pilots. Therefore, with the interest of translating results to flight training, this paper focuses on pictorial cues.

Two repeated-measures studies were conducted to rank the effect of F, H, and R on altitude perception prior to initiating the landing flare. The studies were not sequential but independent of one another. The first study used crude approximations of normal approaches and landings. The second study used realistic daytime and nighttime approaches and landings. Similar to the concept of triangulation in social sciences (Patton, 1999), using multiple research designs with offsetting biases increase validity by counteracting sources of error (Greene, 2007). Therefore, the two studies were combined in order to seek corroboration and convergence. The two studies were conducted at the Ohio Northern University in Ada, OH and approved by the university’s Internal Review Board. Participants signed a consent form and agreed to participate in return for research credits.

Study 1

In the first repeated-measures study, 121 participants viewed eight slideshows depicting a general aviation aircraft on approaches and landings. The slideshows represented a factorial of three pictorial cues (F, H, R), no cues, and all cues (3! + 1 + 1). Participants were asked to initiate the landing flare 10–20 ft (6–8 m) AGL.

Participants

Participants were 121 (65 women and 56 men, mean age = 19.91 years, *SD* = 1.20) undergraduate students from Ohio Northern University. The students were recruited from introductory psychology classes in return for research credits. This convenience sample had normal or corrected vision and no aviation experience. The use of naïve participants was crucial in replicating student pilot visual judgment.

Apparatus

NEC Portable Projector

An LT30 full-color model (resolution = 1,024 × 768 pixels) projector was used to cast an image on an 8.26-ft-wide (252 cm) by 5.77-ft-high (176 cm) screen.

Gateway E4100 Series

A Gateway personal computer (Pentium 4 CPU 2.80 GHz; 496 MB of RAM) with Windows XP Professional Edition (2003) was used to open and run Microsoft PowerPoint.

Microsoft Flight Simulator 2000 (FS2000) Professional Edition

This technologically advanced detailed personal flight simulator program had more than 20,000 airports and 14 aircraft. It provided detailed 3D scenery with 16-bit color based on true elevation data. A resolution of 1,024 × 768 pixels was used for optimal graphics quality and instrument panel readability. The flight simulator was used to capture screenshots of a Cessna Skylane (C 182S) cockpit, with 120° horizontal field-of-view, on approach and landing.

Figure 1. The “All” condition. Flight instruments were disabled. (Microsoft product screen shot, reprinted with permission from Microsoft Corporation). © Microsoft Corp.

Design and Procedure

This study consisted of eight daytime approach and landing scenarios. It was completely automated to control for experimenter bias and counterbalanced to control for order (sequence) effects.

Experimental Conditions

The first author, a commercial pilot with instrument rating, recorded a daytime landing using the flight simulator flight video option. The author placed the aircraft on a 1-nautical mile final to runway (RWY) 6 in Cuyahoga County Airport (KCGF), 057° magnetic heading, and 1,313 ft (400 m) mean sea level (MSL) indicated altitude. Runway 06 had an 883-ft (269 m) MSL indicated elevation and was a 5,502 × 100-ft (1,677 × 30 m) asphalt surface. The aircraft was configured for final approach with the runway numbers as the focal point, a constant (72 KIAS) rate of descent, and centerline alignment. It continued the standardized approach until ground impact.

As shown in Figure 1, H, F, and R were clearly visible and represented the three pictorial cues of interest. At KCGF, the most prominent F was the control tower. According to Goldstein (2007), we judge distances based on prior experience with object sizes. Knowledge of the actual height of objects like hangars and trees (FAA, 2016; Kershner, 1998) influences the perception of height above ground. In order to create permutations of the three pictorial cues, the first author dissected the real-time landing simulation into static screenshots. The first screenshot depicted the aircraft at 1,313 ft and each successive screenshot depicted a 10-ft (3 m) standard decent. This resulted in a 43-screenshot PowerPoint presentation that, when animated, created the effect of a normal approach. The word “CRASH” appeared on slide 44. This original slideshow was labeled the “all” condition.

Next, the first author graphically removed all three pictorial cues. As shown in Figure 2, this resulted in an approach that lacked F, H, or R. This condition was labeled

Figure 2. The “No Cue” condition. (Microsoft product screen shot, reprinted with permission from Microsoft Corporation). © Microsoft Corp.

Figure 3. The Familiar Objects × Horizon condition. (Microsoft product screen shot, reprinted with permission from Microsoft Corporation). © Microsoft Corp.

the "no cue" condition. Finally, the author created factorial slideshow presentations, each with 44 slides, for the three pictorial cues. Figure 3 shows an example of the F × H permutation. This process resulted in six factorial slideshow (3!) presentations, one "all," and one "no cue" slideshow presentation, for a total of eight slideshow presentations.

Tutorial and Demonstration

As noted in the consent form, participants were informed that, "Today, we will test your ability to determine altitude above ground. There are no 'right' or 'wrong' answers, do the best you can. Before we start, we need to explain an aviation term." An illustrative PowerPoint tutorial, including a 4-min video (Cessna, Producer, 2001), was used to explain and show the landing flare. Participants were instructed not to focus too close or too far ahead when attempting to determine height AGL. It was also noted that vision is the primary tool in determining when to initiate the landing flare, but no specific cues were named to avoid biasing participants.

The last slide read:
You are about to view eight slideshows.
Each show simulates a standard landing.
Your task is to FLARE the aircraft.
In other words, say FLARE when the aircraft is 10–20 ft from the ground.
Let's start with an example.

The simulator was configured for an approach and landing to an aircraft carrier off the coast of San Francisco (Flight Simulator 2000 > Activity > Have an Adventure > Carrier landing). This adventure was chosen to minimize practice effects with pictorial cues in a normal airport. Participants were instructed to look outside the cockpit and say "flare" when the aircraft was 10–20 ft from the deck. This scenario was repeated until participants understood what was required of them.

Experimental Slideshow Presentations

The simulator was configured for one of the eight experimental conditions into Cuyahoga County (Co.) Airport. The pre-flight briefing read:
In this slideshow you are landing on Runway 6 at Cuyahoga Co. in Ohio.
Remember, say "FLARE" when the aircraft is 10–20 ft from the ground.
Be assertive, say flare LOUDLY and QUICKLY.
You may NOT change your mind after saying flare.
The aircraft instruments are disabled, look OUTSIDE.
This landing will NOT be repeated.

When a participant said "flare," the slideshow was paused and the next slideshow was presented. This process repeated until each participant experienced all eight conditions. The researcher did not provide feedback and participants had to rely on visual cues because mechanical altitude references were disabled. When no runway existed, the instruction "in this slideshow you are landing on Runway 6 at Cuyahoga Co. in Ohio" was replaced with "in this slideshow you are landing on the grass at..." The rest of the preflight briefing remained the same.

Results

Descriptive Statistics

As shown in Table 1, participants in the runway condition (M = 1,005.01 ft) "flared" closest to the optimal flare altitude (M = 903 ft). The next closest means were for the R × H (M = 1,012.23 ft), H × R × F (M = 1,012.64 ft), and R × F (M = 1,015.77 ft) conditions. The least optimal flares were observed in the horizon condition (M = 1,092.62 ft), followed by H × F (M = 1,082.38 ft) no cue (M = 1,075.77 ft), and familiar objects (M = 1,062.49) conditions. Remember, descent was not simulated in real-time but in 10-ft increments. Thus, Study 1 absolute flare altitudes are not as informative as relative altitudes.

Mean Flare Altitude

A repeated-measures univariate analysis of variance (ANOVA), with significant sphericity and a significant multivariate normality test, was performed to explore the effect of pictorial cues on flare estimates (feet). The analysis was performed because the F test is robust to slight, moderate, and severe departures from normality, regardless of sample size, balanced groups, or group distribution shapes (Blanca et al., 2017; Norman & Streiner, 2014), and because the Q-plot showed a reasonably normal distribution. A review of standardized residuals failed to reveal any residual greater than |3|. Since the assumption of sphericity was violated and since the epsilon level for the interaction was more than .75, the Huynh–Feldt correction was used to adjust the degrees of freedom for the omnibus F test. The main effect of pictorial cues was significant at the .05 level, F (5.91, 710.10) = 59.152, $p < 0.001$, partial eta squared (η_p^2) =.71. Since the main effect was significant, pairwise comparisons were explored while using the Bonferroni correction to guard against Type I error.

As shown in Table 2, findings indicate that perceptions of flare altitude (feet) in the runway condition (1,005.01) were not significantly different from the R × H (M = 1,012.23), H × R × F (M = 1,012.64), or R × F (M = 1,015.77) conditions, but were significantly different from the horizon (M = 1,092.62), H × F (M = 1,082.38), no cue (M = 1,075.77), and familiar objects (M = 1,062.49) conditions. Furthermore, perceptions in the horizon condition (M = 1,092.62) were not significantly different from the H × F

Table 1. Pictorial cues means, standard deviations, confidence intervals, and mean difference standard errors (feet)

				Mean difference *SEs*						
Cue	*M*	*SD*	95% CI	1	2	3	4	5	6	7
1. H × R × F	1,012.64	95.39	995.47, 1029.81							
2. −H × R × F	1,075.77	91.99	1059.21, 1092.33	6.29						
3. H × F	1,082.38	101.23	1064.16, 1100.60	6.04	6.40					
4. R × H	1,012.23	108.74	992.66, 1031.81	6.90	7.78	7.91				
5. R × F	1,015.77	96.80	998.35, 1033.20	4.59	6.85	5.60	7.12			
6. H	1,092.62	96.95	1075.17, 1110.07	7.36	6.78	7.49	8.36	7.48		
7. R	1,005.01	93.29	988.22, 1021.81	4.70	6.21	6.38	6.73	4.69	6.65	
8. F	1,062.49	103.36	1043.89, 1081.10	6.73	6.11	6.79	8.61	7.07	7.43	6.42

Note. F = familiar objects. H = horizon. R = runway.

Table 2. Pictorial cues pairwise mean differences with confidence intervals

Cue	1	2	3	4	5	6	7
1. H × R × F							
2. −H × R × F	63.12**						
	[43, 83]						
3. H × F	69.73**	6.61					
	[50, 89]	[−13, 27]					
4. R × H	−.409	−63.53	−70.14**				
	[−22, 21]	[−88, −38]	[−95, −44]				
5. R × F	3.12	−59.99**	−66.61**	3.53			
	[−11, 17]	[−81, −38]	[−84, −48]	[−19, 26]			
6. H	79.97**	16.84	10.23	80.38**	76.84**		
	[56., 103]	[−4, 38]	[−13, 34]	[53, 107]	[52, 100]		
7. R	−7.63	−70.75**	−77.37**	−7.22	−10.76	−87.60**	
	[−22, 7]	[−90, −50]	[−97, −56]	[−28, 14]	[−25, 4]	[−108, −66]	
8. F	49.85**	−13.27	−19.88	50.26**	46.72**	−30.12*	57.48**
	[28, 71]	[−32, 6]	[−41, 1]	[22, 77]	[24, 69]	[−32, 6]	[36, 78]

Note. F = familiar objects. H = horizon. R = runway. Values in square brackets indicate 95% confidence interval for each correlation. $^*p = .003$. $^{**}p < .0001$. All other values are $.114 \geq p \leq 1.00$. The 95% confidence intervals for differences were adjusted for multiple comparisons.

(*M* = 1,082.38) and no cue (*M* = 1,075.77) conditions, but were significantly different from the runway (*M* = 1,005.01), R × H (*M* = 1,012.23), H × R × F (*M* = 1,012.64), R × F (*M* = 1,015.77), and familiar objects (*M* = 1,062.49) conditions.

Thus, the runway was the most effective, and horizon the least effective, cue in estimating flare altitude. In fact, using the horizon was less efficient than using no cues at all. As shown in Figure 4, a pattern emerged where the presence of a runway cue is associated with optimal, and absence of a runway cue is associated with suboptimal, flare altitude perceptions.

Study 2

In the second, repeated-measures study, 145 participants viewed recorded videos of simulated general aviation aircraft approaches and landings. First, they practiced 10 daytime landings with an instructor (practice "dual" landings). Then, participants performed three daytime and three nighttime solo landings ("solo" landings). As in Study 1, their task was to initiate the landing flare 10–20 ft AGL.

Participants

Participants were 145 (71 men and 74 women, mean age = 19.4 years, *SD* = 1.2) undergraduate students from Ohio Northern University. The students were recruited from introductory psychology classes in return for research credits. As in the first study, the sample had normal or corrected vision and no aviation experience.

Apparatus

The same apparatus used in Study 1 was also used for training and flight simulation in Study 2.

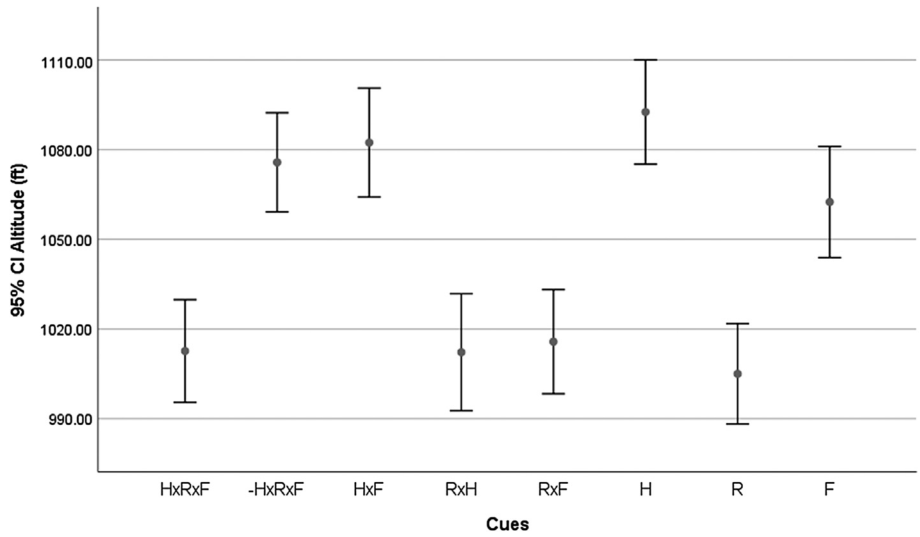


Figure 4. Error bars of eight pictorial conditions.

Design and Procedure

This study consisted of day and night landings that forced participants to use bright light (photopic) or dim light (scotopic) vision. As in Study 1, it was completely automated to control for experimenter bias and counterbalanced to control for order (sequence) effects.

Training Simulation Video

The first author recorded a daytime landing using the Flight Simulator flight video option. The landing scenario was found in Microsoft Flight Simulator Tutorial 7, Situation 3. This landing scenario placed the aircraft on a 2-nautical mile final to RWY 12 in Mojave Airport (KMHV), 122° magnetic heading, and 4,450-ft (1,356 m) MSL-indicated altitude. Runway 12 had a 2,790.04-ft (850.40 m) MSL-indicated elevation and was a 12,500 × 200-ft (3,810 × 61 m) asphalt surface with a 1.0° gradient.

The aircraft was configured for final approach with the runway numbers as the focal point, a constant rate of descent (75 KIAS), and centerline alignment. At the threshold point, the aircraft was 24.7 ft (7.5 m) AGL and the landing flare was initiated 20 ft (6 m) AGL. After the flare, the aircraft was kept at a constant 12° angle of attack until touchdown.

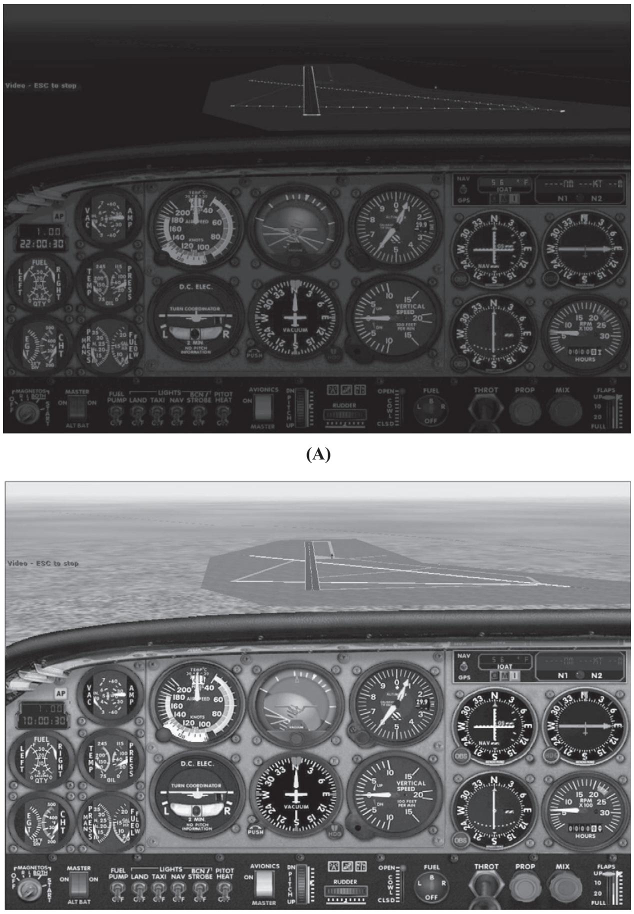
Practice Dual Landings

After watching the same tutorial used in Study 1, participants practiced 10 approaches into RWY 12 at KMHV. The researcher acted as the instructor by reading from a prepared transcript. Initially, the researcher focused on situation awareness and pointed out the runway, taxiways, intersecting runway, and control tower. She also debriefed the landing and used the "Estimating Height and Movement" instructions found in the Airplane Flying Handbook (FAA, 2014). Later, the researcher focused on the landing flare and reminded participants to disregard the instruments, look outside the cockpit, and try to determine 10–20 ft (3–6 m) AGL.

The researcher sat behind the participant and had access to real-time flight data. When the aircraft was 20 ft (6 m) AGL (2,810 ft; 856.48 m MSL), the researcher said "just about now begin to flare." She said this as the aircraft initiated the landing flare for the first five landings. Then, the researcher encouraged the participant to determine when to flare on her/his own. Instructions changed from "look outside and try to flare the aircraft with me" (Landing 6) to "look outside and say flare when the aircraft is 10–20 ft from the ground" (Landing 7), to "look outside and say flare firmly and assertively when the aircraft is 10–20 ft from the ground" (Landing 8–10).

Solo Simulation Video

The first author recorded another landing using the same Microsoft Flight Simulator Tutorial 7, Situation 3. The landing was almost identical to the training video with the exception of the landing flare. Instead of flaring, the aircraft maintained a constant nose-down attitude until ground impact. It was possible to change the recording illumination from daytime to nighttime by moving the simulator clock from 10:00 to 22:00. As shown in Figure 5, at night, the runway lights included runway edge identifying lights (REIL), white edge lights, green light threshold lights, and red runway end lights. The taxiways included blue edge lights. Thus, the runway delineation and, arguably, the horizon were the sole pictorial depth cues common to day and night landings.



(A)

(B)

Figure 5. Nighttime (A) and daytime (B) approaches to RWY 12. (Microsoft product screen shot, reprinted with permission from Microsoft Corporation). © Microsoft Corp.

Solo Landings

In the solo landing phase, participants viewed three daytime and three nighttime landings. Prior to each flight, participants were briefed that:

You are ready to solo the aircraft without your instructor.
The landings are simulated in real time.
Your task is to say "FLARE" when the aircraft is 10–20 ft from the ground.
Be assertive, say flare LOUDLY and QUICKLY.
You may NOT change your mind after saying flare.
The aircraft instruments are disabled, look OUTSIDE.
Any questions? If not, say 'begin'."

When the participant said "flare," the video was paused and the aircraft was suspended in midair. If the participant failed to say flare, the aircraft continued the descent until ground impact. The researcher did not provide feedback and the participant had to rely on depth perception cues because mechanical altitude references were disabled.

Results

Mean Flare Altitude

A dependent (paired) samples test was performed to explore the effect of pictorial cues on landing flare estimates (feet). Standardized residuals greater than |3| and suspicious DAY–NIGHT residuals were scrutinized but none warranted omission. As shown in Figure 6, there was no significant difference in flare estimates (feet) between the photopic (M = 2835.37) and scotopic (M = 2837.19) vision conditions, $t(434) = -1.30$, $p = .193$, 95% CI [−4.44, .920]. The paired samples correlation, depicted in Figure 7, was significant at the .001 level, $r = .83$.

The main finding was that reduced pictorial cues did not impact flare estimates. On average, participants flared the aircraft at 2,836.28 ft (864.50 m) MSL, or 37.98 ft (11.58 m) AGL. This finding highlights the importance of cues common to both photopic and scotopic conditions. As noted earlier, these cues were the runway delineation and, arguably, the horizon.

Variance Flare Altitude

Homogeneity of variance for two dependent samples test (Sheskin, 2004) was performed to explore flare altitude variability (feet) between the photopic (SD = 26.85) and scotopic (SD = 47.45) conditions. The analysis yielded no significant difference in altitude variability among the day and night conditions, $t(433) = .016$, $p > .80$.

The dispersion of flare altitudes is an important indicator of confidence. We would expect to see a large variance in altitude when flaring in a trial-and-error fashion. On the other hand, we would expect to see a small variance when flaring with confidence in altitude perception. In this study, the variance for night and day landing flares was high but not significantly different. This finding suggests that participants were using the same strategy for both daytime and nighttime landings. It also suggests that they flared the aircraft in a trial-and-error fashion regardless of photopic or scotopic conditions (articulating the need for errorless learning in early flight training as discussed by Benbassat & Abramson, 2002b).

Discussion

Summary

In the first study, pictorial cues on approach and landing were manipulated to isolate their effect on landing flare altitude perception. The runway emerged as the most effective cue in determining altitude AGL. Participants initiated the flare 101.38 ft (30.90 m) above the optimal flare altitude. Conversely, when presented alone, the horizon was the

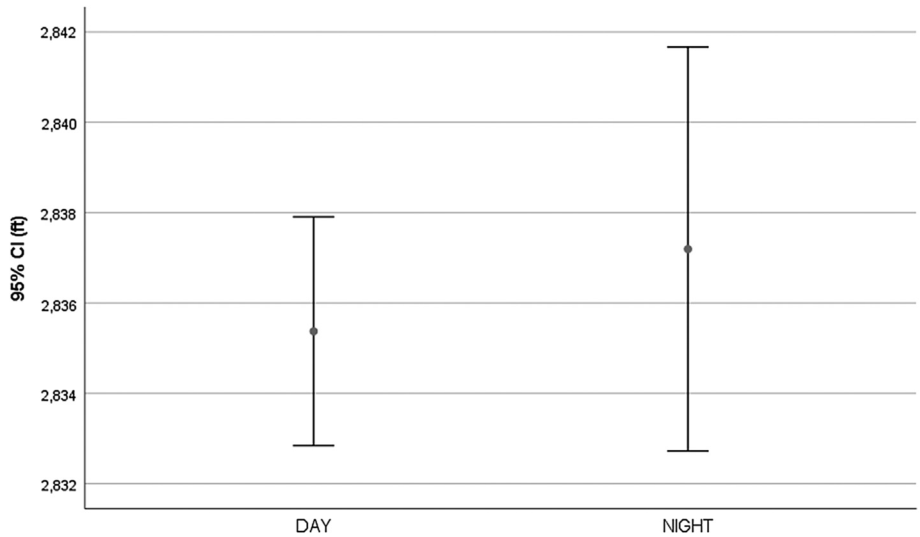


Figure 6. Error bar (or box plot) for day and night landing flares.

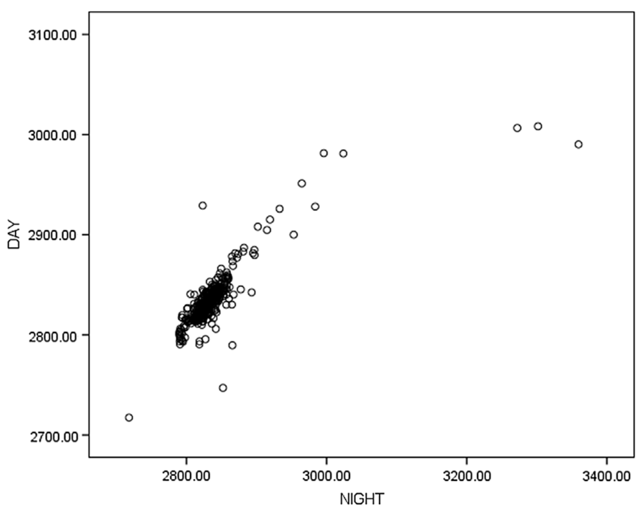


Figure 7. Paired samples correlation of day and night landing flares (feet).

least effective cue. In this scenario, participants initiated the flare 188.99 ft (57.60 m) above the optimal flare altitude. The interaction of cues suggested that more is not better. In other words, adding pictorial cues did not improve altitude perception. Instead, the crucial factor was the presence of the runway cue. The mean difference in observed–optimal flare altitude for scenarios that included the runway was 107.78 ft (32.85 m) compared with 174.68 ft (53.24 m) for scenarios that did not. That is a significant difference of almost 70 ft with obvious implications to aviation safety.

In the second study pictorial cues were indirectly manipulated by using daytime and nighttime landings. At night, the REIL delineation made the runway the most salient pictorial cue. Although obscure, it is impossible to eliminate the possibility that participants also used the horizon. Thus, the use of H × R × F cues in the daytime was compared with R × H cues at night. This comparison failed to detect significant differences in perceptions of altitude AGL. Since the runway was the most salient nighttime pictorial cue, it is plausible that participants used the runway to determine AGL for both daytime and nighttime landings. While not significant, an analysis was conducted to determine why the variance for nighttime landings was higher. The ad hoc comparisons revealed that the culprit was the first nighttime landing. It was significantly higher than the second and third landings. In Study 2, as in reality, participants practiced daytime landings. The unexpected first nighttime landing may have triggered a startle response.

Limitations and Future Research

The most obvious limitations are inherent to simulation fidelity and acuity. The night and daytime approaches and landings were simulated in real time. However, due to technical limitations, it was not possible to manipulate pictorial cues in real time. Instead, static screenshots, representing pictorial cue permutations, were transitioned at a rate that created the effect of a normal approach and landing. However, the loss of fidelity was consistent across permutations and participants and, therefore, represented systematic unexplained variance. In addition, the aim was to detect relative, not absolute, flare altitudes among pictorial cues. A related possible limitation was the acuity of pictorial cues. However, FS2000 introduced 3D elevations that permitted scenery to adjust to elevation, and no issues with 3D acceleration or acuity were experienced. Nevertheless, researchers are encouraged to replicate both studies using real-time manipulation of pictorial cues and higher-end visual graphics. Researchers are also encouraged to replicate the studies using intermediate- and expert-level pilots. It is possible that the use of pictorial cues is mediated by experience level.

Conclusion and Application

Inconsistent landing flare instructions and comments like "just about now begin to flare" or "you're too high!" increase the frustration of not knowing when to initiate the landing flare (Benbassat & Abramson, 2002c; Bramson, 1982; Penglis, 1994). Penglis (1994) concluded that "the reason no student knows where the ground begins is because the method we use to teach landings to students is wrong and does not work" (p. 91). Findings from this paper suggest that the runway is the most important pictorial cue for assessing height AGL prior to initiating the landing flare. This conclusion is based on two different studies with corroborative evidence. Some suggest that flaring the aircraft at the proper altitude depends on visual cues like "the size of familiar objects near the landing area, such as fences, bushes, trees, hangars, and even sod or runway texture" (FAA, 2016, pp. 8–6; Kershner, 1998) and "additional depth cues (such as 3D buildings and an explicit horizon)" (Palmisano et al., 2006, p. xi). Findings from this study suggest that more is not better. Teaching student pilots to focus on multiple pictorial cues may lead to confusion, suboptimal performance, and incidents.

This paper did not endeavor to study the mechanisms by which pictorial cues provide meaningful information. Instead, an applied science approach was taken in the hopes of advancing flight training and safety. Regardless of whether pilots use TTC, absolute angle, or absolute distance to determine what "appears to be" 10–20 ft from the ground, they must use pictorial cues outside the cockpit. Findings from this study suggest that, at least for novice pilots, the most important pictorial cue is the runway. Thus, flight instructors are encouraged to focus on the runway and airport authorities are encouraged to add or improve daytime and nighttime runway delineation.

References

Benbassat, D. (2005). Perceived Importance of the visual cues to the timing of the landing flare. *Journal of Aviation/Aerospace Education & Research, 15*(1), 31–39. https://doi.org/10.15394/jaaer.2005.1512

Benbassat, D., & Abramson, C. I. (2002a). Landing flare accident reports and pilot perception analysis. *The International Journal of Aviation Psychology, 12*, 137–152. https://doi.org/10.1207/S15327108IJAP1202_3

Benbassat, D., & Abramson, C. I. (2002b). Errorless discrimination learning in simulated landing flares. *Human Factors and Aerospace Safety, 2*, 319–338.

Benbassat, D., & Abramson, C. I. (2002c). General aviation landing flare instructions. *Journal of Aviation/Aerospace Education & Research, 11*, 31–38. https://doi.org/10.15394/jaaer.2002.1298

Benbassat, D., Abramson, C. I., & Williams, K. W. (2002). Comparative approach to pilot error and effective landing flare instructions. *International Journal of Comparative Psychology, 15*, 249–255.

Benbassat, D., Williams, K. W., & Abramson, C. I. (2005). General aviation leveloff, roundout, and accident rate analysis. *The International Journal of Aviation Psychology, 15*, 189–203. https://doi.org/10.1207/s15327108ijap1502_5

Blanca, M.J, Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Bramson, A. (1982). *Make better landings*. Van Nostrand Reinhold.

Butcher, R. (1996). *Private pilot flight training manual*. Skyroamers.

Cessna (Producer). (2001). *Cleared for takeoff*. [Motion picture]. (Available from King Schools, 3840 Calle Fortunada, San Diego, CA, 92123-1825).

Christy, J. (1991). *Good takeoffs and good landings* (2nd ed.). Tab Books.

Entzinger, J. O., & Suzuki, S. (2008). Visual cues in manual landing of airplanes. *Proceedings of KSAS – JSASS Joint International Symposium on Aerospace Engineering*, 388–395.

Federal Aviation Administration. (2014). *Airplane flying handbook (Rep. No. FAA-H-8083–3A)*. Skyhorse Publishing.

Federal Aviation Administration. (2016). *Airplane flying handbook (Rep. No. FAA-H-8083–38).* Skyhorse Publishing.

Gibson, J. J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *The American Journal of Psychology, 68*, 372–385. https://doi.org/10.2307/1418521

Goldstein, E. B. (2007). *Sensation and perception* (7th ed.). Thomson Wadsworth.

Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass

Jorgensen, C. C., & Schley, C. (1990). A neural network baseline problem for control of aircraft flare and touchdown. In M. W. Miller & R. S. Sutton (Eds.), *Neural networks for control* (pp. 403–425). MIT Press.

Kershner, W. K. (1998). *The student pilot's flight manual* (8th ed.). Iowa State University Press.

Langewiesche, W. (1972). *Stick and rudder*. McGraw-Hill.

Love, M. C. (1995). *Better takeoffs & Landings*. Tab Books/McGraw-Hill.

National Transportation Safety Board. (2011). *Review of US civil aviation accidents, 2007–2009*. NTSB.

National Transportation Safety Board. (2014). *Review of US civil aviation accidents, Calendar Year 2011(NTSB/ARA-14/01)*. NTSB.

Norman, G. R., & Streiner, D. L. (2014). *Biostatistics: The bare essentials* (4th ed.). People's Medical Publishing House.

Olson, R., & Austin, J. (2006). Performance-based evaluation of flight student landings: Implications for risk management. *The International Journal of Aviation Psychology, 16*, 97–112. https://doi.org/10.1207/s15327108ijap1601_5

Palmisano, S. A., Favelle, S., Prowse, G., Wadwell, R., & Sachtler, B. W. (2006). *Investigation of visual flight cues for timing the initiation of the landing flare*. ATSB Research and Analysis Report (B2005/0119). Australian Transportation Safety Bureau.

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *HSR: Health Services Research, 34*(5), 1189–1208.

Penglis, G. M. (1994). *The complete guide to flight instruction*. Rainbows Books.

Qian, J., Li, J., Wang, K., Liu, S., & Lei, Q. (2017). Evidence for the effect of depth on visual working memory. *Scientific reports, 7*(1), Article 6408. https://doi.org/10.1038/s41598-017-06719-6

Riordan, R. H. (1974). Monocular visual cues and space perception during the approach to landing. *Aviation, Space, and Environmental Medicine, 45*, 766–771.

Sheskin, D. J. (2004). *Handbook of parametric and nonparametric statistical procedures* (3rd ed.). Chapman & Hall/CRC.

History
Received April 17, 2020
Revision received November 15, 2020
Accepted February 9, 2021
Published online May 5, 2021

Acknowledgments
We would like to acknowledge our undergraduate research assistants, Shalagh Frantz, Chad Horstman, Mark Longfellow, Ross May, and the undergraduate students at Ohio Northern University who volunteered to participate. This study would have been impossible without their contribution and dedication. We would also like to thank William Albery and Charles D. Goodyear for their expert advice.

Publication Ethics
The opinions and assertions expressed herein are those of the authors and do not necessarily reflect the official policy or position of the Uniformed Services University or the Department of Defense.

ORCID
Danny Benbassat
https://orcid.org/0000-0001-5878-3667

Danny Benbassat
Department of Physical Medicine & Rehabilitation
School of Medicine
Uniformed Services University of the Health Sciences (USUHS)
4301 Jones Bridge Road
Bethesda, MD 20814
USA
dbenbassat@yahoo.com

Danny Benbassat (PhD) is a Lieutenant Commander with the United States Public Health Service Commissioned Corps and is currently serving as Assistant Professor in the Department of Physical Medicine and Rehabilitation, in the School of Medicine, at the Uniformed Services University in Bethesda, MD. He has extensive academic and industry aviation human factors experience and holds a commercial pilot certificate with an instrument rating. Prior to entering active duty, he supported the FAA in managing high-fidelity human-in-the-loop NextGen simulations and directed the Human Factors Laboratory at Ohio Northern University.

Ting Dong (PhD) is Research Associate Professor in the Center for Health Professions Education at the Uniformed Services University of the Health Sciences. She specializes in cognitive science of teaching and learning and quantitative methodologies for education. Her current research interests include learning processes and outcomes of students during medical school and long-term career outcomes of medical graduates.

Abigail Konopasky (PhD) is Assistant Professor in the Health Professions Education program in the School of Medicine at Uniformed Services University, where her teaching and research focus on linguistic analyses of learner and instructor agency, qualitative and mixed methods, and the psychology of teaching and learning. Her current research projects use functional linguistic methods and a mixed-methods lens to approach the areas of diversity in medicine and academic authorship.

Original Article

Why Learning Opportunities From Aviation Incidents Are Lacking

The Impact of Active and Latent Failures and Confidential Reporting

Sebastian Sieberichs and Annette Kluge

Department of Work, Organizational, and Business Psychology, Ruhr University Bochum, Germany

Abstract: The rising trend of fatal aircraft accidents since 2018 suggests a limited safety capability of airlines in terms of learning from incidents (LFI). We evaluated 2,208 voluntary incident reports from commercial European pilots using qualitatively driven mixed methods to investigate LFI "bottlenecks." The results showed that the report frequency depends on the type of pilots' active failure causing the incident (performance-based errors, judgment and decision-making errors and violations). Learning opportunities were lacking, especially for incidents caused by pilots' inadequate decision-making. Confidential reporting has positive effects on LFI, as these reports contained more information about latent failures. Furthermore, we identified several latent failures that are risk factors for certain unsafe acts. Our results may support airlines in various LFI activities.

Keywords: learning from incidents (LFI), active failures, latent failures, risk factors, confidential reporting

The latest safety report of the International Civil Aviation Organization (ICAO) shows an increasing accident rate since 2018, after a steady decline in the previous years (ICAO, 2019). In 2018, 11 fatal accidents occurred in scheduled commercial operations – the highest number in 5 years (ICAO, 2019). These disturbing statistics show that, despite substantial safety management and organizational learning processes, (preventable) accidents continue to occur, claiming lives, causing financial losses, and impairing the competitiveness of airlines (Stemn et al., 2018). One explanation for this trend may be that in the past, airlines have failed to learn their necessary lessons from accidents, as well as from minor incidents (Drupsteen et al., 2013; Stemn et al., 2018).

The "Value" of Incidents

The latest issue of the *World Safety Journal* describes the management of safety as an integral part of any organization (Bo, 2020). To maintain safety in environments characterized by change and insecurity, organizations need a "safety capability" that includes identifying and controlling the system's destabilizing threats and continually adapting operational routines (Griffin et al., 2015). In particular, high-reliability organizations (HROs), such as airlines, require a constant awareness of emerging threats and of factors that threaten this understanding (Hayes & Maslen, 2015; Weick & Sutcliffe, 2007). Airlines also belong to a class of organizations in which learning from fatal accidents only is not a sufficient strategy (Hayes & Maslen, 2015). This requires a substantial ability and willingness for organizational adaptation, so that in the course of learning processes a change in declarative and procedural knowledge can take place (Argote, 2012; Fiol & Lyles, 1985). Organizational learning is intended to change the behavior of organization members by learning from mistakes (*single-loop learning*) or by modifying the values and norms underlying the behavior (*double-loop learning*; Argyris & Schön, 1996; Putz et al., 2013). The ability to convert experiences from past incidents into behavior and practices in order to prevent similar events in the future can be described as "learning from incidents" (LFI; Drupsteen & Guldenmund, 2014; Jacobsson et al., 2011). The ICAO (2013) defines an incident as an "occurrence, other than an accident, associated with the operation of an aircraft which affects or could affect the safety of operation". Since LFI is possible "regardless of the severity of the consequences" of an incident (Drupsteen & Guldenmund, 2014, p. 83), "learning from weak signals" (precursor signals contributing

https://doi.org/10.1027/2192-0923/a000204

to anticipate occurrences) can be understood as an even more proactive approach to enable organizations to learn successfully (Brizon & Wybo, 2009; Drupsteen & Wybo, 2015). For example, the use of an unapproved departure route would be an incident; the programming of an unapproved departure route during flight preparation, favored by a similarity of the departure route designations, would instead be called a weak signal, if the error was corrected in time. Drupsteen and Wybo (2015) formulate "dealing with everything that may be wrong, from weak signals to incidents" (p. 35) as a fundamental principle of HROs and as an important prerequisite for organizations to learn from their experiences.

In this study, we use the term "incident" as an umbrella term for all types of incidents, near-misses, or weak signals that can provide input for learning and affect the safety of an organization (Drupsteen & Guldenmund, 2014; Rasmussen et al., 2013).

In most cases, incidents are caused by human errors or violations (Wiegmann & Shappell, 2017a). More recent views on human error extend this view to the effect that errors are indeed symptoms of "trouble deeper in the system" of an organization (Dekker, 2006, p. 18). Errors can be defined as "a deliberate action (or the deliberate omission of an action) characterized by the unintended failure to achieve personal goals and/or the unintended deviation from organizational norms and goals which could have been avoided by alternative behaviors of the acting person" (Putz et al., 2013, p. 513). Violations are also intentional acts, but lead to a deliberate, intentional non-compliance with known rules, procedures, or organizational norms (Reason, 2016). To initiate error-related learning processes, the detection of errors in the course of collecting information, the so-called learning product, is an essential step to be able to learn from them (Argyris & Schön, 1996; Cannon & Edmondson, 2001).

Six consecutive phases can be distinguished in the LFI process: "Reporting Incidents, Investigating Incidents, Developing Incident Alerts, Disseminating Information, Contextualizing Information and Implementing Actions" (Littlejohn et al., 2017, p. 82). Even one phase executed improperly can lead to learning becoming ineffective or failing to occur (Drupsteen & Hasle, 2014). Above all, the phases Reporting Incidents and Investigating Incidents have been identified as bottlenecks in several studies (cf., e.g., Drupsteen et al., 2013; Drupsteen & Hasle, 2014; Stemn et al., 2018). We will therefore investigate these two phases more closely in this study.

Reporting Incidents

Within the framework of organizational learning processes, data are derived, for example, from accident investigations, flight data monitoring, crew checks during scheduled flights, or line operations safety audits (LOSA), where safety-relevant findings are generated by various methods such as cockpit observations or crew interviews (Helmreich et al., 2017). The traditional data source for the LFI process of airlines is, however, the written report by an organization member, usually a pilot, about an incident (Margaryan et al., 2017). An aviation safety reporting system was already introduced by the National Aeronautics and Space Administration (NASA) in the mid-1970s (NASA, 1976). The objectives described at that time, namely, to create a reporting system for all members of the organization in which data are stored, evaluated as part of operational routines, and communicated to various stakeholders, continue to form the basis of reporting systems implemented in almost all high-reliability sections, such as nuclear power technology or health care (NASA, 1976; Van der Westhuizen & Stanz, 2017).

Safety management systems (SMS) include a systematic approach to managing safety, including the necessary organizational structures, accountabilities, policies, and procedures (ICAO, 2018b). Within this framework incident reports are used as a source of data to identify risks, develop mitigation measures, and monitor safety (Rasmussen et al., 2013; Van der Westhuizen & Stanz, 2017). The "stories" contained in the reports help operating staff to build their "safety imagination" and evaluate the safety of decisions and are also relevant within the framework of "story-based" learning (Hayes & Maslen, 2015).

Incident reporting can be considered to be a form of change-oriented safety citizenship behavior (SCB; Conchie, 2013). SCB is influenced by person-related antecedents, such as affective engagement and psychological ownership, and situation-related antecedents, such as the safety climate within the organization (Curcuruto & Griffin, 2018; Parker et al., 2010). The safety climate encompasses shared perceptions of safety policies, procedures, and practices among organization members and, in addition to safety behavior mediated by individual motivation, influences forms of safety participation at the discretion of the individual, such as incident reporting (Griffin & Curcuruto, 2016; Griffin & Neal, 2000; Zohar, 2003). The assessment of climatic aspects in relation to the LFI can be conceptualized, for example, with the influence of various environmental factors on learning levels when considering the *error-related learning climate* (Putz et al., 2013).

In the context of LFI, safety-cultural aspects are often discussed as influencing factors that guide the behavior of organizations and their members through underlying assumptions and values (Curcuruto & Griffin, 2018; Reason, 1998). In a learning culture, errors and the resulting incidents are accepted and explicitly seen as an opportunity to learn from (Littlejohn et al., 2014). To achieve this, it is a fundamental requirement that incidents are reported by organization members (Reason, 1998). Incident reporting is encouraged by a *just culture*, which includes an atmosphere of trust, where employees are encouraged and even rewarded to share safety information, but where a clear distinction is also made between acceptable and unacceptable behavior (Dekker, 2018).

Impairments in climatic and cultural aspects, such as a lack of trust and openness, but also fear and shame on the part of organization members, are some of the factors that limit safety participation in terms of incident reporting (cf., e.g., Curcuruto & Griffin, 2018; Drupsteen & Guldenmund, 2014; Gilbey et al., 2016; Jausan et al., 2017; Zabari & Southern, 2018). In order to reduce the influence of these hindering factors and increase the reporting rate, confidential reporting systems are widely implemented (Jausan et al., 2017; Langer, 2016; Merry & Henderson, 2017).

Too Few Incidents Are Reported

There are numerous reasons why incidents are not reported (cf. Jausan et al., 2017). In this study, we address the results of a survey conducted by Sieberichs and Kluge (2017) involving commercial pilots, who observed that the probability of an incident being reported depended on the type of incident. We thereby also follow a recommendation of Hayes and Maslen (2015) to focus further research on the types of incidents reported. This is relevant in the context of LFI, as less frequently or unreported types of incidents limit the learning opportunities that arise from them in the course of single-loop learning and reduce the accuracy of the site risk assessment (Argyris & Schön, 1996; Drupsteen & Guldenmund, 2014; Stemn et al., 2018). For safety management tasks, it is also relevant whether the frequency of various incidents differs in flight phases and route segments (Stolzer et al., 2015; Wheeler et al., 2019). Therefore, we ask:

Research Question (RQ1): Are there incidents that are reported less frequently depending on their causal unsafe acts?

Research Question (RQ1.1): Does the frequency of reported incidents differ between flight phases and route segments depending on their causal unsafe acts?

Investigating Incidents

When investigating incidents, the "immediate and underlying causes of the incident" should be determined (Littlejohn et al., 2017, p. 82). A frequently applied linear framework for investigation is the Swiss cheese model, which serves as the basis for the Human Factors Analysis and Classification System (HFACS; Reason, 1990; Wiegmann & Shappell, 2017b). In this model, accidents or incidents are caused by a chain of organizational and personal factors and are distinguished into *mishap-level factors*, such as *organizational influences* and *unsafe supervision*, and *person-level factors*, such as *preconditions* and *unsafe acts*. Organizational influence, unsafe supervision and preconditions represent *latent failures*, whereas unsafe acts are considered *active failures* (Littlejohn et al., 2017; Wiegmann & Shappell, 2017b). In contrast to this linear view, the system-theoretic accident model and processes (STAMP) approach, incorporates uncertain interactions of different system components and treats safety more as a dynamic control problem (Leveson, 2015).

Even though the Swiss cheese model has been criticized for not being able to capture the real world because it is too static and not specified enough, advantages of the system have been repeatedly highlighted in civil, commercial, and military aviation, but also in other organizations such as hospitals (Cohen et al., 2015; Hollnagel et al., 2006; Larouzee & Le Coze, 2020; Sunaryo et al., 2019). HFACS has also emerged as a reliable system in the framework of incident or weak signal analysis (e.g., Lee et al., 2017; Li & Harris, 2006; Madigan et al., 2016; Miranda, 2018; Munene, 2016). Miranda (2017), for example, was able to identify latent failures that are particularly conducive to certain types of unsafe acts when evaluating major accidents in a military context.

In this study we summarize errors and violations as immediate incident causes with the term "unsafe acts" (cf. Littlejohn et al., 2017; Wiegmann & Shappell, 2017b).

Too Little Information About the Incident Is Given

As stated earlier, confidential reporting systems are widely used to increase the reporting rate (Langer, 2016; Merry & Henderson, 2017). Since critical information about an incident is sometimes withheld in reports (Jausan et al., 2017), in this study we ask whether the use of confidential reporting depends on the type of unsafe act that caused the incident and if confidential reporting also increases the level of information about latent failures. With respect to the aforementioned results of Sieberichs and Kluge (2017), we ask whether the level of information about latent failures in reports depends on whether an error or violation caused the incident.

Research Question (RQ2): Are there incidents that are more frequently reported confidentially depending on their causal unsafe acts?

Research Question (RQ2.1): Do confidential reports contain more information about latent failures than non-confidential reports?

Research Question (RQ2.2): Does the level information about latent failures differ in reports where errors or violations caused the incident?

Latent Conditions Are Not Identified

The identification of latent failures is important to prevent the likelihood of reoccurrence of similar incidents, to facilitate double-loop learning, and to improve safety imagination (Argyris & Schön, 1996; Drupsteen & Guldenmund, 2014; Drupsteen & Hasle, 2014; Hayes & Maslen 2015; Madigan et al., 2016). In addition, identifying the complexity of an incident is important for selecting the necessary learning solutions (Littlejohn et al., 2017). To support airlines in identifying underlying incident causes, we will investigate whether there are latent failures that can be considered risk factors for certain unsafe acts.

Research Question (RQ3): Are there latent failures that are risk factors for the various unsafe acts and the incidents they cause?

With these six research questions we take up three reasons given by Drupsteen and Hasle (2014) on why organizations are not effectively learning from incidents.

Method

Data Selection

To answer the research questions, we evaluated voluntary written reports from pilots to their airline (originating situation), which contained the description of incidents caused by pilot errors or violations. These were exported from the digital safety database of a European airline operating short- and long-haul flights. The evaluation of reports corresponds to a common procedure of the airline in the context of various SMS activities and is in line with its policy. A permission from the airline for evaluation and publication of the data has been obtained.

The database contained no information about the identity of the author. In addition, in the course of an "absolute anonymization," neither the date of the report nor the author's rank was recorded and the information regarding the aircraft type on which the incident occurred was summarized in short- and long-haul (cf. Medjedovic & Witzel, 2010).

The exported reports included all reports that were stored in the database between 2002 and summer 2019 listed under the internally used designation "pilot error." In addition to the description of the incident in text form, each report contained the attributes flight phase (ground, take-off, flight, approach/landing), route segment (short- or long-haul), and report type (confidential or non-confidential). The data export included 2,208 incident reports.

Research Design

The methodology and the step-by-step methods to be applied were defined in advance in a binding research plan (Mayring, 2020). A sequentially linked, qualitatively driven mixed-method context analysis design was defined as methodology (cf. Kansteiner & König, 2020; Kelle, 2019; Mayring, 2020). A method triangulation of qualitative and quantitative document analyses was used (cf. Baur et al., 2017). The aim of the qualitative steps was to prepare a structuring description of the documents that were to be classified according to theoretically meaningful order aspects (Mayring, 2015). The aim of the quantitative steps was to answer the research questions by using descriptive and inferential statistical evaluation procedures. The description dimensions were quantitative variables transformed from the coded reports and the mentioned attributes. We justify the methodology with the complementarity of the procedures used, since we expected further clarification of the qualitatively obtained results through quantitative evaluation steps (Kansteiner & König, 2020; Schoonenboom & Johnson, 2017). According to Kuckartz (2019) and Mayring (2015), qualitatively oriented classifications are also a good starting point for quantitative analyses. We consider reports to be an adequate data basis for answering the research questions, as previous experience with a comparable methodology is available, for example, from Miranda (2018). In terms of scientific theory, we base the methodology on a pragmatic position that follows the paradigm of dialectical pluralism (Baur et al., 2017; Mayring, 2007).

Procedure: Qualitative Steps

For the classification of the reports, a content-analytical process model according to Mayring was chosen as the analysis technique (Mayring, 2015, p. 62). In accordance with a content analytic communication model (Mayring, 2015, p. 59), the description of the subject matter included in the reports was determined as the direction of the analysis. A complete report was defined as the coding unit and context unit. The analysis unit consisted of the 2,208 reports, which were consecutively evaluated. Although the STAMP approach allows for the modeling of nonlinear relationships and is more appropriate in complex socio-technical systems, we use HFACS because we expect it to be more reliable due to its taxonomic structure and this system has proven to be more useful when analyzing a larger number of case studies (cf. Salmon et al., 2012).

The assignment of categories to the reports was done deductively with categories of the Dod-HFACS 7.0 from the "Dod-HFACS 7.0 Guide" (Air Force Safety Center, 2016). To facilitate the use of these categories that were designed for the military context, Scott Shappell has provided us on request with an overview of anchor examples for the civil context. DoD-HFACS 7.0 contains Mishap-Level Factors with categories for Organizational Influences and Supervision and Person-Level Factors with categories for Preconditions and Unsafe Acts (Air Force Safety Center, 2016).

To test the suitability of the category system, 200 randomly selected reports were initially coded in a pilot phase. Since the Mishap-Level Factors turned out to be unsuitable for the coding units, we only used the Person-Level Factors. In the course of this pilot phase, anchor examples and coding rules were added to the categories and definitions given in the "DoD-HFACS 7.0 Guide," and a separate analytical scheme in the form of a coding manual was created to ensure a stable perspective for researchers during analysis (cf. American Psychological Association, 2020). Following the recommendation of Jacobsson et al. (2009), the coding rules were formulated in such a way that factors that "were not directly stated in a (...) report, but that can be deduced following the description of the event" (p. 197) could also be considered.

The coding manual contained 13 categories (including three superordinate categories) within the factor Unsafe Acts and 61 categories within the factor Preconditions.

The category Procedure Not Followed Correctly is an example for an unsafe act within the superordinate category Performance-Based Errors. This category describes a factor when a procedure is performed incorrectly or accomplished in the wrong sequence. It is assigned if a procedural error is explicitly mentioned in the report or the report contains a text passage to which the definition applies (Anchor examples: "We did not apply the Oceanic-Crossing-Procedure correctly and did not check the updated route clearance" [Report 207]; "Cleared for SOBRA 3L, we programmed SOBRA 1S, briefing for SOBRA 3L, we flew SOBRA 1S" [Report 505]).

The category Complacency is an example for a precondition. This category describes a factor when the individual has a false sense of security, is unaware of, or ignores hazards and is inattentive to risks. It is assigned if complacency or an equivalent term is explicitly mentioned in the report or the report contains a text passage to which the definition applies (Anchor examples: "I will never again delegate important call-outs to other than cockpit crew members! Complacency at its best!" [Report 1771]. "The main factor was Complacency in the clearance review" [Report 1311]).

Each coding unit was assigned with one category of the factor Unsafe Acts and the applicable categories of the factor Preconditions. This procedure was repeated for the entire evaluation unit by the main coder. To ensure consistency in the analysis process in terms of developing a stable perspective of the researchers, we defined a period of 12 consecutive weeks for the coding process. The time units for content analysis were limited to 45 min. Between two and four units per day were carried out on at least 5 days per week.

We will illustrate the procedure with the following example:

> Report 999:
>
> Ramp agent (RA): "Please apply brake, pushback completed." I answer: "brake set" after looking out of the window to see if the plane is standing still. The RA replies: "Then the yellow light on the nose gear seems to be defect." I look at the triple indicator and see no brake pressure! After a look at the Parking Brake Selector I see that I have indeed not set the brake. I cannot explain why. Probably classic distraction and fatigue. Fortunately, we had a very alert RA.

The unsafe act Procedure not followed correctly was assigned to this report, because the push-back procedure stipulates that verbal confirmation of the set parking brake may only be given after checking the *triple-indicator*. Distraction and Fatigue were assigned as Preconditions.

Of the 2,208 exported reports, 464 (21.01%) were excluded from further evaluation:

- 278 reports that were incomplete or in the wrong category;
- 113 reports about errors of others (e.g., the author was travelling as a passenger and describes that the pilots did not remove ice from a wing that clearly had to be de-iced);
- 29 reports with measurable damage (these reports are mandatory);
- 25 duplicate reports; and
- 19 reports in which the author was asked to write the report (background: If extreme deviations are detected during flight data monitoring, the airline can require the causing pilot to write a report).

Thus, the description field consisted of 1,744 incident reports.

After the coding of all reports, a second coder was introduced into the coding manual. Both coders jointly coded 14 boundary cases in which the assignment of more than one category of the factor Unsafe Acts would be possible. An example of a boundary case is a report in which a pilot did not correctly apply the *wind-shear-escape-maneuver* and in addition oversteered the aircraft. In this case the category Procedure Not Followed Correctly was selected, because this unsafe act would have the largest relative contribution to the most credible accident scenario, if the same incident happened again (cf. ICAO, 2018a). To determine inter-rater reliability for both coders, a second evaluation unit was created from randomly selected 10% of the reports, which was coded by the second coder with the described content-analytical model using the coding manual.

The main coder is an active civil airline pilot. Moreover, he has experience in the training of pilots and works in the safety department of a major European airline, where he is dealing with risk assessment, root cause analysis, and the evaluation of safety-related reports. The second coder works full-time in the safety department of the aforementioned airline and is responsible for the processing of all safety-related reports. Both coders have attended a training course on aviation accident investigation, which included training in HFACS. Since specific instructions are recommended for a reliable use of HFACS, the coders have completed a preparatory, web-based HFACS training (Clemson University, 2018a, 2018b; Ergai et al., 2016).

Procedure: Quantitative Steps

To prepare the qualitative results of the content analysis for a quantitative evaluation with SPSS (Version 26), the codes were transformed into binary variables. For each coding unit, a data set with 13 variables of the factor Unsafe Acts and 61 binary variables of the factor Preconditions was created. In addition, the aforementioned attributes of the reports from the database export were assigned to the data records as categorical variables. The quantitative analysis was based on 1,744 data sets.

To determine the frequency of reported incidents depending on the causal unsafe acts, the absolute frequencies of the variables of the factor Unsafe Acts were calculated in response to RQ1. To answer RQ1.1 and RQ2, chi-square tests were calculated to investigate whether the observed frequencies of reports in the flight phases deviated from expected frequencies. To determine the level of information about latent failures in the reports, we conducted t tests for independent samples to answer RQ2.1 and RQ2.2. To answer RQ3, logistic regression analysis models for dichotomous dependent variables were calculated (cf. Eid et al., 2015): The variables of the factor Unsafe Acts were each defined as dependent variables, the variables of the factor Preconditions were independent variables that were simultaneously included in the model calculation.

Results

The incidents described in the reports occurred in four different flight phases: 289 incidents (17%) occurred on ground, 223 (13%) during take-off, 461 (26%) in flight, and 771 (44%) during approach or landing. Altogether 1,117 (64%) occurred on short-haul and 627 (37%) on long-haul flights. Of 1,744 reports, 705 (40%) were non-confidential and 1,039 (60%) confidential.

Qualitative Results: Result of the Content Analysis

Each coding unit (N = 1744) was assigned one category of the factor Unsafe Acts: Within the superordinate category Performance-Based Errors (n = 1,310) 76 coding units were assigned with the category Unintended Operation of Equipment, 43 with the category Checklist Not Followed Correctly, 733 with the category Procedure Not Followed Correctly, 431 with the category Over-Controlled/Under-Controlled Aircraft, 21 with the category Breakdown in Visual Scan, and six with the category Rushed or Delayed a Necessary Action. Within the superordinate category Judgment and Decision-Making Errors (n = 152), 56 coding units were assigned with the category Inadequate Real-Time Risk Assessment, four with the category Failure to Prioritize Tasks Adequate, 30 with the category Ignored a Caution/Warning, and 62 with the category Wrong Choice of Action During an Operation. Within the superordinate category Violations (n = 282) 113 coding units were assigned with the category Work-Around Violation, 130 with the category Widespread/Routine Violation, and 39 with the category Extreme Violation – Lack of Discipline.

Each coding unit (N = 1,744) was assigned with the applicable categories of the factor Preconditions. In total, the reports contained 2,691 text passages that were coded with a category of the factor Preconditions. Figure 1 shows how often each category was assigned. For example, 361 reports were coded with the category Not Paying Attention.

The following categories from the factor Preconditions of DoD-HFACS 7.0 could not be assigned to any coding unit: Psychological Problem, Turning/Balance Illusion – Vestibular, Temporal/Time Distortion, Substance Effects (alcohol, supplements, medications, drugs), Loss of Consciousness (sudden or prolonged onset), Trapped Gas Disorders, Evolved Gas Disorders, Hypoxia/Hyperventilation, Inadequate Adaptation to Darkness, Dehydration, Body Size/Movement Limitations, Physical Strength and Coordination (inappropriate for task demands), Vibration Affects Vision or Balance, External Force or Object Impeded an Individual's Movement, Seat and Restraint System Problems.

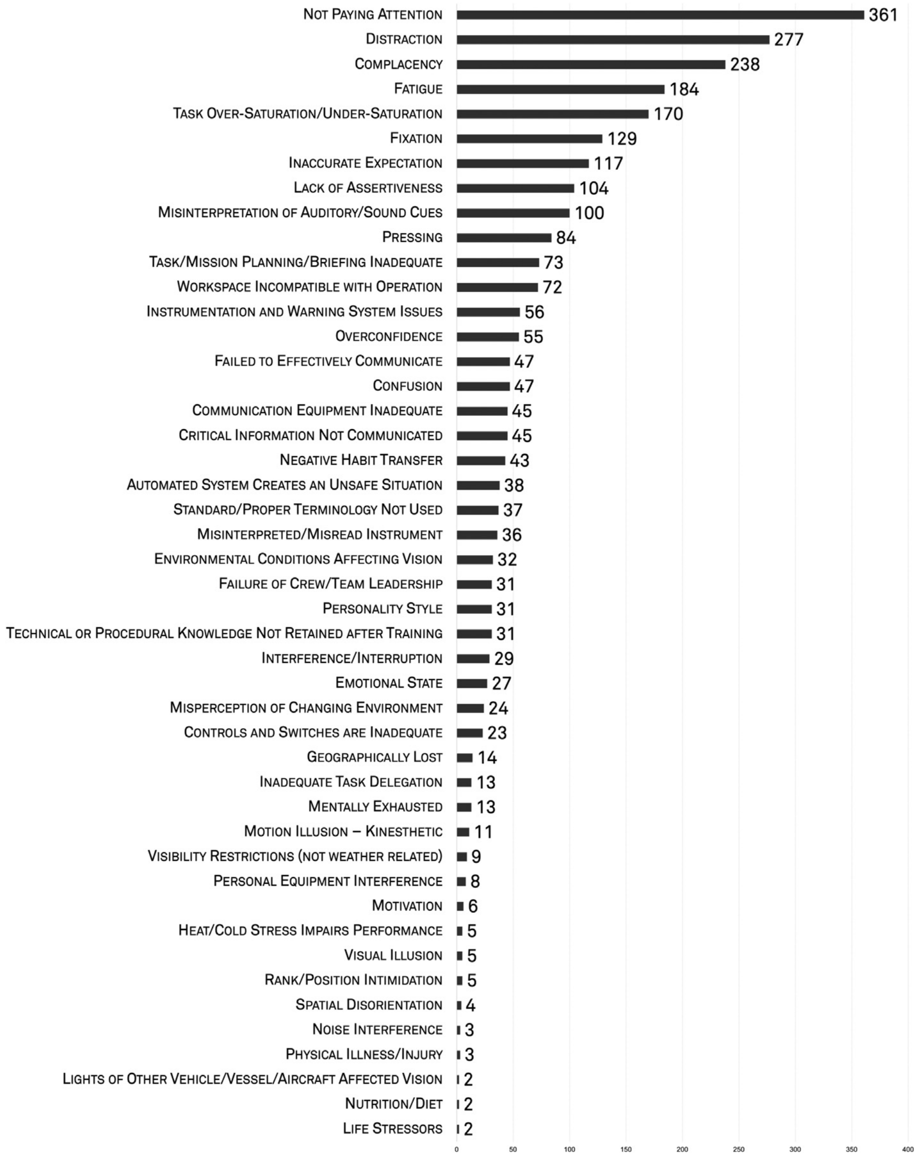


Figure 1. Frequency of preconditions categories.

Cohen's κ was calculated to determine inter-rater reliability for both coders and amounts to κ = .95. According to Landis and Koch (1977), this indicates an almost perfect and according to Altman (1990) a very good reliability.

Mixed Methods Results: Addressing the Research Questions

For a better overview, when answering research questions Q1.1 and Q2.1, we only report results with cell frequencies greater than 5 and at least small effects ($\varphi \geq 0.1$; Cohen, 1988). When interpreting the effect sizes, a small effect can be assumed for $\varphi = 0.1$ or $d = 0.2$, a medium effect for $\varphi = 0.3$ or $d = 0.5$ and a large effect for $\varphi = 0.5$ or $d = 0.8$ (Cohen, 1988). When interpreting pseudo-determination measures (Cox–Snell R^2 and Nagelkerke R^2), a model-fit can be considered acceptable with $R^2 > 0.2$ and good for $R^2 > 0.4$ (Backhaus et al., 2016).

RQ1: Are There Incidents That Are Reported Less Frequently Depending on Their Causal Unsafe Acts?

The frequency of reported incidents varied depending on the causal unsafe acts. The results showed that incidents caused by judgment and decision-making errors are reported less frequently. Incidents caused by judgment and decision-making errors (n = 152) accounted for less than 9% of all reports. Within this superordinate category,

Table 1. Unsafe acts by flight phases

Unsafe acts	Flight phases: On ground	Take-off	In flight	Approach	χ^2, φ
Checklist Not Followed Correctly	M	M	P	P	$\chi^2(3) = 62.25^*$, $\varphi = 0.19$
Procedure Not Followed Correctly	M	M	M	P	$\chi^2(3) = 260.15^*$, $\varphi = 0.39$
Over-Controlled/Under-Controlled Aircraft	M	P	P	M	$\chi^2(3) = 83.15^*$, $\varphi = 0.22$
Wrong Choice of Action During an Operation	P	M	P	M	$\chi^2(3) = 26.48^*$, $\varphi = 0.12$
Work-Around Violation	P	P	P	M	$\chi^2(3) = 88.59^*$, $\varphi = 0.23$
Widespread/Routine Violation	P	P	P	M	$\chi^2(3) = 162.92^*$, $\varphi = 0.31$

Note. If the observed cell frequencies are higher than the expected, a "P"(lus) is shown; vice versa, an "M"(inus). $^*p < .001$.

for example, there were only four reports of incidents where a failure to prioritize tasks adequately was the cause of the incident. About 16% of the reported incidents were caused by violations (n = 282). Within this superordinate category, extreme violations (n = 39) were the least frequently reported. Although performance-based caused more than half of the reported incidents, within this superordinate category there were only six reports of incidents caused by a rushed or delayed necessary action.

RQ1.1: Does the Frequency of Reported Incidents Differ Between Flight Phases and Route Segments Depending on Their Causal Unsafe Acts?

The frequency of reported incidents differed in different flight phases for some of the unsafe acts. The results of the chi-square tests (associations between unsafe acts and flight phases) are presented in Table 1.

If an incident was caused, for example, by an incorrectly followed procedure, the observed frequency of reports in the flight phase Approach was higher than the expected frequency. The opposite applied to the other flight phases.

The results of the chi-square tests (associations between unsafe acts and route segments) showed that the frequency of reported incidents in different route segments did not differ from the expected frequency depending on the causal unsafe acts.

RQ2: Are There Incidents That Are More Frequently Reported Confidentially Depending on Their Causal Unsafe Acts?

The frequency of confidentially reported incidents differed for incidents caused by widespread or routine violations. The results of the chi-square tests (associations between unsafe acts and report type) showed that the observed frequency of confidential reports on incidents caused by widespread or routine violations is higher than the expected frequency, $\chi^2(1) = 59.59$, $p < .001$, $\varphi = -0.19$.

RQ2.1: Do Confidential Reports Contain More Information About Latent Failures Than Nonconfidential Reports?

In nonconfidential reports (M = 1.22, SD = 0.89) fewer latent failures were reported than in confidential reports (M = 1.76, SD = 1.20). Nonconfidential reports contained −0.54 latent failures (95% CI [−0.64, −0.44]) less than confidential reports, $t(1729.26) = -10.73$, $p < .001$, $d = -0.51$.

RQ2.2: Does the Level of Information About Latent Failures Differ in Reports Where Errors or Violations Caused the Incident?

The level of information on latent failures differed between reports in which violations caused an incident (M = 2.00, SD = 1.35) and reports in which errors caused an incident (M = 1.46, SD = 1.04). Reports in which violations caused an incident contained 0.54 more latent failures (95% CI [0.40, 0.68]), $t(1742) = 7.56$, $p < .001$ $d = 0.45$.

RQ3: Are There Latent Failures That Are Risk Factors for Various Unsafe Acts and the Incidents They Cause?

We calculated 13 logistic regression models with all categories of the factor Unsafe Acts as dependent variables and all variables of the factor Preconditions as independent variables. Four models failed the omnibus test of the model coefficients and four models had an unacceptable model-fit. Therefore, we calculated logistic regression models with dependent variables summarized at the superordinate category level (Performance-Based Errors, Judgment and Decision-Making Errors and Violations). Tables 2 to 4 show the results of the regression calculations. For a better overview, only the independent variables with Wald value $p < .05$ are shown.

In all three models, the Hosmer–Lemeshow test of model quality was not significant.

The omnibus test for the model with dependent variable Performance-Based Errors was significant, $\chi^2(46) = 604.54$, $p < .001$. The Cox–Snell index was CS = .293 and the Nagelkerke index was NK = .434.

The omnibus test for the model with dependent variable Judgment and Decision-Making Errors was significant, $\chi^2(46) = 160.62$, $p < .001$. The Cox–Snell index was CS = .088 and the Nagelkerke index was NK = .197.

Table 2. Model summary for performance-based errors

Preconditions	*n* within	% within	*B*	*SE*	Wald	*df*	Exp(*B*)	95% CI for EXP(B) LL	95% CI for EXP(B) UL
Misinterpreted/Misread Instrument	34	2.60%	1.62	0.78	4.31	1.00	**5.05**	1.10	23.28
Not Paying Attention	330	25.19%	1.52	0.23	41.85	1.00	**4.55**	2.88	7.21
Misinterpretation of Auditory/Sound Cues	89	6.79%	1.28	0.38	11.60	1.00	**3.60**	1.72	7.51
Confusion	42	3.21%	1.25	0.54	5.27	1.00	**3.48**	1.20	10.08
Distraction	246	18.78%	1.22	0.23	26.95	1.00	**3.38**	2.13	5.34
Critical Information Not Communicated	33	2.52%	0.96	0.47	4.18	1.00	**2.62**	1.04	6.59
Fatigue	145	11.07%	0.55	0.24	5.32	1.00	**1.73**	1.09	2.75
Task Over-Saturation/Under-Saturation	118	9.01%	−0.46	0.22	4.46	1.00	0.63	0.41	0.97
Pressing	45	3.44%	−0.71	0.29	6.03	1.00	0.49	0.28	0.87
Technical/Procedural Knowledge Not Retained after Training	20	1.53%	−0.88	0.43	4.16	1.00	0.41	0.18	0.97
Overconfidence	15	1.15%	−1.13	0.38	8.97	1.00	0.32	0.16	0.68
Fixation	48	3.67%	−1.62	0.23	49.49	1.00	0.20	0.13	0.31
Complacency	89	6.79%	−1.72	0.18	87.67	1.00	0.18	0.12	0.26
Failure of Crew/Team Leadership	2	0.15%	−2.02	0.93	4.65	1.00	0.13	0.02	0.83
Personality Style	3	0.23%	−2.06	0.72	8.31	1.00	0.13	0.03	0.52
Lack of Assertiveness	14	1.07%	−2.55	0.36	50.35	1.00	0.08	0.04	0.16

Note. Displayed Wald values are significant at $p < .05$. *n within/% within* = frequency/percentage of precondition in incidents caused by performance-based errors; *B* = estimated regression coefficient; *SE* = standard error; Wald = Wald test; Exp(*B*) = odds ratio (values > 1 are in boldface as these preconditions increase the probability of a performance-based error); CI = confidence interval; LL = lower limit; UL = upper limit.

Table 3. Model summary for judgment and decision-making errors

Preconditions	*n*	% within	*B*	*SE*	Wald	*df*	Exp(*B*)	95% CI for EXP(B) LL	95% CI for EXP(B) UL
Lack of Assertiveness	30	19.74%	1.40	0.30	21.71	1.00	**4.06**	2.25	7.31
Technical/Procedural Knowledge not Retained After Training	7	4.61%	1.02	0.48	4.59	1.00	**2.78**	1.09	7.10
Fixation	28	18.42%	1.02	0.26	15.43	1.00	**2.76**	1.66	4.58
Task/Mission Planning/Briefing Inadequate	15	9.87%	0.89	0.34	6.86	1.00	**2.44**	1.25	4.76
Overconfidence	14	9.21%	0.89	0.36	5.91	1.00	**2.43**	1.19	4.96
Pressing	16	10.53%	0.76	0.33	5.27	1.00	**2.13**	1.12	4.06
Fatigue	8	5.26%	−0.90	0.39	5.25	1.00	0.40	0.19	0.88
Distraction	10	6.58%	−0.97	0.35	7.90	1.00	0.38	0.19	0.75
Misinterpretation of Auditory/Sound Cues	3	1.97%	−1.22	0.61	3.96	1.00	0.30	0.09	0.98
Not Paying Attention	7	4.61%	−1.69	0.41	17.35	1.00	0.18	0.08	0.41

Note. Displayed Wald values are significant at $p < .05$. *n within/% within* = frequency/percentage of precondition in incidents caused by judgment and decision-making errors; *B* = estimated regression coefficient; *SE* = standard error; Wald = Wald test; Exp(*B*) = odds ratio (values > 1 are in boldface as these preconditions increase the probability of a judgment and decision-making error); CI = confidence interval; LL = lower limit; UL = upper limit.

The omnibus test for the model with dependent variable Violations was significant, $\chi^2(46) = 420.87$, $p < .001$. The Cox–Snell index was CS = .214 and the Nagelkerke index was NK = .365.

Since the estimated regression coefficient cannot be interpreted meaningfully due to nonlinear relationships, we follow the recommendation of Best and Wolf (2010) and explain the results in terms of direction and strength of the odds ratios. As a summary of the results of the regression analyses, Figure 2 shows the odd-ratio values associated with the respective preconditions of the estimated regression coefficients with positive sign for the three dependent variables.

The preconditions (latent failures) presented can be considered as risk factors for the various unsafe acts and the incidents they cause. Except for fixation and lack of assertiveness, each precondition is only a risk factor for one unsafe act. For example, not paying attention increases the risk of an incident caused by a performance-based error by up to 4.55 times.

Table 4. Model summary for violations

Preconditions	*n* within	% within	*B*	*SE*	Wald	*df*	Exp(*B*)	95% CI for EXP (B)	
								LL	UL
Complacency	126	44.68%	1.99	0.18	120.69	1.00	**7.30**	5.12	10.41
Lack of Assertiveness	60	21.22%	1.26	0.28	20.03	1.00	**3.54**	2.04	6.16
Emotional State	14	4.96%	1.18	0.52	5.16	1.00	**3.27**	1.18	9.07
Personality Style	20	7.09%	1.12	0.49	5.30	1.00	**3.06**	1.18	7.91
Fixation	53	18.79%	1.06	0.24	19.85	1.00	**2.88**	1.81	4.59
Task Over-Saturation/Under-Saturation	38	13.48%	0.64	0.23	7.44	1.00	**1.89**	1.20	2.99
Misinterpretation of Auditory/Sound Cues	8	2.84%	−1.04	0.42	6.06	1.00	0.36	0.16	0.81
Distraction	21	7.45%	−1.05	0.27	15.32	1.00	0.35	0.21	0.59
Not Paying Attention	24	8.51%	−1.12	0.25	19.66	1.00	0.33	0.20	0.54
Critical Information Not Communicated	4	1.42%	−1.84	0.65	7.99	1.00	0.16	0.04	0.57

Note. Displayed Wald values are significant at $p < .05$. *n within/% within = frequency/percentage of precondition in incidents caused by* Violations; *B* = estimated regression coefficient; *SE* = standard error; Wald = Wald test; Exp(*B*) = odds ratio (values > 1 are in boldface as these preconditions increase the probability of a violation); CI = confidence interval; LL = lower limit; UL = upper limit.

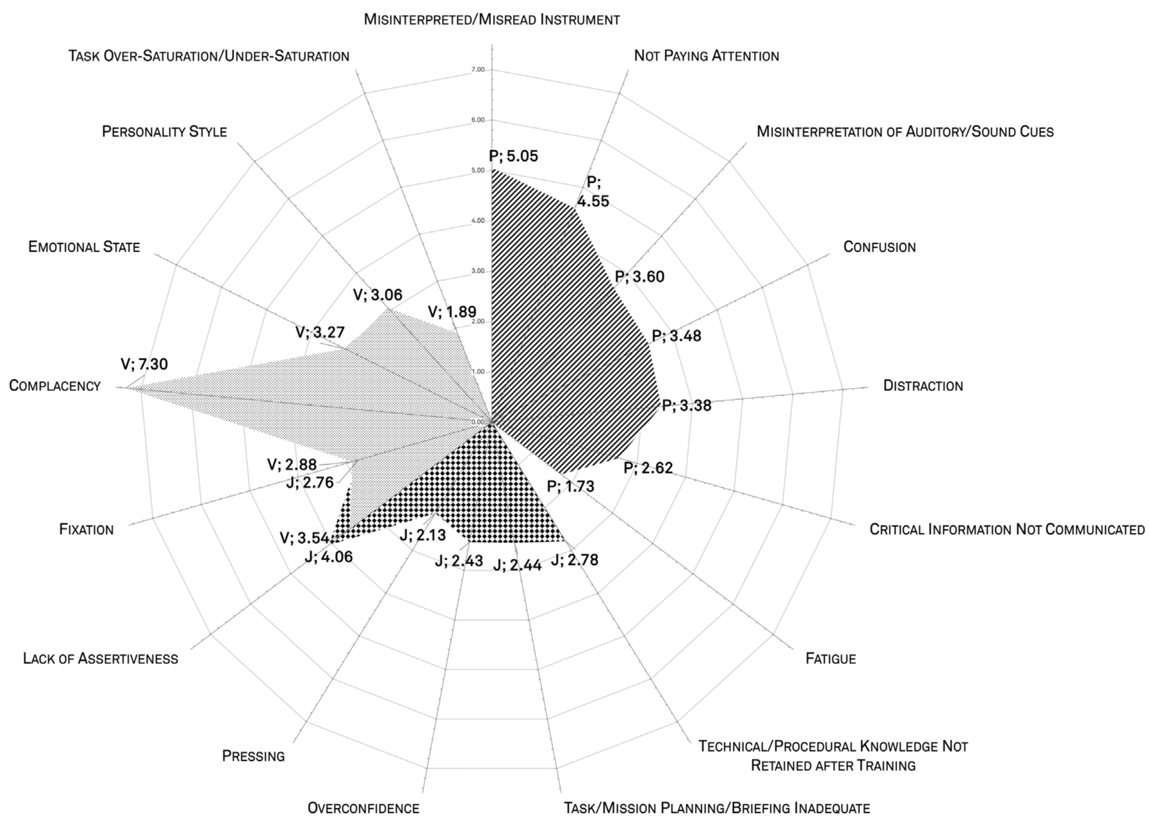


Figure 2. Risk factors for unsafe acts. The numbers shown correspond to the odd-ratio values (Exp(*B*)) of the preconditions in the three regression analyses; the letters indicate the various dependent variables (*P* = Performance-Based Errors, *J* = Judgment & Decision-Making Errors, *V* = Violations).

Discussion

The aim of this study was to further analyze why organizations cannot learn effectively from incidents. By looking at hindering factors in the Reporting Incidents phase, we found that airlines suffer from a lower number of reports on some types of incidents depending on the type of active failure, causing the incident (pilots' unsafe acts) – particularly with regard to judgment and decision-making errors. The frequency of reported incidents for some causal unsafe acts also differed in different flight phases. Confidential reporting had a positive effect on LFI, as these reports contained more information about latent failures than nonconfidential reports; this also applied to reports about incidents caused by violations. Furthermore, confidential reports were more often used to report incidents caused by widespread or routine violations. By looking at hindering factors in the Investigation Incidents phase, we identified a total of 17 person-related latent failures, which can be considered risk factors for various unsafe acts and the incidents they cause.

The identified unequal distribution of incident reports depending on type of active failure causing the incident (pilots' unsafe acts) suggests that a not insignificant proportion of errors remain undetected – today just as 30 years ago (cf. Reason, 1992). A partially comparable distribution of causal errors and violations to our results could, for example, be demonstrated by Munene (2016) when analyzing accident reports. In our study, reports of incidents caused by judgment and decision-making errors account for less than 10% of all reports evaluated. This relatively small number of reports is already an indication of limited learning opportunities in this area. In addition, the number of reports does not necessarily correspond to the number of actual incidents: In a study by Haslbeck et al. (2015), the highest number of unreported incidents was calculated for incidents caused by poor decision-making (e.g., a landing with less residual fuel than legally required). These results were also confirmed by the aforementioned survey by Sieberichs and Kluge (2017): Here, pilots stated that they are rather unlikely to report incidents caused by operational decision errors. Considering this limited willingness to file a report on this type of unsafe act suggests that learning opportunities are lacking although more learning opportunities could be provided by the pilots.

Furthermore, there is the risk that the few reports on judgment and decision-making errors are overshadowed by the disproportionately high number of reports on performance-based errors by distorting the safety imagination of the operating staff (cf. Hayes & Maslen, 2015). This effect is reinforced by the fact that more than half of all incidents recorded in the database are caused by incorrectly applied procedures. However, this overweight is not surprising given the high number of procedures (normal, supplementary, abnormal, etc.) and has already been highlighted in other research studies (cf., e.g., Shappell et al., 2017). Even though incidents are suitable for providing an overview of the site's risk (Stemn et al., 2018), the unequal distribution identified suggests that it is likely that an airline will not be aware of all incidents and, in addition, different frequencies in different flight phases must be taken into account.

In the mid-1970s, NASA emphasized the importance of a confidential reporting option as the first feature of a reporting system (NASA, 1976). The results of our study underline that this feature is still important today, as confidential reports contain more information about latent failures. Our results also support Langer's (2016) findings.

In identifying risk factors for various errors and violations, we were able to confirm certain results of Miranda (2017) at a higher level of abstraction: A misinterpreted instrument is the strongest risk factor for a performance-based error. When evaluating the related reports, it became clear that misread instruments may also occur in connection with so-called mixed-fleet-flying (pilots have the license for aircraft types that differ only slightly from each other). Our results thus underline the analysis of Soo et al. (2016) that, "even the smallest shift in instrument location can cause errors in performances" (p. 454). Lack of attention is the second strongest risk factor for performance-based errors and also the most frequently mentioned risk factor in all evaluated reports. Other studies, such as an investigation of accidents caused by loss of control inflight, also identified issues with flight crew attention as a significant contributing factor (Stephens et al., 2017). To address these "attention-related human performance limitations," special training courses for attention management are being developed (Stephens et al., 2017, p. 1). Furthermore, results of attention studies with eye-trackers are being used for the human-centered design of flight decks (cf. Li et al., 2016).

We could expand the state of research, indicating that lack of assertiveness is not only a risk factor for judgment and decision-making errors, but also for violations (cf. Miranda, 2017). This result is not surprising, considering that decisions in commercial cockpits are mainly made jointly by the pilots. However, a lack of assertiveness is the strongest risk factor for judgment and decision-making errors and the second strongest for violations. This would not be expected, given the results of a NASA (2004) survey, in which about three quarters of the commercial pilots interviewed stated that they had a high degree of assertiveness. In addition, we were able to identify complacency as a risk factor for incidents caused by violations – a factor that is usually discussed in conjunction with *automation surprise* (AS; de Boer & Hurts, 2017). One explanation suggested might be that complacency limits the ability of pilots to actively assess risk, thereby increasing their propensity to commit deliberate deviations (Rascher & Schröder, 2016).

Limitations

We have placed particular emphasis on a transparent design and explication of the research process, but due to the large number of reports we could only partially document an empirical anchoring through textual evidence. The criterion of basing the data on evidence is therefore only fulfilled to a limited extent (APA, 2020). In addition, in terms of methodological rigor, intersubjective plausibility is thus partially limited (Renner & Jacob, 2020). When classifying the reports with preconditions it became clear that the categories of DoD-HFACS 7.0 were not exhaustive, as we identified latent failures in some reports that were not covered by a coding rule in the coding manual. The distinction we have chosen between active and latent failures is based on the Swiss cheese model, which has been increasingly criticized in recent years for being too linear, static, and unspecified to capture the real world (Drupsteen & Hasle, 2014; Hollnagel et al., 2006; Larouzee & Le Coze, 2020). Particularly with regard to the formation of a safety imagination through stories, the evaluation carried out here is therefore less suitable (cf. Hayes & Maslen, 2015). The inter-rater reliability (Cohen's $\kappa = .95$) is very satisfactory, but Rädiker and Kuckartz (2019) argue that the use of reliability coefficients is not necessarily appropriate in the context of qualitative content analyses. For the most part, the effect sizes were only in a small to medium section, which limits the validity of most results obtained. When considering the frequencies of the incidents in different flight phases, our evaluation assumed that all four flight phases are of equal length in terms of time – again, limited validity must be expected. Also, the generalizability seems to be limited in this area, because Wheeler et al. (2019) observed a different distribution of incidents in flight phases. A presentation of results in Q3 with conditional effect plots recommended by Best and Wolf (2010) was not realized due to the high number of independent variables. The assessment of the model quality (Q3) was limited with the indices presented, as these indices did not allow for an interpretation of the explained variance. Although we have evaluated the appropriateness of the method selected in advance (cf. Steinke, 2019), we were unable to determine any temporal changes in the aspects investigated, as the date of the incident was not available. Therefore, we were not able to detect any change in the level of information of the reports about latent failure due to cultural aspects, such as *just culture*; moreover, the generalizability of the results is limited. An absolute anonymization (cf. Medjedovic & Witzel, 2010) can be seen as positive from the perspective of research ethics, but prevents further analyses, such as differences in the reports depending on the author's rank.

Despite the limitations mentioned our findings are informative and meaningful in relation to the current literature and the study objectives (cf. APA, 2020). In particular, the high number of evaluated, mostly confidential, reports over a period of almost 20 years is a special feature of our research.

Implications for Research

We have shown that frequency of reported incidents varies depending on the causal unsafe acts, but in the context of this research we are unable to explain the reasons why or to quantify the actual number of unreported cases. Further research should therefore focus on the factors that influence the overall reporting behavior of pilots. To this end, following preparatory expert interviews in early 2020, we conducted a survey with civil pilots, supported by a European pilot association. Future research should also investigate the impact of just culture on the frequency of incident reports and the level of information about latent failures. For a classification of incidents and accidents in civil aviation we propose the following extension of HFACS:

- Time pressure during daily operations – is a factor where time pressure is caused by external, nonorganizational factors such as predetermined takeoff times or irregularities during ground handling.
- Not relying on gut feeling – is a factor in which the crew has an unarticulated gut feeling regarding a potentially dangerous situation but does not take this into account in the decision-making process.

We suggest checking the transferability of the identified risk factors to other HROs and, due to the aforementioned limitations of linear incident analysis, also by using narrative forms such as storytelling (cf. Maslen & Hayes, 2020).

Implications for Airlines

Airlines should pay particular attention to incidents caused by judgment and decision-making errors in the course of learning and safety management. In the formation of a safety imagination and in assessing the site's risk, the risk of bias due to the high frequency of incidents caused by performance-based errors should be considered. Due to the described weaknesses of linear evaluation methods, the value of the stories contained in the reports should also be considered in the context of airlines' organizational learning processes. In addition, sharing stories can be seen as an effective tool against complacency and, according to our findings, reduces the probability of incidents caused by violations (Hayes & Maslen, 2015). Our research shows that a confidential reporting system – despite a just culture that has emerged in many airlines – has a positive effect on the Reporting Incidents phase and therefore we advocate retaining these reporting options. The risk factors we have identified can serve airlines in the Investigation Incidents phase as a basis for identifying latent failures and as focal points for crew training.

The basic idea of a "*zero accident vision*" is that all (serious) accidents are avoidable (Zwetsloot et al., 2017). If airlines bear in mind the learning potential of less frequently reported incidents and latent failures and recognize the value of confidential reporting, bottlenecks in learning from incidents can be widened and the trend of the aforementioned accident statistics may be reversed.

References

Air Force Safety Center. (2016). *Department of Defense Human Factors Analysis and Classification System (DOD-HFACS), Version 7.0*. https://www.safety.af.mil/Divisions/Human-Factors-Division/HFACS/

Altman, D. G. (1990). *Practical statistics for medical research*. CRC Press.

American Psychological Association (APA). (2020). *Mixed methods article reporting standards (MMARS)*. https://apastyle.apa.org/jars/mixed-table-1.pdf

Argote, L. (2012). *Organizational learning: Creating, retaining and transferring knowledge*. Springer.

Argyris, C., & Schön, D. A. (1996). *Organizational learning II. Theory method and practice*. Addison-Wesley.

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2016). *Multivariate analysemethoden*. Springer.

Baur, N., Kelle, U., & Kuckartz, U. (2017). Mixed Methods – Stand der Debatte und aktuelle Problemlagen [Mixed Methods – State of the debate and current problems]. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 69*(S2), 1–37. https://doi.org/10.1007/s11577-017-0450-5

Best, H., & Wolf, C. (2010). Logistische Regression [Logistic regression]. In C. Wolf & H. Best (Eds.), *Handbuch der Sozialwissenschaftlichen Datenanalyse* (pp. 827–854). Springer. https://doi.org/10.1007/978-3-531-92038-2_31

Bo, Win. (2020). Multiple approaches to managing workplace safety and health. *World Safety Journal, 19*(1), 8–10.

Brizon, A., & Wybo, J.-L. (2009). The life cycle of weak signals related to safety. *International Journal of Emergency Management, 6*(2), 117. https://doi.org/10.1504/ijem.2009.029241

Cannon, M. D., & Edmondson, A. C. (2001). Confronting failure: Antecedents and consequences of shared beliefs about failure in organizational work groups. *Journal of Organizational Behavior, 22*(2), 161–177.

Clemson University. (2018a, August 21). *HFACS unsafe acts* [Video]. YouTube. https://www.youtube.com/watch?v=lc5iohwZBPc

Clemson University. (2018b, August 21). *HFACS Preconditions for unsafe acts* [Video]. YouTube. https://www.youtube.com/watch?v=oYD44TVaYV4

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum Associates.

Cohen, T. N., Wiegmann, D. A., & Shappell, S. A. (2015). Evaluating the reliability of the human factors analysis and classification system. *Aerospace Medicine and Human Performance, 86*(8), 728–735. https://doi.org/10.3357/amhp.4218.2015

Conchie, S. M. (2013). Transformational leadership, intrinsic motivation, and trust: A moderated-mediated model of workplace safety. *Journal of Occupational Health Psychology, 18*(2), 198–210. https://doi.org/10.1037/a0031805

Curcuruto, M., & Griffin, M. A. (2018). Prosocial and proactive "safety citizenship behaviour" (SCB): The mediating role of affective commitment and psychological ownership. *Safety Science, 104*, 29–38. https://doi.org/10.1016/j.ssci.2017.12.010

de Boer, R. J., & Hurts, K. (2017). Automation surprise. *Aviation Psychology and Applied Human Factors, 7*(1), 28–41. https://doi.org/0.1027/2192-0923/a000113

Dekker, S. (2006). *The field guide to understanding human error*. Ashgate.

Dekker, S. (2018). *Just culture: Restoring trust and accountability in your organization*. CRC Press.

Drupsteen, L., & Guldenmund, F. W. (2014). What is learning? A review of the safety literature to define learning from incidents, accidents and disasters. *Journal of Contingencies and Crisis Management, 22*(2), 81–96. https://doi.org/10.1111/1468-5973.12039

Drupsteen, L., & Hasle, P. (2014). Why do organizations not learn from incidents? Bottlenecks, causes and conditions for a failure to effectively learn. *Accident Analysis & Prevention, 72*, 351–358. https://doi.org/10.1016/j.aap.2014.07.027

Drupsteen, L., & Wybo, J. L. (2015). Assessing propensity to learn from safety-related events. *Safety Science, 71*, 28–38. https://doi.org/10.1016/j.ssci.2014.02.024

Drupsteen, L., Groeneweg, J., & Zwetsloot, G. I. (2013). Critical steps in learning from incidents: Using learning potential in the process from reporting an incident to accident prevention. *International Journal of Occupational Safety and Ergonomics, 19*(1), 63–77. https://doi.org/10.1080/10803548.2013.11076966

Eid, M., Gollwitzer, M., & Schmitt, M. (2015). *Statistik und Forschungsmethoden*. Beltz.

Ergai, A., Cohen, T., Sharp, J., Wiegmann, D., Gramopadhye, A., & Shappell, S. (2016). Assessment of the Human Factors Analysis and Classification System (HFACS): Intra-rater and inter-rater reliability. *Safety Science, 82*, 393–398. https://doi.org/10.1016/j.ssci.2015.09.028

Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *Academy of Management Review, 10*(4), 803–813.

Gilbey, A., Tani, K., & Tsui, W. H. K. (2016). Outcome knowledge and under-reporting of safety concerns in aviation. *Applied Cognitive Psychology, 30*(2), 141–151. https://doi.org/10.1002/acp.3179

Griffin, M. A., Cordery, J., & Soo, C. (2015). Dynamic safety capability. *Organizational Psychology Review, 6*(3), 248–272. https://doi.org/10.1177/2041386615590679

Griffin, M. A., & Curcuruto, M. (2016). Safety climate in organizations. *Annual Review of Organizational Psychology and Organizational Behavior, 3*, 191–212. https://doi.org/10.1146/annurev-orgpsych-041015-062414

Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology, 5*(3), 347. https://doi.org/10.1037/1076-8998.5.3.347

Haslbeck, A., Schmidt-Moll, C., & Schubert, E. (2015). Pilots' willingness to report aviation incidents. *Proceedings of the 18th International Symposium on Aviation Psychology*, 596–601. https://corescholar.libraries.wright.edu/isap_2015/6

Hayes, J., & Maslen, S. (2015). Knowing stories that matter: Learning for effective safety decision-making. *Journal of Risk Research, 18*(6), 714–726. https://doi.org/10.1080/13669877.2014.910690

Helmreich, R. L., Klinect, J. R., & Wilhelm, J. A. (2017). System safety and threat and error management: The line operational safety audit (LOSA). In E. Salas, K. A. Wilson, & E. Edens (Eds.), *Critical essays on human factors in aviation. Crew resource management: Critical essays* (pp. 1–6). Ashgate.

Hollnagel, E., Woods, D. D., & Leveson, N. (2006). *Resilience engineering. Concepts and precepts*. Ashgate Publishing.

International Civil Aviation Organization (ICAO). (2013). *ECCAIRS aviation data definition standard*. https://www.icao.int/safety/airnavigation/AIG/Documents/ADREP%20Taxonomy/ECCAIRS%20Aviation%201.3.0.12%20(VL%20for%20AttrID%20%20431%20-%20Occurrence%20Classes).pdf

International Civil Aviation Organization (ICAO). (2018a). *Module 3: Safety reporting & investigation*. https://www.icao.int/MID/Documents/2018/Aerodrome%20SMS%20Workshop/M3-1-SMS_Aerodrome_Safety Reporting and Investigation.pdf

International Civil Aviation Organization (ICAO). (2018b). *Safety management system overview*. https://www.icao.int/MID/Documents/2018/Aerodrome%20SMS%20Workshop/M0-2-SMS%20Overview.pdf

International Civil Aviation Organization (ICAO). (2019). *State of global aviation safety*. https://www.icao.int/safety/Documents/ICAO_SR_2019_final_web.pdf

Jacobsson, A., Ek, A., & Akselsson, R. (2011). Method for evaluating learning from incidents using the idea of "level of learning". *Journal of Loss Prevention in the Process Industries, 24*(4), 333–343. https://doi.org/10.1016/j.jlp.2011.01.011

Jacobsson, A., Sales, J., & Mushtaq, F. (2009). A sequential method to identify underlying causes from industrial accidents reported to the MARS database. *Journal of Loss Prevention in the Process Industries, 22*(2), 197–203. https://doi.org/10.1016/j.jlp.2008.12.009

Jausan, M., Silva, J., & Sabatini, R. (2017). A holistic approach to evaluating the effect of safety barriers on the performance of safety reporting systems in aviation organisations. *Journal of Air Transport Management, 63*, 95–107. https://doi.org/10.1016/j.jairtraman.2017.06.004

Kansteiner, K., & König, S. (2020). The role(s) of qualitative content analysis in mixed methods research designs. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 21*(1). https://doi.org/10.17169/fqs-21.1.3412

Kelle, U. (2019). Mixed methods. In N. Baur & J. Blasius (Eds.), *Handbuch Methoden der empirischen Sozialforschung* (pp. 159–172). Springer.

Kuckartz, U. (2019). Qualitative content analysis: From Kracauer's beginnings to today's challenges. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 20*(3). https://doi.org/10.17169/fqs-20.3.3370

Landis, J. R., & Koch, G. G. (1977). The Measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159. https://doi.org/10.2307/2529310

Langer, C. (2016). Improving safety culture with confidential reporting. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*(1), 1612–1616. https://doi.org/10.1177/1541931213601372

Larouzee, J., & Le Coze, J. C. (2020). Good and bad reasons: The Swiss cheese model and its critics. *Safety Science, 126*, 1–11. https://doi.org/10.1016/j.ssci.2020.104660

Lee, S. Y., Bates, P., Murray, P., & Martin, W. (2017). Training flight accidents: An explorative analysis of influencing factors and accident severity. *Aviation Psychology and Applied Human Factors, 7*(2), 107–113. https://doi.org/10.1027/2192-0923/a000121

Leveson, N. (2015). A systems approach to risk management through leading safety indicators. *Reliability Engineering & System Safety, 136*, 17–34. https://doi.org/10.1016/j.ress.2014.10.008

Li, W.-C., Braithwaite, G., Greaves, M., Hsu, C.-K., & Lin, S.-C. (2016). The evaluation of military pilot's attention distributions on the flight deck. *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero '16*, 1–6. https://doi.org/10.1145/2950112.2964588

Li, W.-C., & Harris, D. (2006). Pilot error and its relationship with higher organizational levels: HFACS analysis of 523 accidents. *Aviation, Space, and Environmental Medicine, 77*(10), 1056–1061.

Littlejohn, A., Lukic, D., & Margaryan, A. (2014). Comparing safety culture and learning culture. *Risk Management, 16*(4), 272–293. https://doi.org/10.1057/rm.2015.2

Littlejohn, A., Margaryan, A., Vojt, G., & Lukic, D. (2017). Learning from Incidents Questionnaire (LFIQ): The validation of an instrument designed to measure the quality of learning from incidents in organisations. *Safety Science, 99*, 80–93. https://doi.org/10.1016/j.ssci.2017.02.005

Madigan, R., Golightly, D., & Madders, R. (2016). Application of Human Factors Analysis and Classification System (HFACS) to UK rail safety of the line incidents. *Accident Analysis & Prevention, 97*, 122–131. https://doi.org/10.1016/j.aap.2016.08.023

Margaryan, A., Littlejohn, A., & Stanton, N. A. (2017). Research and development agenda for learning from incidents. *Safety Science, 99*, 5–13. https://doi.org/10.1016/j.ssci.2016.09.004

Maslen, S., & Hayes, J. (2020). "This is how we debate": Engineers' use of stories to reason through disaster causation. *Qualitative Sociology, 43*, 1–22. https://doi.org/10.1007/s11133-020-09452-1

Mayring, P. (2007). Designs in qualitativ orientierter Forschung [Designs in qualitative oriented research]. *Journal für Psychologie, 15*(2), 1–10.

Mayring, P. (2015). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* [Qualitative content analysis. Basics and techniques]. Beltz.

Mayring, P. (2020). Qualitative Forschungsdesigns [Qualitative research designs]. In G. Mey & K. Mruck (Eds.), *Handbuch Qualitative Forschung in der Psychologie* (pp. 1–17). Springer.

Medjedovic, I., & Witzel, A. (2010). *Wiederverwendung qualitativer Daten: Archivierung und Sekundärnutzung qualitativer Interviewtranskripte* [Reuse of qualitative data: Archiving and Secondary Use of Qualitative Interview Transcripts]. Springer.

Merry, A. F., & Henderson, B. (2017). Incident reporting, aviation and anaesthesia. *Anaesthesia and Intensive Care, 45*(3), 291–294. https://doi.org/10.1177/0310057X1704500304

Miranda, A. T. (2017). Using conditional probabilities to understand "human error" in military aviation mishaps. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 1619–1620. https://doi.org/10.1177/1541931213601889

Miranda, A. T. (2018). Understanding human error in naval aviation mishaps. *Human Factors, 60*(6), 763–777. https://doi.org/10.1177/0018720818771904

Munene, I. (2016). An application of the HFACS method to aviation accidents in Africa. *Aviation Psychology and Applied Human Factors, 6*, 33–38. https://doi.org/10.1027/2192-0923/a000093

National Aeronautics and Space Administration (NASA). (1976). *NASA aviation safety reporting system*. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19760026757.pdf

National Aeronautics and Space Administration (NASA). (2004). *Pilot personality profile using the NEO-PI-R*. https://ntrs.nasa.gov/api/citations/20040191539/downloads/20040191539.pdf

Parker, S. K., Bindl, U. K., & Strauss, K. (2010). Making things happen: A model of proactive motivation. *Journal of Management, 36*(4), 827–856. https://doi.org/10.1177/0149206310363732

Putz, D., Schilling, J., Kluge, A., & Stangenberg, C. (2013). Measuring organizational learning from errors: Development and validation of an integrated model and questionnaire. *Management Learning, 44*(5), 511–536. https://doi.org/10.1177/1350507612444391

Rädiker, S., & Kuckartz, U. (2019). *Analyse qualitativer Daten mit MAXQDA* [Analysis of qualitative data with MAXQDA]. Springer.

Rascher, S., & Schröder, R. (2016). Au, C. v. (2016). Die Gestaltung einer konstruktiven Fehlerkultur als Führungsaufgabe in High Reliability Organizations (HRO) am Beispiel der zivilen Luftfahrt [The design of a constructive error culture as a leadership task in High Reliability Organizations (HRO) using the example of civil aviation]. In C. v. Au (Ed.), *Struktur und Kultur einer Leadership-Organisation* (pp 177–200). Springer. https://doi.org/10.1007/978-3-658-12554-7_10

Rasmussen, H. B., Drupsteen, L., & Dyreborg, J. (2013). Can we use near-miss reports for accident prevention? A study in the oil and gas industry in Denmark. *Safety Science Monitor, 17*(2), 1–12.

Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. *Philosophical Transactions of the Royal Society of London. B: Biological Sciences, 327*(1241), 475–484.
Reason, J. (1998). Achieving a safe culture: Theory and practice. *Work & Stress, 12*(3), 293–306. https://doi.org/10.1080/02678379808256868
Reason, J. (2016). *Managing the risks of organizational accidents*. Routledge.
Reason, J. T. (1992). *Human error*. Cambridge University Press.
Renner, K. H., & Jacob, N. C. (2020). *Das Interview* [The Interview]. Springer.
Salmon, P. M., Cornelissen, M., & Trotter, M. J. (2012). Systems-based accident analysis methods: A comparison of Accimap, HFACS, and STAMP. *Safety Science, 50*(4), 1158–1170. https://doi.org/10.1016/j.ssci.2011.11.009
Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 69*(2), 107–131. https://doi.org/10.1007/s11577-017-0454-1
Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2017). Human error and commercial aviation accidents: An analysis using the human factors analysis and classification system. *Human Factors, 49*(2), 73–88. https://doi.org/10.1518/001872007X312469
Sieberichs, S., & Kluge, A. (2017). Influencing factors on error reporting in aviation – a scenario-based approach. *Advances in Intelligent Systems and Computing*, 3–14. https://doi.org/10.1007/978-3-319-60441-1_1
Steinke, I. (2019). Gütekriterien qualitativer Forschung [Quality criteria of qualitative research]. In U. Flick, E. Kardorff, & I. Steinke (Eds.), *Qualitative Forschung – Ein Handbuch* (pp. 319–331). Rowohlt.
Soo, K., Mavin, T. J., & Roth, W.-M. (2016). Mixed-fleet flying in commercial aviation: A joint cognitive systems perspective. *Cognition, Technology & Work, 18*(3), 449–463. https://doi.org/10.1007/s10111-016-0381-3
Stemn, E., Bofinger, C., Cliff, D., & Hassall, M. E. (2018). Failure to learn from safety incidents: Status, challenges and opportunities. *Safety Science, 101*, 313–325. https://doi.org/10.1016/j.ssci.2017.09.018
Stephens, C., Prinzel, L., Harrivel, A., Comstock, R., Abraham, N., Pope, A., Wilkerson, J., & Kiggins, D. (2017). Crew state monitoring and line-oriented flight training for attention management. *19th International Symposium on Aviation Psychology*,196–201.
Stolzer, A. J., Halford, M. C. D., & Goglia, M. J. J. (2015). *Safety management systems in aviation*. Ashgate Publishing.
Sunaryo, I., Sundara, G., & Widyanti, A. (2019). Adoption of DoD HFACS for medication error analysis. In *Exploring resources, process and design for sustainable urban development: Proceedings of the 5th International Conference on Engineering, Technology, and Industrial Application (ICETIA) 2018*. AIP Publishing. https://doi.org/10.1063/1.5112414
Van der Westhuizen, J., & Stanz, K. (2017). Critical incident reporting systems: A necessary multilevel understanding. *Safety Science, 96*, 198–208. https://doi.org/10.1016/j.ssci.2017.04.004
Weick, K., & Sutcliffe, K. (2007). *Managing the unexpected*. Jossey-Bass.
Wheeler, B., Wertin, R., Ardic, E. G., Fox, G., & Silver, I. (2019). Operational mishap and incident reports by phase of flight. *Journal of Management & Engineering Integration, 12*(2), 105–111.
Wiegmann, D. A., & Shappell, S. A. (2017a). The Human Factors Analysis and Classification System (HFACS). In D. A. Wiegmann & S. A. Shappell (Eds.), *A human error approach to aviation accident analysis* (pp. 45–71). Routledge.
Wiegmann, D. A., & Shappell, S. A. (2017b). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Routledge.
Zabari, M. L., & Southern, N. L. (2018). Effects of shame and guilt on error reporting among obstetric clinicians. *Journal of Obstetric, Gynecologic & Neonatal Nursing, 47*(4), 468–478. https://doi.org/10.1016/j.jogn.2018.03.002
Zohar, D. (2003). Safety climate: Conceptual and measurement issues. *Handbook of Occupational Health Psychology*, 123–142. https://doi.org/10.1037/10474-006
Zwetsloot, G. I., Kines, P., Wybo, J. L., Ruotsala, R., Drupsteen, L., & Bezemer, R. A. (2017). Zero Accident Vision based strategies in organisations: Innovative perspectives. *Safety Science, 91*, 260–268. https://doi.org/10.1016/j.ssci.2016.08.016

History
Received July 16, 2020
Revision received November 13, 2020
Accepted February 9, 2021
Published online May 5, 2021

Publication Ethics
We based the presentation of this study on the Mixed Methods Article Reporting Standards (MMARS) (APA, 2020).

ORCID
Sebastian Sieberichs
https://orcid.org/0000-0002-9338-7325

Sebastian Sieberichs
Department of Work, Organizational, and Business Psychology
Ruhr University Bochum
Universitätsstraße 150
44801 Bochum
Germany
sebastian.sieberichs@rub.de

Sebastian Sieberichs is doctoral candidate at the Department of Industrial and Organizational Psychology at Ruhr University Bochum. In addition to his work as a commercial pilot for a major European airline, he is employed by the airline's flight safety department. His expertise is in human factors and safety and fatigue management.

Annette Kluge is Full Professor for Industrial and Organizational Psychology at Ruhr University Bochum. She obtained her Diploma in I/O Psychology at RWTH Aachen and her doctorate in ergonomics and vocational training at University of Kassel in 1994. Her expertise is in human factors, training science, skill retention, and safety management.

Flight Operation Officers

From Job Analysis to Selection Procedures

Merle V. Herzog and Dirk Stelling

German Aerospace Center, Hamburg, Germany

Abstract: Flight operation officers (FOOs) can be considered the wallflowers in the aviation business. No results from analyses of job requirements have been published for this profession, even though this is recommended as a helpful basis for personnel selection. In this study, 32 active and retired FOOs acted as experts assessing the job requirements for the function of FOO by scoring the Fleishman Job Analysis Survey. The results showed competencies in the cognitive, interactive-social, and sensory domains as the most important ones. The results of the requirements analysis can be translated into selection procedures, in which multiple task tests, attention tests, personality tests, and assessment centers with group tasks are particularly relevant.

Keywords: job requirements analysis, abilities, flight operation officer, dispatch, personnel selection

Role of Flight Operation Officers in an Airline

The flight operation officer (FOO) can be considered the wallflower of the aviation industry. Even in the airline business, many people do not know what kind of tasks are performed by FOOs working in operation control centers (OCC; also called "flight operation center" [FOC] or "airline operation center" [AOC]). Job designations such as *dispatcher*, *operations controller*, and *mission supporter* or *navigation specialist* have been combined under the function of FOO. Dispatch includes the preparation of the flight plan, cooperation with air traffic control, as well as the support of cockpit crews in case of in-flight problems. When preparing the necessary flight execution documents for cockpit crews, the meteorological conditions, the current air traffic situations, fuel calculations, as well as the condition of the aircraft and flight operation facilities at the relevant airports and any restrictions are taken into account (*Deutsche Flugdienstberater Vereinigung* [DFV], n.d.). Operation control offers solutions for flight program disruptions, taking into account safety, economy, and customer satisfaction. All components that contribute to an optimal result in the economic interest of the airline are monitored and supported if necessary. In the event of major disruptions, cooperation with the higher-level authorities of the regulatory authorities and international and national air traffic control is essential (DFV, n.d.). The navigation specialists or mission supporters in the back office do not have direct contact with the cockpit but work in a supportive and preparatory capacity. These task distinctions are not made in every OCC; however, in the present study all of them were considered as part of the target sample and brought under the umbrella term "FOO."

Licensing and Training of FOOs

An international comparison shows differences in the licensing and training of FOOs. For example, in contrast to the United States, Europe does not have a robust legal system regarding the activities of FOOs and has no control of the activities and processes of individual operators (Pazourek, 2013). For instance, there is no common European flight dispatcher license. Germany, like other national authorities, issues a national license, which is based on International Civil Aviation Organization recommendations (Cordes, 2007). The studies by Pazourek (2013) and Cordes (2007) show that there is potential for improvement in Europe regarding licensing and training regulations. According to Pazourek (2013), it is not uncommon for European business aviation operators to operate OCCs with inexperienced personnel.

Another difference in international comparison relates to responsibility. There is a joint-responsibility dispatch system in commercial aviation in the United States and Canada. In this model, the pilot in command and the OCC share the responsibility for the flight (Pazourek, 2013). The non-shared responsibility of the German model means that the aircraft commander alone is responsible, including responsibility for the verification of the flight planning.

https://doi.org/10.1027/2192-0923/a000202

According to the professional association of German FOOs (DFV, n.d.) there are five approved training organizations (ATO) for FOOs in Germany. These must ensure that the aspirant has sufficient competence in mathematics, physics, and English. However, the level and scope of any aptitude test are up to the ATO. Theoretical training takes place at the ATO, that is, the flight school. This theoretical training is strongly oriented toward the training provided for commercial pilots (Cordes, 2007). Practical training takes place at a partner organization, for example, an airline. A potential employer, in turn, will define the criteria that an applicant must meet, such as school-leaving certificate or foreign language skills (DFV, n.d.).

Problems of Standardization

In the United States, personnel selection procedures only have legal validity if the connection between the selection procedures and professional activity can be demonstrated on the basis of a work analysis (Brannick & Levine, 2002). In Germany, the *Deutsches Institut für Normung* (DIN) 33430 (2016) provides recommendations that also emphasize the necessity of a job requirements analyses. A job analysis is a descriptive process in which information about work is collected, organized, analyzed, and documented (Wilson, 2014). Requirements in the airline aviation industry are a well-studied topic. Several job analyses concerning the job profiles of pilots and air traffic controllers have been reported (e.g., Goeters et al., 2004). Results of job analysis surveys have been used to create a very well thought-out pilot selection process at the German Aerospace Center (DLR; Goeters & Maschke, 1998; Oubaid, 2013). However, to date, such an analysis has not taken place for German FOOs. A hierarchical task analysis based on an interview with a former flight dispatcher was previously undertaken but this was limited to describing the dispatch of an airplane in detail (Boo, 2016). It also refers to the American market and thus assumes a shared responsibility. Moreover, the other areas of the FOO's responsibilities are not covered and the study cannot be used as a basis for deriving personnel selection procedures.

An OCC simulation study has provided the first indications that memory functions and teamwork including problem solving, coordination, and information utilization are important factors for successful professional practice (Littlepage & Wertheimer, 2017). Thus, observer-rated teamwork was meaningfully related to two effectiveness measures, namely, trigger effectiveness and delay loss. Observer-rated teamwork also mediated the relationships between transactive memory and both effectiveness measures. It can be assumed that other factors are also relevant for successful professional practice. These could be identified with a job analysis.

Since FOOs are not immune to the upheavals in aviation and in the general economy, it is even more important to define the current requirements for the activity. Computers are already taking over the preparation of flight routes, which changes the task spectrum for dispatchers. As so-called third pilots, flight dispatchers can intervene in the case of in-flight problems and thus further contribute to safe flight operations (Boo, 2016). The FOO profession could change if an OCC is used to remotely control aircraft, such as intra-city air taxis (Nneji et al., 2018) or drones (Prats et al., 2008). This could change the requirements in the FOO profession, for example, by placing greater importance on psychomotor abilities. However, changing the requirements in a task area can only be addressed with specific training if the original requirements are known.

In summary, the current research situation is limited. The FOO profession has so far received too little attention. Currently, neither personnel selection nor training is based on an analysis of the required abilities and skills. For this reason, the purpose of the present paper is to define the job requirements for an FOO and give recommendations for selection. This could provide the basis for adapting the requirements for licensing and the specifications for other selection and training standards.

Method

Participants and Sampling Procedure

Participation was voluntary and independent of any airline. Multiple German airlines were contacted and the professional association of German FOOs (DFV) provided additional support and helped to disseminate the survey. A total of 58 questionnaires were distributed to FOOs as subject matter experts, of whom 32 responded, resulting in a return rate of 55%.

Eight female and 24 male FOOs with an average of 12.64 years (SD = 11.36) of work experience and a mean age of 43.94 years (SD = 15.54 years) participated in the study. Half of the participants were working as trainers and one-quarter had a management function. Ten participants worked in small-sized airlines (up to 25 aircraft), 13 worked in medium-sized airlines (25–200 aircraft), while seven were employed by large airlines (over 200 aircraft; two were missing). In total, the 32 respondents were employed by nine airlines. Most of the participants reported additional experience in the aviation business, less than a quarter reported none.

The majority of participants were licensed by the German federal aviation authority (*Luftfahrtbundesamt*, LBA) and additionally two had an American FAA (Federal Aviation Administration) license. Six participants indicated that they had no licenses at all. Most of the participants worked as FOO at the time of the survey, but three of them were retired and made their ratings looking back at the time of their work. A supplementary analysis showed no differences in the rating of requirements between licensed versus not licensed and retired versus active FOOs; thus, the different groups were combined.

Questionnaire

The job holders acted as experts, and they assessed their own job demands. For this purpose, the participants answered a German adaptation of the Fleishman Job Analysis Survey (F-JAS; Kleinmann et al., 2010), an established job analysis instrument (Wilson, 2014). This trait-oriented and standardized tool is sufficient for assessing (required) human abilities. The revised version shows satisfactory psychometric properties (Kleinmann et al., 2010). The F-JAS consists of 73 rating scales covering five domains: cognitive, psychomotor, physical, sensory, and interactive-social skills. In the present study, two additional competency scales were added to the cognitive domain: operational monitoring and vigilance. Both scales were developed to anticipate future requirements in aviation jobs (Eißfeldt et al., 2009; Eißfeldt, 2016). The rating scales of the F-JAS are 7-point behaviorally anchored scales. Beside every behavioral indicator, a detailed definition and conceptual delimitations are provided to assist the respondent. The scores are indications of the level of ability or skill that is required to perform the job properly. Higher values indicate greater importance of this ability for doing the job. The recommended required number of at least 15 respondents (Kleinmann et al., 2010) was exceeded in this study.

Only abilities with an average rating greater or equal to 4 are considered to be of general significance for a particular activity (Kleinmann et al., 2010). Since stricter criteria are useful for more demanding professions, competency scales with means above 5.50 are considered relevant abilities in the present study.

Results

The overall mean ratings exhibited a descending rank order of job requirements from cognitive abilities to physical abilities. This ranking was verified by a nonparametric Friedman analysis of variance (χ = 106.45, *df* = 4, $p < .001$). The most relevant competency scales are shown in Table 1. They are from the cognitive (*M* = 5.15, *SD* = 0.55, range = 5.91–4.09), interactive-social (*M* = 5.21, *SD* = 0.84, range = 6.38–4.06), and sensory domain (*M* = 4.04, *SD* = 0.85, range = 5.81–2.59). None of the physical (*M* = 1.87, *SD* = 0.91, range = 2.03–1.43) or psychomotor (*M* = 2.47, *SD* = 1.04, range = 3.30–2.00) domain scales were found to be relevant.

Table 1. Means (*M*) and standard deviations (*SD*) for job requirements of flight operation officers

Competencies	*M*	*SD*
Reliability	6.38	0.75
Mental stamina	5.94	1.01
Simultaneous information processing	5.91	0.78
Operational monitoring	5.91	1.17
Problem sensitivity	5.84	0.88
Emotional control	5.84	0.85
Speech recognition	5.81	0.69
Selective attention	5.69	1.00
Speed of closure	5.66	1.00
Oral comprehension	5.66	0.97
Vigilance	5.59	1.10
Written comprehension	5.56	1.05
Independence	5.56	1.05
Coordination	5.53	1.02
Auditory attention	5.53	1.02

Note. The 15 most important competencies (*M* > 5.50) across all domains are presented (*N* = 32).

Further analysis of airline size did not result in any significant difference in the assessed requirements. Also, there were no differences between trainers or managers in the evaluation of the scales compared with other FOOs. A distinction between dispatch, operation control, and back office also did not lead to significant differences in the evaluation of the requirements.

Discussion

Top Job Requirements for FOOs

The results of the survey suggest that cognitive and interactive-social requirements are the most important for professionally performing the job of FOO. With regard to the cognitive domain, operational monitoring, simultaneous information processing, problem sensitivity, selective attention, speed of closure, oral and written comprehension, and vigilance are all important. Whereas oral and written comprehension is quite self-explanatory, the other scales need further explanation (Kleinmann et al., 2010). Problem sensitivity is the ability to determine when something is going wrong and includes understanding the problem and its components. Speed of closure is the ability to quickly recognize the meaning of visual or auditory information that initially appears to be without meaning or order. Selective attention is the ability to concentrate on a (possibly boring) task without being distracted. Simultaneous information processing is the ability to switch back and forth between multiple information sources.

The two scales presented in addition to the F-JAS were also among those relevant for the FOO role. Vigilance is the ability to track information thoroughly and continuously over a long period of time, when active action is only rarely required (Eißfeldt et al., 2009). Operational monitoring is the ability to track meaningful information coming from different sources (e.g., automation) without having to act directly (Eißfeldt & Gayraud, 2015).

The most important social-interactive skills for a FOO are reliability, mental stamina, emotional control, independence, and coordination, and according to Kleinmann et al. (2010) these are defined as follows. Reliability is the ability to act reliably and responsibly toward others, for example, to be disciplined, conscientious, and trustworthy when fulfilling tasks. Mental stamina is the ability to maintain an optimal level of effort until work tasks are successfully completed, even if fatigue, distraction, or boredom make this difficult. Emotional control is the ability to stay calm and collected in stressful or unexpected situations. Independence is the ability to work in an unstructured environment with few instructions and control, for example, to make decisions without consulting others. Coordination is the ability to structure work plans and activities taking into account the pace, style, and timing of work.

The necessary sensory skills are mainly limited to basic communication skills and computer literacy. Compared with other operators in aviation (Goeters et al., 2004), the low level of psychomotor skills required is striking. No psychomotor requirement was found to be important for an FOO. The physical requirements seem to be negligible. The people should be healthy, but do not have to be athletic. The requirements do not differ with regard to airline size or position.

Practical Recommendations for Selection

As there are no differences in the assessment of the requirements between dispatch versus operation control versus back office, a common selection of personnel seems to be reasonable. This is also supported by the fact that this distinction is not made in all OCCs and that many FOOs change the focus of their activities during the course of their careers.

To use the results from the job analysis in a selection process, various selection process recommendations are suggested. Ability tests, multiple-task tests, and monitoring tasks should be performed to cover the cognitive domains. With a multiple-task test, simultaneous information processing and problem sensitivity can be assessed. Both vigilance and operational monitoring can be tested using a monitoring task. Selective attention and speed of closure can be covered by visual and auditory attention tests. An auditory task should be used to cover the sensory requirement for auditory attention. Oral and written comprehension should be tested by means of abilities tests. The English language should be used, since in the FOO profession a great deal of communication takes place in English. In addition, computer ability should be assessed by the computer presentation of tasks.

In the FOO profession there are always instances of working alone but also times when teamwork is required. To cover the interactive-social requirement, personality questionnaires and AC group tasks should be carried out. Reliability, including conscientiousness and emotional control, should be assessed using personality questionnaires. During the AC group tasks, behavioral observations should be carried out by trained personnel. Mental stamina can be evaluated during the AC tasks by assessing whether the optimal level of work is invested in a task until it is successfully completed. Independence can be judged in the AC task according to how the participants behave when there are few guidelines and decisions have to be made during the group task. To prove their coordination skills, participants should be able to demonstrate that they can manage time and materials to synchronize their tasks with the other participants. Since an AC group task requires the participants' ability to communicate, the sensory domain can also be addressed.

The AC group task could be implemented using the computer-based Group Assessment of Performance and Behavior (GAP) tool (Zinn et al., 2020), which is currently used in the pilot selection at DLR. GAP is a turnkey solution for conducting fully digital group tasks comprising software and complementary hardware. Touch screens are used as input devices by both candidates and observers. On the candidates' screens, instructions and an operating area for the ongoing task are displayed. Each candidate is supposed to adapt to a specific profile that includes individual goals and weaknesses. The candidates have to address all important information in the face-to-face group discussion. The candidates permanently have to enter the (intermediate) results achieved in the discussion on their individual touch screens. The candidates' screens and entries are displayed on the observers' screens in real time.

GAP usually comprises four sequences, addressing both planning and conflict-oriented aspects. The applicants should find cooperative group solutions and at the same time achieve work goals (Oubaid, 2013). In a conflict task, the applicants could be asked to decide which member of the group should be promoted. In a planning task, the applicants could be requested to relocate passengers in order to meet their individual wishes regarding their seats within a given timeline. During the sequences, the candidates must continuously solve low-level mathematical tasks displayed on their screen. This additional task increases candidates' mental load with an apparent impact on their interactive performance (Zinn et al., 2020).

During the entire session, the observers press digital anchor buttons on their task screens reflecting behavioral indicators of the following competencies: leadership, teamwork, communication, adherence to procedures, and workload management (Zinn et al., 2020). After the session a summary of the candidates' performance is displayed on the observers' screens. Each observer individually carries out their evaluation on the candidates' performance in each of the competencies on a 6-point rating scale. GAP also features candidates' self-assessment regarding the given competencies. All results are forwarded automatically in a database.

In the selection process, the results of the AC group tasks should be considered together with the results of the described performance and personality tests. In an interview, the motivation for the job and the biographical background should be evaluated. If the personnel selection for FOO is implemented in the manner described here, the important requirements of the profession will be met and by taking these requirements into account, DIN 33430 (2016) will be complied with.

References

Boo, J. (2016). Enhancing a flight dispatcher display for safer flight operations. *Open Access Thesis, 751*. https://docs.lib.purdue.edu/open_access_theses/751

Brannick, M. T., & Levine, E. L. (2002). *Job analysis: Methods, research, and applications for human resource management in the new millennium*. Sage Publications. https://doi.org/10.5860/choice.39-5899

Cordes, A. (2007). *Job profile and training requirements for European Flight Dispatchers*. City University London.

Deutsches Institut für Normung e.V. (2016). *Anforderungen an berufsbezogene Eignungsdiagnostik: DIN 33430* [Requirements for job-related aptitude diagnostics: DIN 33430]. https://doi.org/10.31030/2514220

Deutsche Flugdienstberater Vereinigung e.V. (n.d.). https://flugdienstberater.org

Eißfeldt, H. (2016). Operational monitoring in aviation – a scale and some findings. In EAAP (Ed.), *Proceedings of the 32th Conference of European Association of Aviation Psychology, Cascais, Portugal, September 26–30* (pp. 386–396). EAAP.

Eißfeldt, H., & Gayraud, K. (2015). Shifts in ability requirements as perceived among aviation operators. *Procedia Manufacturing, 3*, 2543–2549. https://doi.org/10.1016/j.promfg.2015.07.549

Eißfeldt, H., Grasshoff, D., Hasse, C., Hörmann, H.-J., Schulze Kissing, D., Stern, C., Wenzel, J., & Zierke, O. (2009). *Aviator 2030 – Ability requirements in future ATM systems II: Simulations and experiments*. DLR Forschungsbericht 2009-28. DLR.

Goeters, K.-M., & Maschke, P. (1998). Anforderungsprofil bei Linienflugzeugführern 1997 [Requirement profile for airline pilots 1997]. *FlightcrewInfo, 3*(98), 10–17.

Goeters, K.-M., Maschke, P., & Eißfeldt, H. (2004). Ability requirements in core aviation professions: Job analysis of airline pilots and air traffic controllers. In K.-M. Goeters (Ed.), *Aviation psychology: Practice and Research* (pp. 99–119). https://doi.org/10.4324/9781315261843

Kleinmann, M., Manzey, D., Schumacher, S., & Fleishman, E. A. (2010). *Fleishman Job Analyse System für eigenschaftsbezogene Anforderungsanalysen: F-JAS; deutschsprachige Bearbeitung des Fleishman Job Analysis Survey by Edwin A. Fleishman* [Fleishman Job Analysis System for trait-based job analyses: F-JAS; German adaptation of the Fleishman Job Analysis Survey by Edwin A. Fleishman]. Hogrefe.

Littlepage, G. E., & Wertheimer, M. (2017). Teamwork and emergent cognitive states as predictors of routine and adaptive performance in flight dispatch centers. *19th International Symposium on Aviation Psychology*, 125–130. https://corescholar.libraries.wright.edu/isap_2017/47

Nneji, V. C., Cummings, M. L., Stimpson, A. J., & Goodrich, K. H. (2018). Functional Requirements for Remotely Managing Fleets of On-Demand Passenger Aircraft. AIAA Aerospace Sciences Meeting, January 8–12, 2018, AIAA.

Oubaid, V. (2013). Maßgeschneiderte Verfahren psychologischer Eignungsdiagnostik am Beispiel der Pilotenauswahl [Tailor-made procedures in psychological aptitude diagnostics using the example of pilot selection]. In M. Hinsch & J. Olthoff (Eds.), *Impulsgeber Luftfahrt* (pp. 249–270). Springer. https://doi.org/10.1007/978-3-642-32669-1_11

Pazourek, M. (2013). The challenges of operations control centres (OCC) in European business aviation. *European Business Aviation, 8*, 60–66.

Prats, X., Pastor, E., Royo, P., & López, J. (2008). *Flight dispatching for unmanned aerial vehicles*. Paper presented at the AIAA Guidance, Navigation and Control Conference and Exhibit. Honolulu, Hawaii. https://doi.org/10.2514/6.2008-6636

Wilson, M. A. (2014). A history of job analysis. In L. L. Koppes (Ed.), *Historical perspectives in industrial and organizational psychology* (pp. 219–241). Psychology Press. https://doi.org/10.4324/9781315820972

Zinn, F., Goerke, P., & Marggraf-Micheel, C. (2020). Selecting for cockpit crew. In R. Bor, C. Eriksen, T. L. Hubbard, & R. King (Eds.), *Pilot selection – psychological principles and practice* (pp. 21–34). CRC Press.

History
Received May 26, 2020
Revision received September 15, 2020
Accepted January 18, 2021
Published online May 5, 2021

Merle V. Herzog
German Aerospace Center
Sportallee 54a
22335 Hamburg
Germany
merle.herzog@dlr.de

Merle V. Herzog (MSc, Psychology) works as an aviation psychologist at the DLR Department of Aviation and Space Psychology in Hamburg, Germany. She has been involved in the selection of pilots since 2018. Her field of expertise in empirical research encompasses job requirements, personality assessment, and new cockpit automation concepts.

Dirk Stelling (PhD, Psychology) has worked at the German Aerospace Center (DLR), Department of Aviation and Space Psychology since 1990. He is head of the Crew Performance & Transport division and works as Professor for Differential Psychology and Diagnostics at Fresenius University Hamburg.

Research Note

Display on Demand Method Increases Time Spent Looking Outside the Cockpit

Testing a Training Method During Visual Flight Rules Flights

Colin Blättler[1], Ludovic Fabre[1], and Gregory Froger[2]

[1]Centre de Recherche de l'École de l'Air, Salon Air, France
[2]University of Aix Marseille, France

Abstract: Modern on-board instrumentation can lead to distraction, particularly by absorbing attention inside the cockpit, which reduces air safety. The display on demand (DoD) method tested here aims to impede that problem for glider and visual flight rules (VFR) pilots. In total, 21 students were assigned to either an experimental or a control group in a pre- and post-test design. In the experimental group, the cockpit instruments were displayed on the participants' demand to allow for the orientation of visual attention out of the cockpit. Three types of basic exercises were tested. Skills acquisition was measured while evaluating the ability to follow flight indications given by the instructor such as airspeed, and visual attention was measured by an eye tracker. All participants improved their performance after training. Compared with classic training, the DoD method allows participants to spend more time looking outside the cockpit without any impact on the subjective workload. This is a promising method for ab initio flight training.

Keywords: HMI, aerospace, training, eye tracking, cognitive process

The objective of the present study was to test a training method to increase the time spent looking outside the cockpit and suggest its integration into future modern cockpit training programs. Specifically, this aimed at determining an application framework dedicated to basic visual flight skills (VFR) for gliders and light aircraft. *The Aeronautical Information Manual* (Federal Aviation Administration [FAA] AIM 8.1.6.C.3, 2020) recommends that "the time a pilot spends on visual tasks inside the cabin should represent no more than 1/4 to 1/3 of the scan time outside." However, cockpits are becoming more and more digital, either through cockpit-integrated instrumentation or by carrying portable electronic devices. These devices increase distraction and air safety issues (Funk et al., 1999; Kelly & Efthymiou, 2019; National Transportation Safety Board [NTSB], 2020) by increasing the amount of time spent in the cabin. For example, Johnson and coworkers (2006) reported that the time spent in the cabin without such instruments (e.g., GPS) increases from 40% to 80% with such instruments. Casner (2005, 2006) showed that the use of GPS for VFR navigation reduces situational awareness because the pilots in his study no longer took their bearings in the outside world. The NTSB reported that the widespread use of these instruments in light aviation has increased the number of fatal accidents compared with aircraft equipped with analog cockpits (NTSB, 2010).

Studies have shown that, in the laboratory, it is possible to modulate the allocation of visual–spatial attention to artificial tasks that mimic aeronautical tasks (e.g., Froger et al., 2018). Froger et al. (2018) tested a training method to improve visual attention sharing as a means to support the implementation of the FAA recommendations. They used a specific virtual environment developed to expose participants to a dual-task situation that mimics aeronautical activity. This dual-task condition consisted of one task located in the upper part of the screen (i.e., visual search representing the see-and-avoid safety task within the visual range) and one task located in the lower part of the screen representing the system management activity. These two tasks, although presented at the same time, cannot be carried out simultaneously and require switching from one to the other. Eye movements were recorded to measure the duration of eye fixation on each task. In the control condition, it was found that participants spent 60% of their time looking at the bottom task and 40% of their time on the top task. In the experimental condition, the bottom task was masked as soon as the participant spent more than 2 s on this task, forcing the participant to take their eyes off the

https://doi.org/10.1027/2192-0923/a000201

bottom task and make it reappear. Participants were trained in this condition for 12 min. The duration of eye fixation was re-evaluated immediately after the training session and 24 hr later without the bottom task being masked. The results showed that these participants spent 60% of their time looking up and 40% down after training. The authors concluded that it is possible to permanently modify the allocation of visual-attention resources using a method that displays information for a short period.

Although the use of eye monitoring has the potential to improve learning outcomes, its intensive use is limited by several technological barriers (e.g., sensitivity to sunlight). It seems interesting, therefore, to find an alternative to the eye-tracking option to ensure that student pilots acquire basic flying skills, namely, basic flight maneuvers and visual attention to the outside world.

An alternative procedure to eye-tracking is to use a cockpit without information and make information accessible for a short time at the request of the student pilots so as to involve the student more actively in his or her learning. Chen and Singer (1992) demonstrated that strategic input or an adapted method from coaches or instructors was necessary for learning. In this vein, we proposed to test the display on demand (DoD) method. The DoD consists of an explicit act by the pilot to obtain the information he/she needs to perform his or her task. DoD is fully in line with the self-regulated strategies (SRS) theory defined as "actions occurring during the actual performance of a cognitive task that allows an individual to control, or direct his own activity through self-imposed rules or regulations that better adapt his performance to different circumstances or surroundings" (Ferrari et al., 1991, p. 139). SRS contribute to the perception of self-control that has been demonstrated to improve learning through more in-depth information processing (McCombs, 1989). In education and cognitive psychology, studies have indicated that this higher level of information processing is achieved by allowing learners to participate actively and independently in the learning process through the use of SRS. In this context, participants process information associated with their own meaning. Craik and Tulving (1975) argued that meaning is the main factor influencing in-depth information processing, resulting in better memorization

SRS in the form of DoD information allows for the preactivation of mapping rules for the information that will be processed immediately afterward. This pre-activation of the mapping rules is a priming of the subsequent task, which is likely to improve performance in the realization of the task by allowing faster processing (Maquestiaux, 2012). Thus, the DoD method is expected to produce deeper and faster information processing, which should lead to less time spent overall on visual tasks inside the cabin after training. This method would allow for more time to be spent looking outside the cockpit, which would meet the FAA recommendations.

Method

The goal of the present study was to test a glider flight training method in a simulated environment to develop basic flying skills (i.e., basic flight maneuvers and outside world visual attention). The experiment was designed to compare two ab initio training lessons by including a DoD method and using the cockpit display in a glider simulator. The participants were divided into two groups. In the first group, the participants had three on-board instruments at their disposal during all the training sessions. In the second condition, the instruments were hidden. As soon as the participant asked aloud that the instruments be displayed, the experimenter pressed a button that triggered the display of the instruments for 2 s. For each trial, the participant could only request the instruments a maximum of three times. This condition was meant to allow the participants to look at the outside world. Airspeed deviation was used as the dependent variable; the higher the airspeed deviation, the weaker the performance. The two hypotheses were (1) both groups increased their performance, and (2) more time was spent looking outside by the experimental group in the final test.

Participants

In total, 21 participants from the French Air Force Academy with a mean age of 28.87 years (SD = 7.74) and no previous experience of aircraft flight displays took part in the experiment. They were randomly assigned to one of two groups: 13 participants were assigned to the control group (CG) and eight to the experimental group (EG).

Apparatus

The flight simulator was run using the commercially available Xplane 10.42 (32 bits) software. The simulated aircraft was a two-seater modified ASK21 glider (ref Xplane: *ASK21-Metric_V2.1*). The external view was projected onto a white background to create a view angle of 170° (horizontal) by 60° (vertical) located in front of the participants. The cockpit flight instruments presented basic flight instruments such as the airspeed indicator, heading indicators, and the altimeter. Depending on the experimental condition, each instrument was either presented or not (Figure 1). The glider simulator was connected to a flight stick pro (CH products) and rudder pedals (CH products). Gaze

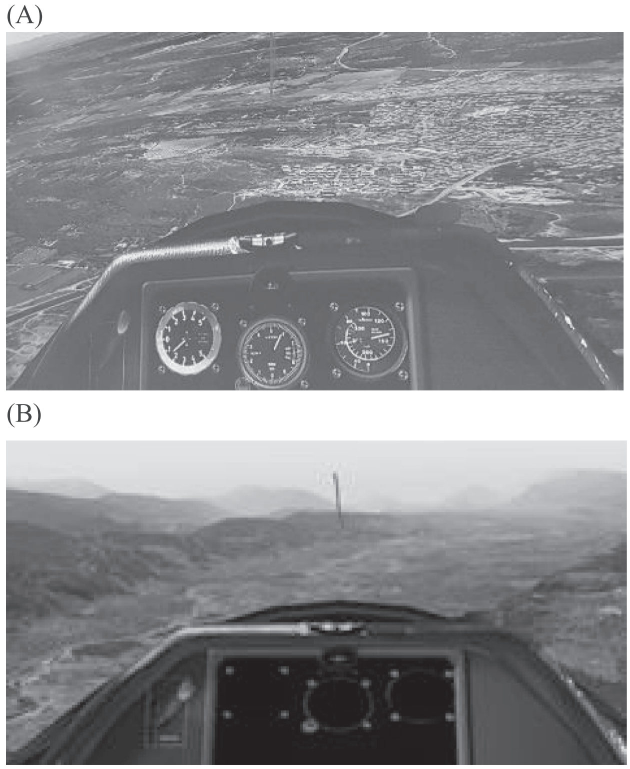


Figure 1. Example of cockpit, with information (A) and without information (B).

positions data was captured by a Tobii Pro Glasses 2 eye-tracker cadenced at 100 Hz.

Procedure

The participants were provided with a basic understanding of how to fly a glider. In a phase of familiarization, they tried out the effects of the glider controls. Immediately after the familiarization phase, the experiment was composed of three learning conditions (straight-line, 360° right and left turns, and final approach) and a final test session. In the straight-line condition, participants had to maintain airspeed (100 km/hr) along six straight lines for 60 s. In the turning condition, participants had to complete four 360-degree right and four 360-degree left turns with a 30° bank angle maintaining constant airspeed. In the final approach condition, participants had to complete six approaches maintaining a minimum airspeed of 90 km/hr to avoid stall.

The participants were given a briefing sheet that included the objectives of the flight. After each trial, the experimenter (with a glider pilot license) gave feedback on whether or not the airspeed was correct. For each condition, the first trial was used as a baseline and compared with the final test session. The participants in both groups were equipped with an eye-tracker for all trials to record the time spent looking inside the cockpit. Subjective measures of workload (NASA-TLX) were collected at the end of the experiment. The experiment lasted approximately 45 min.

Results

An ANOVA was conducted with Group (CG vs. EG) as between subjects factor on airspeed deviation for straight lines, turns, and approach situations in the baseline condition. No group difference was found, $F(1,19) < 1$; $p = .444$, meaning that performances were comparable between the two groups at the beginning of the experiment. To test the impact of training on flight performances, an ANOVA was conducted with Group as between-subjects factor on airspeed deviation per condition and Training (baseline vs. final test) as within-subjects factor. For the straight-line condition, a significant main effect of Training was found, $F(1,19) = 6{,}42$; $\eta^2 = 0.25$; $p = .02$. The participants were better at maintaining airspeed in the post-test ($M = 3.82$; $SD = 2.37$) than in the baseline condition ($M = 7.07$; $SD = 6.18$). The effect of Group condition and the interaction was not significant, $F(1,19) < 1$; $p = .63$ and $F(1,19) < 1$; $p = .64$, respectively. Regarding the turning condition, a significant main effect of Training was also observed, $F(1,19) = 8.09$; $\eta^2 = 0.30$; $p = .0103$). The participants were better at maintaining airspeed in the post-test ($M = 7.60$; $SD = 5.08$) than in the baseline condition ($M = 17.50$; $SD = 13.98$). The effect of Group condition and the interaction was not significant, $F(1,19) = 1.38$; $p = .25$ and $F(1,19) < 1$; $p = .95$, respectively. Concerning the final approach, no significant effect was observed.

Eye-tracker recordings revealed that only EG using the DoD method decreased the time spent looking inside the cockpit after training compared with baseline (Figure 2). The ANOVA showed no main effect or interaction. Fisher's (LSD) post hoc analyses were conducted. The following pairs were significantly different: EG baseline ($M = 38.75$; $SD = 18.75$) and EG final test ($M = 23$ .17; $SD = 13.78$).

Perceptions of workload were measured with the NASA-TLX scale. No significant differences were found: EG: $M = 64.29$, $SD = 10.91$; CG: $M = 62.74$, $SD = 6.28$; $t(19) = .41$; $p = .8$.

Discussion and Conclusion

The current study aimed to propose a flight training method favoring the time spent looking outside the cockpit. As expected, all participants improved their performance after training. Compared with classic training, the DoD method enabled participants to spend more time looking outside

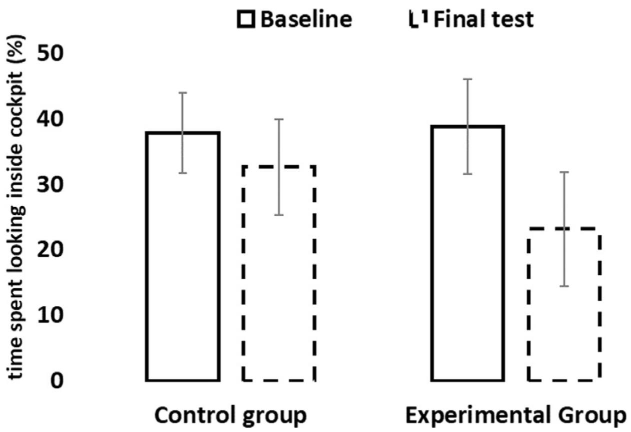


Figure 2. Time spent looking inside the cockpit (%) by group for the baseline and final test.

the cockpit without any impact on the subjective workload and significant influence on performance.

Thus, these results, which are in line with SRS theory (Chen & Singer, 1992), offer a possible new instructor–student interaction in the process of learning basic VFR skills. This is a promising way to impede the absorption of attention caused by the elements of glass cockpits.

The DoD method seems to be an effective way to overcome the technical barriers associated with eye-tracking methodology. However, the present experiment was based on only one training session which lasted approximately 45 min; therefore, future research must evaluate the increasing or decreasing effects of the DoD method in multiple-session training. Moreover, the DoD method could be tested during real flights to possibly design a "training function" in the future cockpit system.

The DoD learning method based on a glider flight simulator and numerical displays during the first stage of pilot training (ab initio) contributes to methodological advances in the training, learning, and cognitive engineering fields (Salas et al., 1998). This DoD learning strategy encourages the acquisition of a visual attention pathway. This is an additional reason for promoting the development of specific courses (part-task training) using flight simulators, in a more comprehensive training program, dedicated to specific skills (e.g., visual attention scan pattern, action schemata). However, the positive transfer of the DoD method and the acquisition of more complex skills, such as the articulation of basic skills in a landing circuit, should be validated in future work.

The DoD learning method is part of an overall approach to improving air safety. This method seems compatible with all generations of instruments (analog and digital). Hence, DoD should contribute to the design of flight instruments dedicated to the acquisition of flight-specific knowledge.

References

Casner, S. M. (2005). The effect of GPS and moving maps displays on navigational awareness while flying under VFR. *International Journal of Applied Aviation Studies, 5*(1), 153–165.

Casner, S. (2006). Mitigating the loss of navigational awareness while flying with GPS and moving map displays under VFR. *International Journal of Applied Aviation Studies, 6*(1), 121.

Chen, D., & Singer, R. N. (1992). Self-regulation and cognitive strategies in sport participation. *International Journal of Sport Psychology, 23*, 277–300.

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268

Federal Aviation Administration. (2013). *Industry training standards (FITS) more*. http://www.faa.gov/training_testing/training/fits/

Federal Aviation Administration. (2020). *Aeronautical information manual*. https://www.faa.gov/air_traffic/publications/atpubs/aim_html/index.html

Ferrari, M., Pinard, A., Reid, L., & Bouffard-Bouchard, T. (1991). The relationship between expertise and self-regulation in movement performance: Some theoretical issues. *Perceptual and Motor Skills, 72*(1), 139–150. https://doi.org/10.2466/pms.1991.72.1.139

Froger, G., Blättler, C., Dubois, E., Camachon, C., & Bonnardel, N. (2018). Time-interval emphasis in an aeronautical dual-task context: A countermeasure to task absorption. *Human Factors, 60*(7), 936–946. https://doi.org/10.1177/0018720818783946

Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteguh, C., & Owen, G. (1999). Flight deck automation issues. *The International Journal of Aviation Psychology, 9*(2), 109–123. https://doi.org/10.1207/s15327108ijap0902_2

Johnson, N., Wiegmann, D., & Wickens, C. (2006). Effects of advanced cockpit displays on general aviation pilots' decisions to continue visual flight rules flight into instrument meteorological conditions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(1), 30–34. Sage Publications. https://doi.org/10.1177/154193120605000107

Kelly, D., & Efthymiou, M. (2019). An analysis of human factors in fifty controlled flight into terrain aviation accidents from 2007 to 2017. *Journal of Safety Research, 69*, 155–165. https://doi.org/10.1016/j.jsr.2019.03.009

Maquestiaux, F. (2012). La simultanéité des actes psychiques: Apports du protocole PRP. *L'Année psychologique, 112*(4), 631–662. https://doi.org/10.4074/S0003503312004058

McCombs, B. L. (1989). Self-regulated learning and academic achievement: A phenomenological view. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement. Springer series in cognitive development* (pp. 51–82). Springer. https://doi.org/10.1007/978-1-4612-3618-4_3

National Transportation Safety Board. (2010). *Safety study: Introduction of glass cockpit avionics into light aircraft.* NTSB. Report: NTSB Publication No. SS-10/01. http://www.ntsb.gov/doclib/safetystudies/SS1001.pdf

National Transportation Safety Board. (2010). *Most wanted list of transportation safety improvements*. https://www.ntsb.gov/safety/mwl/Documents/2019-20/2019-20-MWL-SafetyRecs.pdf

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The International Journal of Aviation Psychology, 8*(3), 197–208. https://doi.org/10.1207/s15327108ijap0803_2

History
Received November 15, 2019
Revision received July 24, 2020
Accepted August 20, 2020
Published online December 4, 2020

ORCID
Colin Blättler
https://orcid.org/0000-0002-7415-0269

Colin Blättler
Centre de Recherche de l'École de l'Air
Ecole d'air
Base aérienne 701
13661 Salon Air
France
colin.blattler@ecole-air.fr

Colin Blättler is an associate professor in human factors. He carries out his activity within the Centre de Recherche de l'École de l'Air (CREA), Salon Air, France. Its work is organized around the use of new technologies for civil and combat flight training.

Ludovic Fabre is an associate professor in human factors. He carries out his activity within the Centre de Recherche de l'École de l'Air (CREA), Salon Air, France. His research focuses on three themes: cognitive fatigue, the influence of emotions on cognition, and the contribution of new technologies in training.

Grégory Froger is a PhD student of psychology at the University of Aix-Marseille, France. He received his master's degree in psychology from Dijon University in 2015.

News and Announcements

EAAP2020 Online Week

Building Bridges: Raising Aviation Psychology to the Next Level!

Michaela Schwarz

European Association for Aviation Psychology

From September 28 to October 2, 2020 the EAAP Board organized its first inaugural online meeting in the history of EAAP with more than 160 participants worldwide and 30 speakers in different time zones.

The EAAP2020 Online Meeting and Virtual Workshop Week 2020 was organized as a replacement for the 34th EAAP Conference, that unfortunately had to be postponed due to COVID-19 restrictions.

The online week was offered free of charge so support EAAP members in difficult economic times during the pandemic. To make this possible EAAP was relying on invited speakers and facilitators to offer their expertise free of charge. At this point the EAAP Board on behalf of all members would like to say a big thank you to all speakers for their willingness to contribute in their spare time with such an enthusiasm and degree of professionalism.

The online week via the platform Zoom could also not have been realized without the University of Applied Sciences in Berlin (HMKW), a longstanding EAAP partner, who kindly agreed to sponsor EAAP by providing four Zoom rooms and the associated licenses free of charge. HMKW offers an international master program in business psychology with a strong focus on Human Factors: https://www.hmkw.de/en/university/departments/psychology/department-of-psychology.

The EAAP Board would like to thank Prof. Dr. Ronald Freytag and Dr. Harald Kolrep and their teams for supporting the EAAP Online Week.

The main objectives of the online week were:

- to stay in touch with EAAP members during difficult times.
- to facilitate networking amongst members.
- to offer continuous education and practical exchange for EAAP members accounted towards their accreditation.
- to report on recent EAAP activities including five dedicated working groups.

The first day of the week was dedicated to bringing together regional aviation psychology associations to hold their annual business meetings online and discuss future issues and cooperation. The EAAP family has been growing over the years with about 15% of members living and working outside the boundaries of Europe. It was important to the EAAP Board to offer multiple opportunities for regional partners to get together and distribute time slots throughout the day for the benefit of members in different time zones.

The most popular session on day 1 was the regional meeting with invited presidents from aviation psychology associations from Germany, Austria, Spain, Italy, Australia, Brazil, Argentina, Uruguay, and Chile. Next to introducing the different associations and teams to each other, popular topics such as the accreditation of aviation psychologists, the new EASA regulations in relation to aviation mental health and psychological (risk) assessments, pilot support initiatives, and future aviation psychology issues were discussed. Thank you to all presidents for their time and effort in preparing the meeting, and ensuring an exceptionally smooth, efficient, and highly productive meeting.

Day 1 concluded with the renowned basic civil aviation knowledge course held by Dr. Rainer Brorsen aimed at students and newcomers to aviation psychology. However, many experienced aviation psychologists were also happy to share their wisdom. Time really passed quickly in the 1,5 hours session – thanks to Rainer making the course very interactive including myth buster on-board, as well as engaging the audience in guessing airport designators based on runway layouts.

Day 2 was dedicated to the fundamentals of setting up and developing peer support programs and aeromedical considerations and opportunities with invited speakers and founders of the European Pilot Peer Support Initiative (EPPSI), including representatives from the European Cockpit Association (ECA), the European Society of

https://doi.org/10.1027/2192-0923/a000206

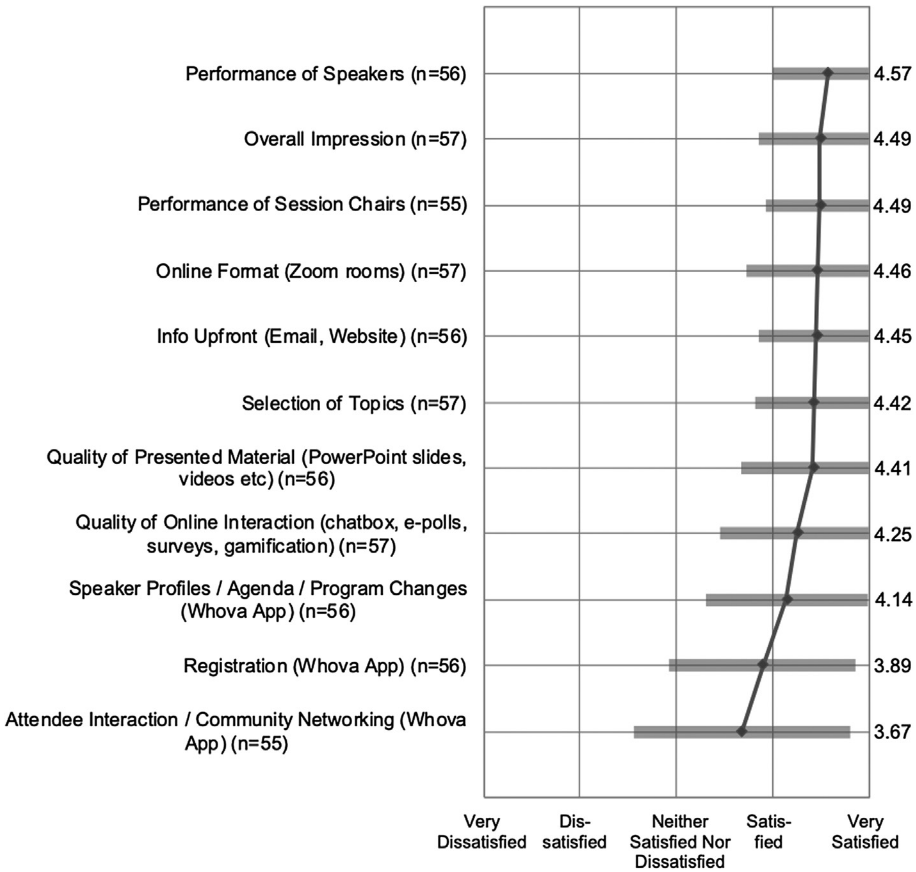


Figure 1. Online week evolution graph.

Aviation Medicine (ESAM), the Mayday Foundation, and the Center for Aviation Psychology (CAP). Day 2 concluded with the introduction of the new Code of Professional Conduct for EAAP members developed by EAAP members in the associated work group led by Board member Dr. Robert Bor. The draft Code of Professional Conduct was reviewed and challenged by participants and the revised version will be put forward to EAAP members for endorsement at the upcoming EAAP Business Meeting in May 2021.

Day 3 started of with discussing the new postgraduate aviation curricula developed by the Austrian and Spanish Aviation Psychology Association (AAPA / AEPA). The session chair led participants through a lively discussion via chat related to topics which should be covered in the curriculum and to what extent. A big challenge considering different levels of psychology education and different academic requirements in Europe and beyond.

The parallel session informed members about the recent lived experience and well-being project conducted by the Centre for Innovative Human Systems (CIHS) School of Psychology, Trinity College Dublin, Ireland. Results were astonishing and most relevant to all participants, raising many questions in the chat box, which were all answered live.

Day 4 was designed to be aimed at members with a specific Human Factors background kicking off with EUROCONTROL and DeepBlue on automation in ATC, followed by training needs for aviation psychologists and human factors specialists. In parallel, CAP was running a session on cognitive assessment of aircrew for aviation psychologists. The end of the day was dedicated to students and professionals engaged in writing scientific papers. Thank you to former board member and former Editor-in-Chief of the EAAP journal *Aviation Psychology and Applied Human Factors,* Prof. Dr. Don Harris, for leading the way. For all who missed this session, please refer to Don’s guidebook on writing human factors research papers (2012).

Day 5 came very quickly, featuring an excerpt of the popular Human Factors in Flight Safety Course held by Brent Hayward and Alan Hobbs in fond memory of Rob Lee. In parallel, Vsevolod Peysakhovich from ISAE-SUPAERO in Toulouse shared his knowledge and experience in cognitive psychology with amazing insights in the possibilities on what modern eye tracking technology can offer to aviation psychology and human factors studies.

Following good standard practice of EAAP, the online week was evaluated using an online evaluation form for quality assurance purposes. Next to the overall online format, the performance of speakers and session chairs, selection of topics, quality of presented material and handouts were rated on a 5-point rating scale. Special thanks at this stage go to former EAAP student and now experienced Human Factors Specialist Ms Zsofi Berkes, who has been managing the conference evaluation for the third time in a row. More than 50 participants took the opportunity to evaluate the online week, delivering evidence that aviation conferences and meetings could be realised online in the future if needed (see Figure 1).

Areas of improvement pointed towards investing in an even more interactive platform and offering more recent research, case studies, and lessons learnt in addition to basic lectures. The EAAP Board will certainly take these on-board for the next conference.

Finally, I would like to express my sincere thanks to all speakers, session chairs, support staff, sponsors, and participants for making this online week a great success. I look forward to meeting all of you hopefully live and in person at the next EAAP Conference now planned for

September 26–30, 2022
at Yacht Hotel Sunborn in Gibraltar, UK.

All details are available on the conference's website https://conference.eaap.net/

My very last words I save for Jenny, Renée, Karina, Gunnar, Mickaël, and Rob, who, since their election in September 2018 grew to be not only an extraordinarily strong team, but an amazing EAAP Board, dedicating an enormous amount of their spare time to EAAP business and leading the way to raising aviation psychology to the next level for the benefit of the EAAP family. Thank you, team, I am very proud to working with you guys!

Published online May 5, 2021

Michaela Schwarz, PhD
European Association for Aviation Psychology
EAAP
president@eaap.net

Michaela Schwarz (PhD, University of Graz, Austria) is an accredited aviation psychologist and human factors expert working toward improving human performance in aviation and rail. Her main research focus and expertise is the assessment and improvement of safety and just culture, the integration of human performance elements in safety management systems, and the development and delivery of human factors training programs. Michaela is President of the European Association for Aviation Psychology (EAAP) and Vice-Chair of the Austrian Aviation Psychology Association.

News and Announcements

Meetings and Congresses

International Symposium on Aviation Psychology (Virtual Meeting)
May 18–21, 2021
Contact: ISAP, https://aviation-psychology.org/
ISAP is a venue for research on human performance within aviation systems, and for design solutions tailored to human aptitudes and constraints. We welcome proposals on any topics within aviation psychology, and on any basic or applied research that speaks to issues of human performance in socio-technical systems. Behavioral, neuroscientific, computational, and other approaches are all appropriate.
Our format will be a blend of live and pre-recorded presentations, synchronous Q&A sessions, and virtual salons. All presenters will be invited to contribute a 5-page paper to conference proceedings.

Course: Initial Human Factors in Flight Safety Barcelona, Spain
October 1, 2021
Contact: EAAP, https://www.eaap.net
This EAAP-recognised Human Factors in Flight Safety training course has been rescheduled for later this year, hopefully late September/early October. The course will be delivered by the experienced team of Brent Hayward and Alan Hobbs, together will special guest presenters.
The course will be planned and conducted in a "COVID-Safe" manner, and will proceed if international travel restrictions and other conditions related to the pandemic allow that. The training team will monitor the situation closely in upcoming months and advise when a final decision on holding the course has been made.
An information and registration brochure is in preparation and will be released once it is confirmed that the course can proceed safely. In the meantime those interested in participating can register their interest directly with Brent Hayward: bhayward@dedale.net

PACDEFF CRM and Aviation Human Factors Conference
Sydney, Australia
October 21–22, 2021
Contact: PACDEFF, E-mail wayne@pacdeff.com, Web https://pacdeff.com/
PACDEFF 2021 will be a combined hybrid conference with the Australian Aviation Psychology Association. PACDEFF is the largest CRM, NTS and Aviation Human Factors Conference of its type in the world, with around 300 attendees generally attending pre-COVID. The forum is intended as a non-profit, non-partisan opportunity for Human Factors practitioners to meet and discuss contemporary issues in the Human Factors field, with an emphasis on airline training. Please note that the conference is planned to be a hybrid event where some international speakers in particular, will appear via Zoom, with a limited physical audience of around 150–200 people, due to social-distancing guidelines. It is also intended that those unable to attend in person will be able to do so via Zoom as well.

4th International Conference on Human Factors Munich, Germany
t.b.a.
Contact: Lufthansa Aviation Training, Web https://www.human-factors-conference.com/conference/
Due to the current COVID-19 situation, the conference had to be rescheduled to a later date. The theme of this 4th conference is "The human factors enigma." This conference should help you to solve parts of the mystery behind the human factor. What are some of the challenges out there and how can and should we deal with them. What leadership skills are needed and why empathy is a skill that is needed more than ever. How to think critically, how to teach and strengthen all sides of our brain, why cooperation is key and what to do when you are in an extreme situations. The conference is an EAAP recognized course.

HFES – Annual Meeting Europe Chapter
Liverpool, UK
April 20–22, 2022
Contact: Secretary Europe Chapter of the HFES, Groningen, The Netherlands, E-mail secretary@hfes-europe.org, Web https://www.hfes-europe.org/annual-meeting/
Under the present circumstances, the HFES has decided to postpone the 2020 conference. The Human Factors and Ergonomics Society, Europe Chapter, is organized to serve the needs of the human factors profession in Europe. Its purpose is to promote and advance through the interchange of knowledge and methodology in the behavioral,

https://doi.org/10.1027/2192-0923/a000203

biological, and physical sciences, the understanding of the human factors involved in, and the application of that understanding to the design, acquisition, and use of hardware, software, and personnel aspects of tools, devices, machines, equipment, computers, vehicles, systems, and artificial environments of all kinds. The Chapter is an affiliate of the Human Factors and Ergonomics Society, Inc., a nonprofit corporation chartered by the State of California, USA, and organizes scientific meetings every year.

1st International Conference in Aerospace Medicine
Paris, France
September 22–24, 2022
Contact: Aerospace Medical Association, Web https://www.asma.org/news-events/events/1st-international-conference-in-aerospace-medicine
ICAM 2022 will be the first international forum to bring together civilian and military specialists involved in Aerospace Medicine, airlines medical operations, repatriation, and occupational health. The Scientific Committee is preparing an exciting program. This meeting was originally postponed to September 2021 due to the COVID-19 pandemic and has now been postponed to September 2022 to ensure the pandemic is under control and restrictions on international travel and large group gatherings have been removed. This conference is under the auspices and joint sponsorship of: The Aerospace Medical Association – AsMA, The International Academy of Aviation and Space Medicine – IAASM, The European Society of Aerospace Medicine – ESAM, and La Société Francophone de Médecine Aérospatiale – SOFRAMAS.

34th EAAP Conference
Gibraltar, UK
September 26–30, 2022
Contact: European Association for Aviation Psychology (EAAP), E-mail secretarygeneral@eaap.net, conferences@eaap.net, Web https://conference.eaap.net
EAAP34 conference is postponed to 2022. The new call for Papers will be published October 2021. The theme of the 34th EAAP conference is “Building bridges: Enabling human performance.” The objective of EAAP is to promote the study of psychology and the scientific pursuit of applied psychology in the field of aviation. The much appreciated and well attended EAAP conferences provide a forum for professionals and students from all domains of aviation psychology and aviation human factors. Professionals and students from related fields, like aerospace medicine, are also very welcome. For details about the program and registration, please visit the above mentioned website.

Instructions to Authors

Aviation Psychology and Applied Human Factors (APAHF) publishes innovative, original, high-quality applied research covering all aspects of the aerospace domain. In order to make the journal accessible to both practitioners and scientific researchers, the contents are broadly divided into original scientific research articles and papers for practitioners.

The fully peer-reviewed articles cover a variety of methodological approaches, ranging from experimental surveys to ethnographic and observational research, from those psychological and human factors disciplines relevant to the field, including social psychology, cognitive psychology, and ergonomics. High-quality critical review articles and meta-analyses cover particular topics of current scientific interest. APAHF in Practice consists of shorter, less technical, but still fully peer-reviewed articles covering a wide range of topics, such as comments on incidents and accidents, innovative applications of aviation psychology, and reviews of best practices in industry.

***Aviation Psychology and Applied Human Factors* publishes the following types of articles:** Original Articles, Research Notes, Review Papers, APAHF in Practice, Book Reviews, News and Announcements.

Manuscript Submission: Manuscripts for any section must be submitted online at http://www.editorialmanager.com/apahf. Should you have any technical queries please contact Juliane Munson, Hogrefe Publishing at production@hogrefe.com. Detailed instructions to authors are provided at www.hgf.io/apf

Copyright Agreement: By submitting an article, the author confirms and guarantees on behalf of him-/herself and any coauthors that the manuscript has not been submitted or published elsewhere, and that he or she holds all copyright in and titles to the submitted contribution, including any figures, photographs, line drawings, plans, maps, sketches, tables, and electronic supplementary material, and that the article and its contents do not infringe in any way on the rights of third parties. ESM will be published online as received from the author(s) without any conversion, testing, or reformatting. They will not be checked for typographical errors or functionality. The author indemnifies and holds harmless the publisher from any third-party claims.

The author agrees, upon acceptance of the article for publication, to transfer to the publisher the exclusive right to reproduce and distribute the article and its contents, both physically and in nonphysical, electronic, or other form, in the journal to which it has been submitted and in other independent publications, with no limitations on the number of copies or on the form or the extent of distribution. These rights are transferred for the duration of copyright as defined by international law. Furthermore, the author transfers to the publisher the following exclusive rights to the article and its contents:

1. The rights to produce advance copies, reprints, or offprints of the article, in full or in part, to undertake or allow translations into other languages, to distribute other forms or modified versions of the article, and to produce and distribute summaries or abstracts.
2. The rights to microfilm and microfiche editions or similar, to the use of the article and its contents in videotext, teletext, and similar systems, to recordings or reproduction using other media, digital or analog, including electronic, magnetic, and optical media, and in multimedia form, as well as for public broadcasting in radio, television, or other forms of broadcast.
3. The rights to store the article and its content in machine-readable or electronic form on all media (such as computer disks, compact disks, magnetic tape), to store the article and its contents in online databases belonging to the publisher or third parties for viewing or downloading by third parties, and to present or reproduce the article or its contents on visual display screens, monitors, and similar devices, either directly or via data transmission.
4. The rights to reproduce and distribute the article and its contents by all other means, including photomechanical and similar processes (such as photocopying or facsimile), and as part of so-called document delivery services.
5. The right to transfer any or all rights mentioned in this agreement, as well as rights retained by the relevant copyright clearing centers, including royalty rights to third parties.

Hogrefe OpenMind: Information about the open access publishing program Hogrefe OpenMind, including the article processing fee and the Creative Commons license under which the article will then be published, are given at www.hogrefe.com/j/apahf

Online Rights for Journal Articles: Guidelines on authors' rights to archive electronic versions of their manuscripts online are given in the "Guidelines on sharing and use of articles in Hogrefe journals" on the journal's web page at www.hgf.io/apf

April 2021


EAAP
EUROPEAN ASSOCIATION FOR AVIATION PSYCHOLOGY