Article:

Becker, T. and Ayton, P. (2024) Effects of flight crew role assignment on aviation accidents and incidents: Evidence of a systemic safety issue. Safety Science, 170. 106352. ISSN: 0925-7535

https://doi.org/10.1016/j.ssci.2023.106352

Effects of Flight Crew Role Assignment on Aviation Accidents and Incidents: Evidence of a Systemic Safety Issue

Tom Becker[1] & Peter Ayton[2]

[1]Department of Psychology City, University of London

[2]Centre for Decision Research, University of Leeds

Author Note

Tom Becker: https://orcid.org/0000-0002-0354-6338

Peter Ayton: https://orcid.org/0000-0003-2285-4608

Declaration of interests: none.

Correspondence concerning this article should be addressed to

Tom Becker, Im Holze 7b, 28355 Bremen, Germany.

Email: mail@tom-becker.net

Abstract

To investigate whether a reported association between flight crew role assignment and accident/incident frequency has persisted subsequent to reforms introduced to address it, we analyze global civil aviation data on 841 accidents and safety critical incidents for the period 2000-2020 resulting in 5318 fatalities. We find that significantly more such events occur, and significantly more fatalities result, when the Captain (Pilot-in-Command) rather than the Co-pilot (Second-in-Command) is acting on the controls as “Pilot Flying” rather than acting as “Pilot Monitoring” – prima facie evidence that the regular combination of command and control (Pilot-in-Command as Pilot Flying) presents a significant systemic safety risk. The yearly proportion of events with the Pilot-in-Command as Pilot Flying significantly increased over 2000-2020. Moreover most (72.2%) events occurred in the absence of any emergency, in technically airworthy aircraft and most (87.9%) were also judgmentally assessed as preventable by the flight crew. Our findings are consistent with accounts of the crew assignment effect that invoke role-dependent status hierarchy effects interfering with effective monitoring and with cognitive overload debilitating the Pilot-in-Command when assigned as the Pilot Flying. We interpret these findings as evidence that measures specifically introduced to implement effective flight crew teamwork, such as Crew Resource Management training, fail to prevent ineffective flight crew teamwork. We consider the implications of our findings for the aviation industry and present options for mitigation and issues for further research

Keywords: Human Factors, Teamwork, Hierarchy, Monitoring & Intervention, Risk Management

1. Introduction

The importance of effective flight crew teamwork for the prevention of aviation accidents has been confirmed by various findings of accident and incident investigations (Broome, 2011; Jentsch, Barnett, Bowers, & Salas, 1999; NTSB, 1994; Orlady, 1982) and safety research (Driskell & Salas, 1992; Mosier, Fischer, & Orasanu, 2011; Paris, Cannon-Bowers, & Salas, 2000; Salas, Burke, Bowers, & Wilson, 2001; Salas, Wilson, Burke, Wightman, & Howse, 2006). By the late 1970s research by NASA found that several jet transport accidents in the US involved failures of decision making, leadership, pilot judgment, communication and crew coordination (Cooper, White, & Lauber, 1980). According to Helmreich (2006) this sparked the development of specific teamwork training for pilots, known today as Crew Resource Management (CRM) training and which emerged in the early 1980s (Helmreich & Foushee, 2010) to address the human factor in accidents as initially highlighted by Beaty (1969).

As described by Helmreich (2006) CRM training was initially focused on correcting behavioral deficits primed by status hierarchy effects such as a lack of assertiveness by junior Co-Pilots and authoritarian behavior by the pilot in command (PIC) of the flight. One prominent incident, one of many exemplifying the need for this training (see Cooper et al., 1980), was the world's worst ever aviation accident on Tenerife (1977) with 583 fatalities when two B747 aircraft collided on the runway. A contributing factor identified in the investigation[1] of this accident was a steep cross-cockpit-authority gradient manifested in suppression by the PIC of critical concerns raised by the other crew members.

Since its introduction more than four decades ago CRM-training has evolved and, in 1995, became a mandatory training item for all pilots (International Civil Aviation Organization (ICAO), 1998). In 2005 CRM-training was augmented by training in a practical

[1] https://www.skybrary.aero/accidents-and-incidents/b742-b741-tenerife-canary-islands-spain-1977

concept called threat and error management (TEM) to help pilots with their risk-management (Klinect, 2005; Merritt & Klinect, 2006); accordingly, for the past twenty years, CRM/TEM-training covering pertinent cognitive and social skills such as effective communication, workload-management, decision-making, monitoring and intervention has been viewed as comprehensive human factor training and recognized as necessary to safely operate an aircraft. Collectively referred to as non-technical skills (NOTECHS) these skills complement the required technical skills such as manual flying and procedural skills (European Union Aviation Safety Agency (EASA), 2023; Flin et al., 2003). These technical and non-technical skills, combined with defined appropriate knowledge, make up the set of competencies that all professional pilots, irrespective of their hierarchical position or functional role within a flight crew, are required to be trained and assessed in so as to gain and annually revalidate their pilot license as those globally agreed competencies are presumed necessary for safe aircraft operation (ICAO, 2020).

In the interests of providing a coordinated workflow on the flight deck the teamwork of a dyadic flight crew is additionally organized by a prescribed role framework. Firstly, there are the hierarchical positions of the Pilot-In-Command (PIC), also known as Commander[2], and the Second-In-Command (SIC), also known as the Co-Pilot, defining legally that the PIC has the final responsibility for the safe operation of the aircraft (ICAO, 2022, sec. 4.5.1). Secondly, there are two distinct functional intra-cockpit roles, each allocated to one of the two crew members: (1) the pilot flying (PF), sometimes called the handling pilot or pilot on controls and (2) the pilot monitoring (PM), previously - until 2003 (Federal Aviation

[2] To act in the role of the PIC on a certain aircraft type, pilots require an airline specific command training and a respective endorsement in their pilot license. On successful completion of such training pilots usually upgrade in their rank within an airline from Co-Pilot to Captain. There is only one Captain rank, but within Co-Pilots there might be different ranks such as Second Officer, First Officer, or Senior First Officer depending on internal airline regulation, flight experience and training. However, these Co-Pilot ranks are not a regulatory requirement. When two Captains (PIC-rated pilots) are flying together one is designated as the PIC and the other as the SIC/Co-Pilot. For flying medium and large transport aircraft, PICs need an Airline Transport Pilot License (ATPL) which requires a minimum of 1500 hours of flight experience (EASA, 2022; FAA, 2023).

Administration (FAA), 2003) called the Pilot-Not-Flying (PNF). The PF operates the controls of the aircraft, either manually or via the autopilot system, while the PM is responsible for monitoring the aircraft's flight path, the aircraft systems and the actions of the PF (FAA, 2017; Flight Safety Foundation (FSF), 2014). The decision as to which of the pilots acts as PF and as PM is ultimately up to the PIC, but it is common practice in commercial aviation that the two pilots alternate these roles by each flight sector (Limor & Borowsky, 2020, p. 45; NTSB, 1994, p. 37; Scott, 2013, sec. 3.4.4.6; Sumwalt, Cross, & Lessard, 2015, p. 15). In only rare cases is there a requirement that the PIC must act as PF, e.g. during certain low visibility weather conditions or within operation from or to certain airports when mandated by the regulatory authorities of the airport's state or the aircraft operator (airline), e.g. Aspen or Funchal. Nevertheless, the subsequent flight will then usually be flown with the SIC as PF. Accordingly in approximately one half of the flights performed by commercial flight crews, the higher-ranking and usually (though not always) more experienced crew member, the PIC, performs the role of the PF, and the lower-ranking crew member, the SIC, performs the role of the PM. In the other half of the flights, the roles are reversed.

Evidence that the quality of the flight crew's team performance depends on the crew assignment to these roles (whether PIC or SIC is assigned as PF or PM) – and that flight safety is affected by the role assignment – has been reported by a number of sources. An early quantitative study was reported by Orlady (1982) who analyzed 245 pilot-reported incidents and found that more near midair collisions, takeoff anomalies, and crossing altitude deviations were reported when the PIC was PF; however more altitude deviations, near midair collisions during approach, and landing incidents occurred when the SIC was PF. Orlady also reported that a monitoring failure by the PM preceded a significant operational anomaly in most cases.

In 1994 the US National Transportation Safety Board (NTSB) reported an analysis of all the (37) major flight-crew-related accidents involving US carriers that occurred between

1978 and 1990 (NTSB, 1994). The crew assignment at the time of the accident was determined for all 37 accidents which revealed that in 30 of the 37 accidents the captain was the PF at the time of the accident (NTSB, 1994, p. 38) . Noting this striking inequality – that in more than 80% of the 37 accidents reviewed the PIC was the PF - the report states that: "The Safety Board was unable to determine any particular significance to, or draw any conclusions from, this finding." Despite this inconclusiveness they also found that among the most common errors identified in these events was the "failure to monitor or challenge the other crewmember's errors"; monitoring and challenging errors were found in the majority (31) of the 37 accident cases.

Further research into the issue of flight crew role assignment was conducted by Jentsch et al. (1999) who analyzed 221 reports of US air carrier incidents in which an error or errors occurred that could have compromised flight safety. The analysis used evaluations made by aviation professionals in order to determine loss of situational awareness and tactical decision-making errors in relation to flight crew role assignment. These authors found that more incidents (142 of 221 [64%]) happened with the PIC as PF than the SIC as PF (79 of 221 [36%]) and also found that PICs lost situational awareness more often and made more tactical decision-making errors when they were acting as PF compared to when acting as the PM which they attributed to the additional workload for the PIC as PF. Jentsch et al. concluded that the combination of aircraft control and problem solving, at times when the situation demands the full use of the PIC's cognitive skills, might be a burden for effective accident prevention and could account for both their observation that incidents occur when the PIC is PF and the NTSB finding that more serious accidents occur when the PIC is PF. Such an effect may be aggravated by their finding, as in the Orlady (1982) and NTSB (1994) studies, of failures of monitoring: "…(Co-Pilots)… seemed to be somewhat ineffective at monitoring and challenging the captains' errors, especially when the captains were at the controls" (Jentsch et al., 1999, p. 11).

In a simulator study exploring the relationship between hierarchical role (PIC/SIC) and functional flight deck position (PF/PM) Palmer, Lack, & Lynch (1995) found that the status difference between PIC and SIC significantly influenced pilot behavior in the functional roles of PF or PM: PICs initiated more in-flight transfers of control (i.e. from PF to PM or vice versa) and gave more direct commands in either functional role more often than SICs.

A systematic review of the empirical literature on the effect of crew role assignment on flight safety outcomes by Beveridge, Henderson, Martin, & Lamb (2018) summarized substantial evidence of an effect of crew role assignment on performance measures of three factors directly affecting flight safety – namely the flight crew's monitoring, situational awareness and decision-making. Almost all of the reviewed studies (16 of 18) observed significant relationships between crew role assignment and their outcome measurement. The two studies that did not find any such effect were accident/incident data reviews (Boss, 2012; Khatwa & Roelen, 1997) though two other included accident/incident reviews did report effects (Khatwa & Helmreich, 1999; NTSB, 1994).

Beveridge, et al. attributed crew role assignment effects to three specific differences between the PIC and SIC: (1) differences in the cognitive workload of each pilot (when the PIC is PF the cognitive burden of aircraft control restricts the PIC's ability to maintain situational awareness and impacts PIC decision-making); (2) differences in the relative expertise of the PIC and SIC (due to inexperience SICs are more likely to lose situational awareness - critical for effective monitoring and compounding the workload effects on the PIC when the PIC is PF); (3) differences in the relative status and authority of the PIC and SIC (e.g. status and authority factors may inhibit the SIC as PM from communicating any observations or concerns; lack of assertiveness or corrective action by the SIC was cited as significant in the majority of the accident/incident review studies). Beveridge, et al.

interpreted the research findings as indicating that there may be a greater number of inherent obstacles for optimal crew performance with the PIC as pilot flying

Subsequent to the publication of Beveridge, et al.'s review Behrend & Dehais (2020) also found evidence for differences in the behavior of the PIC and SIC in a simulator study on go-around[3] decision-making of 62 commercial pilots. In a simulated landing, where aircraft displays showed a tailwind speed for which company procedures permitted continuation of the approach, but mandated a go-around decision, only about half the pilots chose to go-around. When the pilot being tested was in the PF-role, decisions did not significantly differ between pilots of different rank; but, when in the PM-role, significantly more of the PICs than SICs initiated a go-around.

This brief review of the empirical literature indicates that both ineffective flight crew teamwork *per se,* as well as its association with flight crew role assignment, are long-standing safety issues in aviation. The measures taken by the industry to date to address this issue have been the establishment and development of human factor training together with initiatives to highlight the importance of the role of the PM which includes the re-labelling of the PNF to PM in 2003 as described above (FAA, 2003). This change was initially proposed by Sumwalt, Thomas, & Dismukes (2002) who remarked that: "Although this change may seem small, we felt that it was important because it described what that pilot should be doing (monitoring) versus what he/she is not doing (not flying)." In justifying the name change the FAA (2003) explicitly acknowledged that "Studies of crew performance, accident data, and pilots' own experiences all point to the vital role of the non-flying pilot as a monitor. Hence, the term pilot monitoring (PM) is now widely viewed as a better term to describe that pilot."

The 2003 Federal Aviation Administration paper also advocated the improvement of crew monitoring performance via the development and implementation of effective Standard

[3] A "go around" is when, during an approach or landing the flight crew elects to abort the approach or landing. It is a normal flight procedure for which pilots are trained and assessed annually on a recurrent basis.

Operating Procedures (SOPs) to support monitoring and cross-checking functions, by training crews on monitoring strategies, and by pilots following those SOPs and strategies. Subsequently further industry initiatives have targeted the monitoring behavior of pilots by promoting strategies for improving monitoring performance (Civil Aviation Authority (CAA), 2013; FSF, 2014) .

In specific regard to the effects of crew assignment there have also been calls for crew assignment to be altered. Thus, Jentsch et al. (1999) concluded that: "The current results point toward a fundamental yet probably unwelcome change for captains: Having the captain at the controls may not always be the best course of action. Instead, the results seem to indicate that captains should consider letting their FOs fly while they concern themselves with the big picture." (p. 12). A similar recommendation was expressed in a report edited by Sexton (2004) considering how teams operating in high-risk-environments should work together to deal with crisis situations to best effect. The report argued that captains are overloaded with multi-tasking when they try to accomplish both the pilot flying and the pilot in command duties and proposed that: “The data suggest that if a crew encounters a high workload situation when the captain is pilot flying, it is best to cede control of the aircraft to the first officer” (Sexton, 2004, p. 18).

Such thinking has had at least some impact on policy. Jentsch et al. (1999) reported that: “At least one major U.S. airline already advises its captains in training to consider handing over control to the FO in an emergency”. They also cited Stewart (1992) who, even before publication of the NTSB (1994) paper, revealed that a major European airline had instituted a policy mandating that, regardless of who is assigned as the PF for the flight, the FO (SIC) is at the controls from the enroute descent point to decision altitude. “Such a procedure leaves the captain more free to monitor the progress of the flight in the more difficult flying environments” (Stewart, 1992, p. 262)." (A description of this “monitored approach” procedure can be found on www.picma.info).

In order to determine the extent to which these efforts have mitigated the issue of ineffective flight crew teamwork and its association with flight crew role assignments, we sought empirical data about aviation accidents and incidents covering more recent periods – specifically since CRM training was fully implemented in pilot training (1995) and the label change from PNF to PM had been implemented (2003). The only available accident/incident review carried out subsequent to these reforms (Boss, 2012) found no association between crew assignment and accident frequency; however, this study was rather limited in scope. Although Boss (2012) analyzed all major U.S. air carrier accidents between 1991 and 2010 for which the NTSB conducted a major investigation, this was a small set of events (50) and included only 19 that occurred after the year 2000 – too few to analyze statistically. For this reason, we attempted a larger scale analysis of more events. To that end, as well as studying reports of accidents, we also studied reports of safety critical incidents, as investigated by earlier research (Jentsch et al., 1999; Khatwa & Helmreich, 1999; Orlady, 1982). We analyzed 841 events – substantially more than in any other comparable study[4] - that occurred globally over the period 2000-2020 to investigate whether the association between flight crew role assignment and accident/incident frequency has persisted subsequent to the reforms introduced to address it.

2. Method and Data

2.1. Database Selection

Like other high-risk-domains the commercial aviation industry routinely analyzes their mishaps to learn lessons for future prevention. The conduct of the investigation and reporting of aviation accidents and incidents is regulated on a global basis stipulating that the state in which the event occurred is responsible for the investigation (ICAO, 2020a). Therefore,

[4] The largest number of events included in any comparable study is 245 incidents (Orlady, 1982). Khatwa & Roelen (1997) analysed 156 fatal accidents though were only able to identify the PF for 24. Khatwa & Helmreich (1998) analysed crew assignment for 71 fatal accidents and 5 serious incidents but could only identify PF for 58 of these.

records of investigation reports are, initially at least, only collated by national investigation bodies. These often only contain selected events, and, in some cases, reports are not drafted in English making it difficult to gather comprehensive global data for research. Although other industry stakeholders such as aircraft manufacturers, regulators and trade associations also maintain accident and incident databases for safety relevant events only ICAO offers a single central and freely accessible repository that theoretically records all aviation accident and incident reports on a global basis. However the ICAO database is incomplete: some earlier paper based records are not all included and it is also restricted to final reports which are not always produced for every incident or accident as recently highlighted by an independent safety organization (FSF, 2022, p. 5).

Notwithstanding these shortcomings we identified a more comprehensive database managed by the Jet Airliner Crash Data Evaluation Centre (JACDEC) in Hamburg, Germany, which not only collates information about safety relevant events in aviation from various different available sources, (e.g. the official websites of the accident investigation branches sometimes provide interim reports for events with no final report), but also presents the raw data of the reports collated in a format usable for statistical analysis. The database is professionally maintained, such that the content is based on the internationally agreed Accident/Incident Data Reporting (ADREP)[5] standards and offers a commercial service for researchers, travelers, and insurance companies. It focusses on commercial aviation including only aircraft with >19 seats. Their website[6] claims that the database includes "all known accidents, hull losses[7], serious incidents and incidents in civil aviation… back to 1969".

ICAO defines an accident as "an occurrence associated with the operation of an aircraft …. in which a person is fatally or seriously injured…the aircraft sustains damage or

---

[5] Further information on the details of ADREP and how an accident/incident is defined in aviation can be found in https://skybrary.aero/articles/icao-adrep

[6] https://www.jacdec.de/accident-incident-database/

[7] A hull loss defines an accident where an aircraft was destroyed or damaged beyond repair.

structural failure…the aircraft is missing or is completely inaccessible…”. An incident is defined as “an occurrence, other than an accident, … which affects or could affect the safety of operation”, a serious incident is defined as “an incident involving circumstances indicating that there was a high probability of an accident …” (ICAO, 2020a).

Note that the JACDEC event categories differ somewhat from those used by ICAO: rather than using the ICAO category of “accidents”, JACDEC identifies those accidents where an aircraft was destroyed, or damaged beyond repair, as “Hull losses”. Accidents not resulting in a hull loss are included in the category “Serious incidents”, which also includes events that are not, by the ICAO definition, “accidents”, but serious incidents.

At the time of our analysis in April 2021 the JACDEC database comprised 17,795 worldwide occurrences which comprised 4,885 events categorized by JACDEC as hull loss accidents, 7,220 as serious incidents and 5,690 as incidents (which included some accidents according to ICAO classification). It therefore provided the largest set of global aviation event data freely accessible for research that we were able to find. Access to this database can be purchased so that information about events can be downloaded from it.

2.2. Event selection

Not all events within the database are appropriate for the study of flight crew role assignment. For events occurring during maneuvers on the ground (e.g. runway incursions, collision with other aircraft, vehicles, or obstacles on the taxiway or the ramp) the role assignment is usually fixed to be the PIC as PF, either as a result of the technical design of certain aircraft not providing any technical means for the SIC to taxi the aircraft (e.g. B737 aircraft), or by virtue of company regulation as most airlines in the world prohibit the SIC from controlling the aircraft on the ground even if the technical means for this are available.

Because our focus is on events for which the role assignment is interchangeable between the PIC and SIC, we therefore excluded those ground events not associated with takeoff or landing operations. We also excluded all events exclusively attributed to technical malfunctions or emergencies[8] as the primary accident factor in the database so that no obvious pilot actions were identified with either the cause or the outcome of the events. We also excluded all events in which more than a single aircraft was involved (e.g. for air traffic control related events such as loss of separation or other air proximity events) to avoid the complication of multiple determinations of role assignments for a single event. This left us with events in the database coded as runway excursions (RE), collisions with obstacles during take-off or landing (CTOL), abnormal runway contact (ARC), under- or overshoots (USOS), controlled flight into terrain (CFIT), loss of control inflight (LOC-I) and low altitude (LALT)[9]. After application of these exclusion criteria and restricting events to those in the period 2000-2020 we obtained records for 3,335 events, 405 of which were coded as incidents, 1,931 as serious incidents and 999 as hull losses, together accounting for 11,780 fatalities.

We then applied three further data exclusion criteria: firstly we excluded all military aircraft; secondly, to ensure that we only investigated events involving civil commercial aircraft operated by an airline flight crew consisting of at least two qualified pilots, we excluded all small aircraft (Maximum Take-off Weight < 15t, e.g. small turboprop or business jet aircraft) as some of these might also be operated by a single pilot or the type of operation may not be comparable to commercial airlines in terms of pilot qualification and training; thirdly we excluded aircraft requiring additional crew members such as navigators or radio operators to ensure comparability of events regarding task sharing and teamwork. This left a final sample of 2,293 events comprising 370 Incidents, 1,459 Serious Incidents and 464 hull losses which together accounted for 9,256 fatalities.

---

[8] JACDEC allows to filter and select events with technical malfunctions or inflight fires only as it uses the official ADREP acronym for technical aircraft malfunctions such as System Component Failures (SCF) - non-power plant (NP) and power plant (PP) and Non-Impact Fires (F-NI).

[9] According to the International Civil Aviation Organization (ICAO) the accident types of RE, CFIT and LOC-I represent the highest accident frequency and fatality risk in commercial aviation (ICAO, 2020c, p. 27).

2.3. Event coding

Our analysis studied the relative frequency of events as a function of four factors: (1) *Role Assignment* is a binary variable reflecting whether the PIC or SIC was the PF as gleaned from the context or investigation reports. We read all the investigation reports and event descriptions of the 2,293 events. Given the limitations of the information provided in the event reports (e.g. not always containing relevant information on the role assignment as it is still not a mandatory investigation requirement) we were able to make assessments of *role assignment* for 841 (36.6%) of the events. For those events when, during the flight, there was a planned or unplanned role change from the original role assignment (n=55) we determined the PF as the pilot whose control input was consequential for the event. E.g. if after a bounced landing by the SIC the PIC took over control and subsequently crashed the aircraft, we coded the PIC as PF. When the SIC took away control from the PIC it was coded as SIC as PF.

The aircraft in our sample of events have flight controls at both pilot stations though, normally, only the PF operates her/his controls. However, there were some cases of *dual input* (n=55) when both the PF and the PM were acting on their respective controls simultaneously. These were coded as PIC as the PF. Dual inputs can either be mutual, e.g. by both pilots pushing the controls such as the aircraft elevator or brakes in the same direction (n=33), or in opposition, e.g. by pilots pushing the controls in different directions (n=22), which is possible on most aircraft equipped with an electronic side stick as aircraft control.

(2) *Mode of Operation* is a binary variable reflecting whether the aircraft was in *Normal Operation* or *Non-Normal Operation* as defined by the aircraft's technical and flight status. *Non-Normal Operation* was indicated when the aircraft was not technically airworthy or there was an aircraft related emergency (i.e. inflight smoke or fire). In the absence of any technical failure or aircraft related emergency the event was deemed to have occurred in *Normal Operation.* For the purposes of this study, we considered threats due to weather, terrain or time pressure, provided there was no aircraft related technical malfunction or

emergency, as occurring in normal operation: they should always be safely manageable by a well-trained flight crew using given TEM-strategies. Although some aircraft operating manuals include certain flight maneuvers (such as windshear- or upset-recovery, terrain- or traffic avoidance or rejected take-off procedures) in their non-normal operation section these maneuvers are applicable regardless of whether the aircraft is technically airworthy or not. Accordingly, we coded events in which these maneuvers were used as *Non-normal* only when the aircraft was not technically airworthy or there was an aircraft related emergency (i.e. inflight smoke or fire). We were able to determine the *mode of operation* for 1,076 (46.9%) of the 2,293 events in our sample.

(3) *Teamwork Behavior* is a binary variable indicating whether or not “poor CRM” or “poor teamwork” was a contributory or causal factor in the event. This may be explicitly referred to in investigation reports but in some cases, when no such direct information or conclusion was provided in the report, but when, in the opinion of the first author, sufficient information included in the investigation report or event description clearly merited it, a judgment was made by the first author that *Teamwork Behavior* was an issue in the event. Given the dyadic nature of the flight crew setup this included cases when mutual intervention was obviously missing or ineffective, but nevertheless was evidently realistically possible, e.g. when errors or deviation from standard operating procedures by one pilot occurred but were not trapped and mitigated by the other pilot although he/she had an option to do so. Notwithstanding such cases it should be noted that errors by one pilot not trapped and mitigated by the other pilot were not necessarily coded as a failure of *Teamwork behavior;* if, for example, there was insufficient time or opportunity for a mitigating response by the other pilot then poor CRM or poor teamwork would not be recorded. The first author is an active airline pilot with experience as an airline safety manager, accident investigator, training captain and human factors facilitator. The reliability of these judgements was tested and confirmed by comparison with judgments made by an independent expert on a subset of the

judged cases (see details described below). We were able to determine the *teamwork behavior* for 834 (36.4%) of the 2,293 events in our sample.

(4) *Preventability*, is a binary variable denoting whether or not there was a realistic opportunity for preventing the event by pilot behavior, either by an individual pilot or by the flight crew team. Preventability was determined from investigation reports; events were coded as pilot-preventable if the report explicitly indicated this - e.g. by stating “the pilots should have initiated a go-around” or the “accident could have been prevented”. As with the coding of *Teamwork Behavior*, in cases when no such explicit information or conclusion was provided a judgment was nonetheless made by the first author that an event was pilot-preventable when information included in the investigation report or description of the event made this clear. Examples include cases when it was obvious that a more risk-averse option or a realistic option for error prevention or effective mutual intervention had been available to the flight crew. We were able to determine the *preventability* for 907 (39.6%) of the 2,293 events in our sample.

Complete listings of the downloaded events selected for each of the reported analyses are available online at OSF https://osf.io/h5aj8/files/osfstorage/64c69585c7ab290ca3d4df62.

2.3.1 Data coding reliability

To cross-check the reliability of the data coding assessments a sample of 32 events was assessed by an independent expert briefed as to the determinations that should be made according to the event coding detailed above. The independent expert is an active and experienced airline and ‘check pilot’ (senior examiner for Airbus and Boeing aircraft types). The expert’s assessments for *Mode of operation*, *Crew assignment* and *Preventability* were in complete (100%) agreement with the researchers’ assessments. For *Teamwork Behavior* two events were rated differently resulting in high (93.8%) but somewhat less than perfect levels of agreement. Although the assessments of *Teamwork/Behavior* and *Preventability* may be critiqued as judgmentally determined, the high levels of inter-rater agreement support the validity of these assessments.

3. Results

3.1. Role assignment

Table 1 depicts the 841 events in our sample of 2,293 events for which we were able to determine the crew role assignment. Given the assumed equal distribution of flight sectors between the PIC and SIC, there is no reason to expect any difference in the frequency of events as a function of the two different possible role assignments - unless the crew role assignment itself influences the propensity for accidents and serious incidents. Nonetheless, plainly the two proportions of events are substantially different. Almost two and a half times as many events occurred when the PIC was acting as PF (n = 597, 71.0%) as when the SIC was acting on the controls (n = 244, 29.0%). A binomial test for equal proportions confirms that the events were not equally frequent for each role assignment: significantly more events occurred when the PIC was acting on the controls, $p < .001$.

Table 1

Frequencies (n) and Percentages (%) of Events and Fatalities by Role Assignment

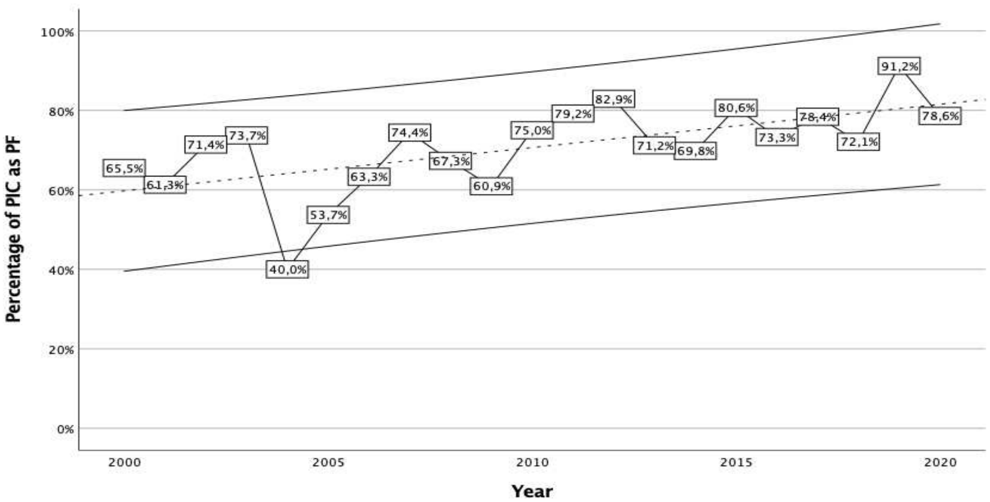
	Role Assignment											
	Events						Fatalities					
	PIC as PF		SIC as PF		TOTAL		PIC as PF		SIC as PF		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	163	79.9	41	20.1	204	100	4078	76.7	1239	23.3	5317	100
Serious Incidents[a]	394	68.2	184	31.8	578	100	1	-	0	-	1	-
Incidents[a]	40	67.8	19	32.2	59	100	0	-	0	-	0	-
All events	597	71.0	244	29.0	841	100	4079	76.7	1239	23.3	5318	100

[a] Percentages of fatalities for serious incidents and incidents are not computed due to a lack of data.

Binomial tests confirm that this difference is also statistically significant for each event type depicted in table 1; significantly more hull losses, significantly more serious incidents and significantly more incidents occurred when the PIC was acting on the controls than when the SIC was acting on the controls (Hull losses: PIC as PF (n = 163, 79.9%), SIC as PF (n = 41, 20.1%), $p < 0.001$; Serious incidents: PIC as PF (n = 394, 68.2%), SIC as PF (n = 184, 31.8%), $p < 0.001$, Incidents: PIC as PF (n = 40, 67.8%), SIC as PF (n = 19, 32.2%), $p = 0.009$). Nevertheless, and indicating that more serious outcomes result when the PIC is PF, a chi-squared test reveals that the association between the role assignment and the frequency of events is stronger for hull losses than for serious incidents ($\chi^2(1, N = 782) = 10.13$, $p = 0.001$).

To investigate whether the role assignment effect had changed over the 21-year period covered by our dataset we plotted the percentage of events when the PIC was acting as PF for each year over the period 2000-2020 (see Figure 1). Counter to the expectation that reforms to crew resource management and the pilot label change would reduce the role assignment effect, Kendall's tau revealed a significantly *increasing* monotonic trend in the annual proportion of events where the PIC was PF (n = 21) over the period 2000-2020 ($\tau_b = 0.46$, $p = .003$).

Figure 1: *Percentages of the PIC as PF per year*



Note: The dotted line depicts the regression fit line. The two solid lines depict the 95% confidence interval around that line.

Within the sample of 841 events for which we were able to determine the role assignment, 90 events (10.7%) were fatal, accounting for 5,318 fatalities; the other 751 events (89.3%) were non-fatal. Binomial tests show that significantly more fatalities result with the PIC as PF (n = 4,079) than with the SIC as PF (n = 1,239), $p < .001$, and that the PIC was acting as PF significantly more often in both event types (Fatal: PIC as PF (n = 72, 80.0%), SIC as PF (n = 18, 20.0%), $p < .001$; Non-Fatal: PIC as PF (n = 525, 69.9%), SIC as PF (n = 226, 30.1%), $p < .001$. A chi-squared test reveals that the pattern of the PIC as PF is also significantly stronger for fatal than for non-fatal events, ($\chi^2(1, N = 841) = 3.98$, $p = 0.046$) consistent with the notion that more serious outcomes result when the PIC is PF. Given the strong manifestations of a crew role assignment effect on events, fatal events and fatalities, it is not surprising that Mann-Whitney U tests indicate that the number of fatalities *per event* ($M = 6.3$, $SD = 27.6$) differs significantly by role assignment for all events in the sample (PIC as PF = 6.8, SIC as PF = 5.1), z = -1.99, $p = .047$. However, there is no such difference in the number of fatalities per event ($M = 59.1$, $SD = 63.6$) for the smaller set of 90 fatal events (PIC as PF = 56.7, SIC as PF = 68.8), z = 0.06, $p = .956$.

We also investigated when *changes* of role assignment had been initiated, either by an explicitly announced change of role assignment or *de facto* by simultaneous input on the controls by both pilots (dual input). In aggregate a total of 110 (13.1%) of the 841 events for which we were able to determine the role assignment involved role changes; 94 of these (85.5%) were events where the SIC had originally been the PF at the controls. In half (55) of the 110 events there was an announced control change from one pilot to another. Consistent with the notion that at least some PICs judge it best that they, rather than the SIC, be at the controls in demanding situations (but inconsistent with the recommendation of Sexton (2004), the vast majority of these announced control changes involved the PIC taking control from the SIC (52, 94.5%) rather than vice-versa (3, 5.5%) - a difference that a binomial test reveals is unlikely to be due to chance ($p < .001$). For the other 55 of the 110 events there was a deviation from the prescribed division of tasks between the PF and PM such that both pilots gave input to the controls *simultaneously*: for 33 of these events both pilots were acting on their controls in the same direction (e.g. during braking on the runway during landing or inflight when deemed necessary to aggregate a control input) and in 22 events they were acting in opposition to each other, despite the fact that aircraft operating procedures do not allow contradictory control inputs by the pilots but require the PM to *officially* take over control by announcing “I have control”. For both types of dual control cases the SIC was originally assigned as PF in significantly more of these events: for the 33 events in which we detected input in the same direction the SIC was originally the PF in 25 (75.8%), $p = 0.005$; for the 22 events in which we detected contradictory control inputs the SIC was originally the PF in 17 (77.3%), $p = 0.017$.

Note that the presence of inflight role changes does not allow for the role assignment effect to be substantially attributed to changes of control such that the PIC spuriously more often appears to be the PF because of inflight role changes: even allowing for all of the 107 inflight role changes that resulted in the PIC being coded as PF and considering these as SIC as PF, there are still significantly more events with the PIC as PF than the SIC as PF (PIC as PF = 490; SIC as PF = 351, $p < .001$).

Finally, given that the logic of our inferences regarding the role assignment effect depends critically on the assumption that the two crew assignments (PIC as PF, SIC as PM and SIC as PF, PIC as PM) operate equally frequently, we exploited an opportunity to subject this assumption to empirical test. One way this might be evaluated is by looking at the distribution of crew assignments for events that are independent of any pilot behaviors – this should reflect the underlying distribution of crew assignments. One category of event that can be reasonably assumed to be unrelated to role assignment - and hence to be unaffected by any selection effects stemming from pilot behavior - is bird or other wildlife strikes.

Within our sample of 841 events for which we were able to determine the crew role assignment, only 2 events were recorded by JACDEC as involving bird or wildlife strikes, therefore we conducted a separate download of JACDEC events that involved bird or wildlife strikes (n = 421). However, only 38 of those events had investigation reports available so we additionally searched another event database (www.skybrary.aero) for bird or wildlife strike events over the same time period (2000-2020). As these kinds of events are only infrequently investigated in depth, we were able to determine the role assignment for only 30 of the total 61 events with a report available. Given the assumption that a bird or wildlife strike event is independent of pilot behavior we have no reason to hypothesize that they would be more prevalent with one role assignment than another - unless that role assignment was in fact more prevalent in practice. In short, the relative frequency of bird strikes for each crew assignment should reflect the relative frequency of each role assignment in practice.

When we investigated the role assignment in these 30 bird or wildlife strike events, we found that there were 16 events occurring with the SIC as PF and 14 with the PIC as PF. A binomial test showed no statistically significant difference between the two role assignments, $p = .856$. This confirms our assumption, consistent with airline practice of alternating crew role assignment, that each role assignment is indeed operated equally often in commercial aviation practice.

3.1.1 Role Assignment by geographic region

Our analyses of the role assignment effect are based on a global sample of events. Nonetheless the role assignment effect might not be a global phenomenon; for example, it might, conceivably, be restricted to some parts of the world. To investigate this issue and elicit possible geographical differences in the role effect we split the sample based on ICAO defined regions. ICAO has segmented the world into the following eight regions[10] (in

[10] https://www.skybrary.aero/articles/icao-regions

alphabetical order): Africa-Indian Ocean (AFI), Asia and Middle East (MID / ASIA), Caribbean (CAR), European (EUR), North American (NAM), North Atlantic (NAT), Pacific (PAC) and South American (SAM). As the role effect might possibly vary as a function of possible influences stemming from different norms and cultures within different airlines around the world, we determined, for each of the 841 events, the location of the operator's (airline's) headquarters and sorted them to the relevant regions. For simplicity we diverted slightly from the original segregation of ICAO by dropping the NAT and PAC regions (the only affected headquarters (Iceland) was included in Europe), and merged the CAR and SAM regions to a single SAM region. Additionally, we removed the southern part of the EUR Region, within which ICAO includes some of Africa's most northern countries, so that Europe only included the European continent and combined it with AFI so as to include the whole African continent. Table 2 exhibits the findings which confirm by means of binomial tests the presence of a statistically significant effect in all regions in the world, $p < .01$.

Table 2

Frequency of Events by Role Assignment and Geographic Region

	Role Assignment		
	PIC as PF	SIC as PF	TOTAL
Region	Events	Events	Events
Africa	41 (75.9%)	13 (24.1%)	54 (100%)
Asia and Middle East	204 (76.1%)	64 (23.9%)	268 (100%)
Europe	214 (66.9%)	106 (33.1%)	320 (100%)
North America	103 (65.6%)	54 (34.4%)	157 (100%)
South America and Caribbean	35 (83.3%)	7 (16.7%)	42 (100%)
Total	597 (71.0%)	244 (29.0%)	841 (100%)

3.2. Mode of Operation

As shown in table 3 the majority - 777 (72.2%) - of the 1,076 events for which we were able to determine the mode of operation (normal or non-normal) occurred during normal operation, meaning with technically airworthy aircraft and no emergency present; only 299 (27.8%) of the events occurred in the presence of any aircraft-related abnormality or emergency; a binomial test for equal proportions showed that this difference was statistically significant, $p < .001$. Another binomial test revealed that significantly more of the 6102 fatalities occurred in normal operation (4,181, 68.5%) than in non-normal operation (1,921, 31.5%), $p < .001$.

Table 3

Frequencies (n) and Percentages (%) of Events and Fatalities by Mode of Operation

	Mode of Operation											
	Events						Fatalities					
	Normal		Non-Normal		TOTAL		Normal		Non-Normal		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	158	60.5	103	39.5	261	100	4179	68.5	1920	31.5	6099	100
Serious Incidents[a]	551	77.3	162	22.7	713	100	2	-	1	-	3	-
Incidents[a]	68	66.7	34	33.3	102	100	0	-	0	-	0	-
All events	777	71.0	299	29.0	1076	100	4181	76.7	1921	23.3	6102	100

[a] Percentages of fatalities for serious incidents and incidents are not computed due to a lack of data.

Recall however that in our selection of events as described in section 2.2 we excluded all events attributed solely to technical aircraft malfunctions and other emergencies such as onboard fires when we downloaded the set of events to analyze the role assignment effect. As a result, we were not able to determine how many of the events solely attributed to technical failures or emergencies as primary event causation were excluded. As these events would plainly all qualify as occurring in non-normal operation their exclusion will clearly affect our measure of the proportions of events occurring in normal and non-normal operation. Therefore, in order to estimate the proportions of events in normal and non-normal operation that would have resulted had we not excluded all events attributed solely to technical aircraft malfunctions and emergencies, we undertook analysis of a separate sample of events previously downloaded from the JACDEC database. In contrast to the sample used for our analyses this other dataset included all types of accidents and incidents over the period 1977 to 2019 and comprised 13,142 events. We then applied the same selection criteria which we had used to retrieve our sample of 2,293 events underlying the analyses presented in this paper. Accordingly, we excluded: events occurring during maneuvers on the ground; ground events not associated with takeoff or landing operations; all events in which more than a single aircraft was involved; all small aircraft (Maximum Take-off Weight < 15t); all military aircraft; events prior to the time period 2000-2019. This resulted in a sample of 4,496 events of which only 530 (11.8%) events were solely attributed to technical failures or inflight fires. Assuming a similar proportion of events were excluded from our initial sample of 3,335 events (see 2.2) would imply it excluded 394 events and so, together with the 1,076 events for which we were able to determine the mode of operation, would have resulted in a total of 1,470 events with 777 (52.9%) occurring in normal operation and 693 (47.1%) occurring in non-normal operation. After allowing for the events excluded for being solely attributed to technical failures or inflight emergencies a binomial test shows that significantly more events occur in normal than in non-normal operation ($p = .03$).

It may also be noted that we were only able to establish mode of operation for 1,076 of the 2,293 total events; we have no way of estimating the proportions of the two modes of operation for the other 1,217 events. Although unlikely, assuming that *all* 1,217 of the 1,217 events for which we were unable to determine the mode of operation occurred in non-normal operation and combining those events with the 394 events that we estimate were excluded as events solely attributed to technical failures produces, from a total of 2,687 events, 1,910 (71.1%) in non-normal operation and 777 (28.9%) in normal operation. Hence, after making allowances for potential sampling errors on our analyzed set of events, it is possible that a minority of events occur in normal operation; nevertheless, the conclusion that operation with technically airworthy aircraft and in the absence of any onboard emergency is not without risk is inescapable.

3.2.1 Mode of operation and role assignment

We were able to determine both the mode of operation and the role assignment for 837 events (see table 4). The PIC was the PF in proportionately more events in both non-normal operation (76.3%) and normal operation (69.6%) but chi-squared tests reveal there was no statistically significant association between role assignment and mode of operation for all events ($\chi^2(1, N = 837) = 2.74, p = .098$) or within any of the three types of event: Hull losses ($\chi^2(1, N = 203) = .99, p = .319$); Serious incidents ($\chi^2(1, N = 575) = 2.31, p = .128$); Incidents ($\chi^2(1, N = 59) = 2.16, p = .141$).

Table 4

Frequencies (n) and Percentages (%) of Events by Mode of Operation, Role Assignment and Event type.

	Mode of Operation											
	Normal Operation						Non-Normal Operation					
	PIC as PF		SIC as PF		TOTAL		PIC as PF		SIC as PF		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	112	81.8	25	18.2	137	100	50	75.8	16	24.2	66	100
Serious Incidents	330	66.8	164	33.2	494	100	61	75.3	20	24.7	81	100
Incidents	32	64.0	18	36.0	50	100	8	88.9	1	11.1	9	100
All events	474	69.6	207	30.4	681	100	119	76.3	37	23.7	156	100

The finding that the role-assignment effect did not vary with mode of operation could be viewed as somewhat surprising. Given the longstanding recommendation cited earlier (Jentsch et al., 1999; Sexton, 2004), that whenever a crew encounters a high workload situation it is best to have the SIC on the controls, it might be expected that the proportion of events with the SIC as PF would be higher in non-normal operation than normal operation. Of course, complex and high workload situations may also arise in normal operation (e.g. due to weather, time pressure or demanding air traffic control clearances); moreover some non-normal situations might be perceived as easy to handle. Nonetheless, given the unambiguous indicators of non-normal operation, one might expect the crew to more frequently implement a change of the role-assignment in non-normal operation when the PIC was PF.

Accordingly, in order to investigate this further, in addition to the analysis of role changes reported in section 3.1 above, which comprised all events for which we were able to determine the role assignment, we also investigated changes of role assignment by mode of operation. Within the 782 cases for which we were able to determine both the mode of operation and whether or not there was an announced change of control, we found 633 events in normal operation of which 38 (6%) involved announced role changes, 36 (94.7%) from the SIC to the PIC and only 2 (5.3%) from the PIC to the SIC. Of the 149 events in non-normal operation 17 (11.4%) involved announced role changes; of those role changes 16 were from the SIC to the PIC and only 1 from the PIC to the SIC. If pilots were compliant with the Sexton (2004) recommendation one would expect to see more announced control changes from the PIC to the SIC. In fact, we found quite the opposite: nearly all announced control changes in both modes of operation were from the SIC to the PIC. These findings indicate that pilots refrain from adopting the recommendation, perhaps due to a lack of knowledge about the recommendation. The fact that the announced control changes are nearly all in the opposite direction to that recommended, and that there are proportionately more such control changes in non-normal operation than normal operation, suggests a lack of trust by PICs in the competencies of the SIC as PF or perhaps, as especially in non-normal operation, a feeling of responsibility or perceived social expectation in PICs given their role as the accountable pilot. Despite being numerically fewer there are proportionately more announced control changes in non-normal operation than normal operation ($\chi^2(1, N = 782) = 5.39$, $p = .02$).

3.3. Teamwork Behavior

We were able to determine coding of the teamwork behavior variable for 834 of the 2,293 events (see table 5). Binomial tests for equal proportions confirmed that the flight crew's teamwork behavior was a contributory factor in a significantly larger proportion of events - 730 (87.5%) - than in events showing no issues with flight crew's teamwork behavior - 104 (12.5%), $p < 0.001$. This same difference is evident in the same comparison for each of the three event types: Hull losses (192, 88.9% vs 24, 11.1%, $p < .001$); Serious incidents (495, 86.7% vs 76, 13.3%, $p < .001$); Incidents (43, 91.5% vs 4, 8.5%, $p < .001$).

Table 5

Frequencies (n) and Percentages (%) of Events and Fatalities by Teamwork Behavior

	Teamwork Behavior											
	Events						Fatalities					
	Issue		No Issue		TOTAL		Issue		No Issue		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	192	88.9	24	11.1	216	100	5491	96.3	212	3.7	5703	100
Serious Incidents[a]	495	86.7	76	13.3	571	100	1	-	0	-	1	-
Incidents[a]	43	91.5	4	8.5	47	100	0	-	0	-	0	-
All events	730	87.5	104	12.5	834	100	5492	96.3	212	3.7	5704	100

[a] Percentages of fatalities for serious incidents and incidents are not computed due to a lack of data.

We were able to determine *both* teamwork behavior and role assignment for 779 of the 834 events (see table 6).

Table 6

Frequencies (n) and Percentages (%) of Events by Teamwork behavior, Role Assignment and Event type.

	Teamwork Behavior											
	Issue						No Issue					
	PIC as PF		SIC as PF		TOTAL		PIC as PF		SIC as PF		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	143	82.2	31	17.8	174	100	12	57.2	9	42.8	21	100
Serious Incidents	322	68.2	150	31.8	472	100	49	71.0	20	29.0	69	100
Incidents	26	66.7	13	33.3	39	100	3	75.0	1	25.0	4	100
All events	491	71.7	194	28.3	685	100	64	61.5	30	38.5	104	100

Somewhat at odds with the notion that teamwork is impaired when the PIC is PF, we found no evidence for an association between *role assignment* and *teamwork behavior* over all events ($\chi^2(1, N = 779) = .52$, $p = .47$). Nevertheless, and although an analysis split by the event types also revealed no association between *role assignment* and *teamwork behavior* for either incidents ($\chi^2(1, N = 43) = .12$, $p = .74$) or serious incidents ($\chi^2(1, N = 541) = .22$, $p = .64$), there is indeed a significant association between *role assignment* and *teamwork behavior* for the more serious 195 events involving hull losses, χ^2 $(1, N = 195) = 7.21$, $p = .007$. This association is such that the unequal preponderance of the two role assignments (more events with the PIC as PF than the SIC as PF) is more extreme in events with a teamwork issue than the events where there was no teamwork issue. This finding is consistent with the notion that teamwork is impaired when the PIC is PF.

We were also able to determine both the teamwork behavior and mode of operation for 831 events. 678 (81.6%) of these were in normal operation and for the vast majority (646, 95.3%) of the 678 events in normal operation there were issues with the flight crews' teamwork behavior. By contrast, among the 153 events happening in non-normal operation, only just over half (82, 53.6%) of the events had issues with the flight crews' teamwork behavior. This significantly greater association of teamwork behavior as a contributory factor for events in normal operation (χ^2 (1, N = 831) = 199.77, $p < .001$). Although this might be taken as evidence that flight crew teamwork is better in response to a technical malfunction or emergency, this association might also be a result of a selection effect acting on the set of events. Events in normal operation may be more likely to be tagged as showing poor teamwork because they originated as a *consequence* of poor teamwork than events in non-normal operation where a technical failure or onboard emergency would likely be the primary origin of the event.

3.4. Preventability of Events

We were able to determine the preventability for 907 events (see table 7). Binomial tests for equal proportions confirmed that, across all occurrence types, significantly more events - 797 (87.9%) - were deemed pilot-preventable than not pilot-preventable - 110 (12.1%), $p < 0.001$. This same pattern was also confirmed for each of the three event types: Hull losses ($p < 0.001$); Serious incidents ($p < 0.001$); Incidents ($p < 0.001$). The vast majority of the 5715 fatalities (5501, 96.3%) are due to preventable events rather than the events deemed not preventable (214, 3.7%).

Table 7

Frequencies (n) and Percentages (%) of Events and Fatalities by Preventability

	Preventability											
	Events						Fatalities					
	Preventable		Not Prevent.		TOTAL		Preventable		Not Prevent.		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	203	88.2	27	11.8	230	100	5500	96.3	212	3.7	5712	100
Serious Incidents[a]	538	87.5	77	12.5	615	100	1	-	2	-	3	-
Incidents[a]	56	90.3	6	9.7	62	100	0	-	0	-	0	-
All events	797	87.9	110	12.1	907	100	5501	96.3	214	3.7	5715	100

[a] Percentages of fatalities for serious incidents and incidents are not computed due to a lack of data.

We were able to determine both role assignment and preventability for 796 of the 837 events (see table 8).

Table 8

Frequencies (n) and Percentages (%) of Events by Preventability, Crew Assignment and Event type.

	Preventability											
	Pilot-Preventable						Not Pilot Preventable					
	PIC as PF		SIC as PF		TOTAL		PIC as PF		SIC as PF		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	149	83.2	30	16.8	179	100	9	47.4	10	52.6	19	100
Serious Incidents	342	68.4	158	31.6	500	100	33	68.8	15	31.2	48	100
Incidents	30	65.2	16	34.8	46	100	3	75.0	1	25.0	4	100
All events	521	71.9	204	28.1	725	100	45	63.4	26	36.6	71	100

Across all events we found no evidence of an association between role assignment and preventability ($\chi^2(1, N = 796) = 2.26, p = .132$). Nevertheless, and although an analysis split by the event types also revealed no association between role assignment and preventability for either serious incidents ($\chi^2(1, N = 548) = .002, p = .96$) or incidents ($\chi^2(1, N = 50) = .16, p = .692$), there is indeed a significant association between role assignment and preventability within hull losses ($\chi^2(1, N = 198) = 13.71, p < .001$). Many more hull loss events judged pilot-preventable occurred with the PIC as PF (149) than with the SIC as PF (30), yet there was hardly any difference in the preponderance of the two crew assignments for those events judged not preventable by the flight crew (9 with PIC as PF; 10 with the SIC as PF). This pattern is consistent with the notion that preventability is impaired when the PIC is PF to the extent that *preventable* adverse events are more frequent when the PIC is PF.

The findings for *Preventability* are very similar to those for *Teamwork Behavior*: more events, and more events of each type, are both preventable and attributable to ineffective teamwork behavior; moreover, both *Preventability* and *Teamwork behavior* are similarly associated with more egregious outcomes and with the PIC as PF role assignment. However, despite their similarities, these factors are both conceptually and empirically distinguishable; we were able to determine both *Preventability* and *Teamwork behavior* for 830 events and identified 33 (4.0%) where *Preventability* and *Teamwork behavior* were dissociated. While only 2 (2.7%) of the 74 events judged not pilot-preventable nevertheless also involved inappropriate teamwork, 31 (4.1%) of the 756 events judged pilot-preventable showed no issue with flight crew teamwork. Those events included cases where an event was judged preventable by virtue of the fact that an individual error by one pilot was realistically avoidable (e.g. by better workload- or risk management), but the other pilot had no realistic opportunity to effectively intervene (e.g. due to a lack of time). Other cases included events in which the flight crew showed good teamwork by mutually agreeing on a decision, which turned out to be consequential for the event, but which, realistically, could have been taken differently given the context of the event.

Finally, we investigated whether there was an association between preventability and mode of operation. We found 905 events for which we were able to determine both preventability and mode of operation (see table 8). As might be envisaged for events that occur in the absence of any technical failure or aircraft related emergency, almost all (703, 98.5%) of the 714 events that occurred during normal operation were deemed preventable[11]; of the 191 events that occurred in non-normal operation a far smaller proportion were deemed pilot-preventable (93, 48.7%). A chi-squared test confirmed that preventability varied significantly as a function of mode of operation ($\chi^2(1, N = 905) = 352.3$, $p < .001$).

As shown in table 9 that pattern of association - greater preventability in normal operation than non-normal operation - is apparent for all three event types. Of the 52 incidents in normal operation 51 (98.1%) were deemed pilot-preventable and only 1 (1.9%) deemed not pilot-preventable, but of the 10 incidents occurring in non-normal operation 5 were deemed preventable and 5 not preventable ($\chi^2(1, N = 62) = 22.2$, $p < .001$). Of the 507 serious incidents happening in normal operation 499 (98.4%) were judged pilot-preventable and only 8 (1.6%) as not pilot-preventable, but of the 107 serious incidents happening in non-normal operation only 39 (36.4%) were deemed pilot-preventable and 68 (63.6%) were judged as not pilot-preventable (χ^2 (1, N = 614) = 312.9, $p < .001$). For hull losses in normal operation 153 (98.7%) were pilot-preventable and only 2 (1.3%) were found to be not pilot-preventable, but of the 74 hull losses in non-normal operation 49 (66.2%) were deemed pilot-preventable and 25 (33.8%) not pilot-preventable (χ^2 (1, N = 229) = 50.9, $p < .001$).

[11] Of the 11 events deemed non-preventable in normal operation 9 involved unforeseeable interventions from wildlife, water, humans or vehicles on the ground leading to contact during landing with people, wildlife or equipment (e.g. cows, a calf, owls, a car, a person). The other 2 events involved mistakes in aircraft operating procedures and aircraft loading.

Table 9

Frequencies (n) and Percentages (%) of Events by Preventability, Mode of Operation and Event type.

	Preventability											
	Pilot-Preventable						Not Pilot-Preventable					
	Normal Operation		Non-Normal Operation		TOTAL		Normal Operation		Non-Normal Operation		TOTAL	
Event type	n	%	n	%	n	%	n	%	n	%	n	%
Hull losses	153	75.8	49	24.2	202	100	2	7.4	25	92.6	27	100
Serious Incidents	499	92.8	39	7.2	538	100	8	10.5	68	89.5	76	100
Incidents	51	91.1	5	8.9	56	100	1	16.7	5	83.3	6	100
All events	703	88.3	93	11.7	796	100	11	10.1	98	89.9	109	100

Discussion

The events collated and analyzed here provide clear evidence of a systemic safety risk in civil commercial aviation. Although accident statistics confirm that the safety of air travel has improved over the years[12] our analysis indicates that the flight crew role assignment is strongly associated with the frequency of aviation accidents and safety critical incidents. Specifically, our analysis shows that, despite the convention that captains and co-pilots usually take it in turns by flight sector to be the PF or PM such that half of flight sectors feature the PIC as PF and half the SIC as PF, almost two and a half times as many events occurred when the PIC was acting on the controls. Four times as many events involving fatalities and more than three times as many fatalities occur when the PIC rather than the SIC is PF. The clear implication is that the crew role assignment is itself a factor influencing the propensity for accidents and serious incidents.

[12] The industry has been able to consistently decrease the fatal accident rate from around 3 per million flights in the 1960s to below 1 per million flights by 1990 and further below 0.5 per million flights from the year 2000 onwards (Airbus, 2023; Boeing, 2022; International Air Transport Association (IATA), 2022).

Furthermore, our research provides clear evidence that the association between flight crew role assignment and accident/incident frequency has persisted subsequent to the reforms introduced to address it; indeed, and counter to any expectation that these reforms might, over time, gradually improve the deleterious impact of role assignment on accidents and safety critical incidents, the data show a significantly *increasing* trend in the crew role assignment effect over time between the years 2000 and 2020. The crew assignment effect is also not a localized phenomenon; it is present in each of the five continental regions we identified where civil aviation operators are headquartered.

Despite training pilots in CRM for decades, introducing a label change from PNF to PM and launching industry initiatives on monitoring the flight crew team setting with the PIC as PF and with the SIC as PM is associated with more severe accident outcomes and fatalities than the team setting with the SIC as PF and the PIC as PM, a pattern which suggests that current aviation regulation and airline policies regarding flight crew role settings are ill-fated and dysfunctional and which threatens the industry's announced goal of zero fatalities by 2030 (ICAO, 2020c). As crew role assignment is mostly under the control of aircraft operators there is a clear basis for mitigating the risk associated with this factor - a point to which we return in discussing the policy implications of our findings.

We also found that, even in the absence of any technical malfunction or emergency, the risk for harm in commercial aviation is substantial: within the sample of all the (1,076) events for which we were able to code *Mode of Operation* as normal or non-normal, more than twice as many events and more than twice as many fatalities occurred in normal operation - i.e. with technically airworthy aircraft and no emergency present - than in non-normal operation. Even after allowing for the events omitted from our sample that were exclusively attributed to technical failure as well as conceding the possibility that *all* the events for which we were unable to code the mode of operation occurred in non-normal operation, there was a substantial proportion of events occurring in normal operation. This finding suggests weaknesses in flight crew risk management when dealing with routine operational challenges, e.g. weather, time pressure or distraction, which is consistent with findings by EASA (2020, p. 28) who found that the most common underlying cause for accidents between 2015 and 2019 was flight crew's management of challenging circumstances created by technical failure or poor weather conditions. Further research on the distribution of all safety relevant events in aviation based on our proposed methodology for defining the context of events by *mode of operation* may help provide a better risk picture for the aviation industry.

Our evaluations of *Teamwork Behavior* further support this notion, with "poor CRM" or "poor teamwork" identified as a contributory factor in a large majority (95.3%) of the events occurring in normal operation while only just over half (53.6%) of the events in non-normal operation showed issues with teamwork. Although we were unable to detect an association between flight crew role assignment and the manifestation of problematic *Teamwork Behavior* across the whole set of events, there was just such a discernible relationship for the most serious events (hull losses): consistent with the notion that flight crew teamwork is impaired when the PIC is PF, failures in flight crew teamwork were significantly more prevalent for hull losses when the PIC was PF.

Our assessment of the preventability of events indicated that a large majority of the events (87.9%) – were deemed preventable by more risk-averse flying or decision-making by the flight crew. While slightly less than half of the events (48.7%) occurring in non-normal operation were preventable by the flight crew, almost all (98.5%) of the events in normal operation were determined to be preventable by the flight crew. Notably the vast majority of the fatalities (96.2%) were as a result of pilot-preventable events. As for the evidence of an association between flight crew *Role Assignment* and *Preventability* our finding was similar to that for the association between flight crew *Role Assignment* and *Teamwork Behavior:* across the whole set of events there was no association between *Role Assignment* and the

Preventability of events. Nevertheless, there was a strong relationship between these factors for the most serious events (hull losses): almost four times as many hull loss events judged pilot-preventable occurred with the PIC as PF than with the SIC as PF.

The variation in the association of the role assignment effect by event type for both *Teamwork Behavior* and *Preventability* may reflect the influence of sampling characteristics of our limited sample of events. Note that the distribution of role assignments for both the *Teamwork Behavior* and *Preventability* variables are strongly skewed - the small numbers of events where Teamwork was *not* an issue and events were *not* preventable, together with the relatively few events where the SIC was PF, reduces the sensitivity of our analyses to probe the role assignment effect by event type for both *Teamwork Behavior* and *Preventability*. Note also that the number of *incidents* is relatively small. This doubtless reflects a feature of the reporting of events - incidents that don't involve any damage, injuries or fatalities will, unlike accidents, be reported at the carriers' discretion (Stamolampros, 2022). As a consequence, incidents are a category of events that will inevitably, to some extent at least, be under-reported and under-analyzed. In the JACDEC database 32% of events are coded as incidents (as we report above, 5,690 incidents out of a total of 17,795 events). Our multiple exclusion criteria, including the focus on events between 2000-2020, reduced the number of events to 2,293 of which only 370 (16%) were incidents. The availability of crew role assignment information further significantly reduced the number of incidents we were able to include in our analyses; role assignment was determinable for 841 events (37%) of the 2,293 events but only 59 (7%) of these were incidents.

The small number of incidents raises concerns that they may be under-represented in the database; certainly, given that not every incident results in an accident, one would expect more incidents than accidents. The small numbers of incidents obviously compromise the power of our statistical analyses; moreover, there is also a risk that any process of exclusion of incidents from the dataset will also produce non-representative samples. For example, any tendency by the PIC to under-report incidents where the PIC was the PF more often than when the SIC was PF would clearly undermine our analysis. While we have no way of knowing whether or not, and to what extent if any, this happened, such an effect would help account for our findings that the associations between flight crew *Role Assignment* and *Teamwork Behavior and* flight crew *Role Assignment* and *Preventability* were only evident for hull losses.

Our research approach may serve as an example for the aviation industry on how to better study commercial aviation's safety performance by including consideration of all kinds of events instead of only (fatal) accidents. We would also advocate that records of events be improved so that they are more comprehensive and include as many details are available as standard. For example, and despite the evidence we present for its relevance, crew role assignment is not currently a standard feature of accident and incident reports. For all we know there may well be other features of events that currently go unrecorded that, if made available, would be revealed as pertinent. In the digital age there seems no reason why event reports could not link to all available data about for example all crew members detailed flying history. In epidemiological research developments in information technology and the development of digital databases have proved enlightening (Hripcsak et al., 2015; Zhang et al., 2017); there are similar opportunities in aviation pending reforms to the design of event reports and the curating of aviation event archives.

In contemplating the policy implications of our research it is worth noting that our findings on the significant influence of crew assignment are consistent with the conclusions of previous research on flight crew role assignment (Behrend & Dehais, 2020; Beveridge et al., 2018; Jentsch et al., 1999; NTSB, 1994; Orlady, 1982; Sexton, 2004) in indicating that having the PIC active on the controls is not the best option for accident prevention, may even impede effective flight crew teamwork and also suggest a particular route for reform as we outline below. Along with the crew assignment effect, other of our findings indicate that there is significant scope for improvement in aviation safety. Specifically, the findings that substantial proportions of events occur in the absence of any technical failure or aircraft related emergency, involve "poor CRM" or "poor teamwork" as a contributory factor and are judged pilot-preventable all point to the potential for reforms to mitigate these outcomes.

Given its evidently pernicious nature it is clearly important to establish what course of action could mitigate - and even entirely eliminate - the crew role assignment effect. We noted in our review of literature in the introduction that a number of different mechanisms have been proposed as underlying the effect, namely: (1) a greater cognitive workload on the PIC when the PIC is PF negatively impacting PIC decision-making; (2) differences in the relative expertise of the PIC and SIC (SICs are more likely to lose situational awareness due to a lack of experience); (3) differences in the relative status and authority of the PIC and SIC inhibiting the SIC as PM from communicating any observations or concerns. While the relative contribution of these three mechanisms remains to be established some reforms could potentially address all three possibilities. For example, one might assume that the deleterious influence of the crew role assignment effect could be tackled by relieving the PIC of the task of operating the controls by mandating the routine combination of *Command and Monitoring* assigning the PIC as PM on all flights. However careful evaluation of the impact of reforms is needed as any reform might potentially have unintended consequences; for example, if PICs never took the controls this might, over time, lead to a decline in their flying skills.

Given the repeated concerns raised about the quality of monitoring and intervention (Barshi & Bienefeld, 2018; Bienefeld & Grote, 2012; Civil Aviation Authority (CAA), 2013; Dismukes & Berman, 2010; Flight Safety Foundation (FSF), 2014, 2021; Foushee, 1984; Noort, 2020; Noort, Reader, & Gillespie, 2021a, 2019, 2021b; Perkins et al., 2022; Sumwalt & Morrison, 1997; Sumwalt et al., 2002) and its often-noted association with the crew assignment effect (Behrend & Dehais, 2020; Besco, 1995; Beveridge et al., 2018; Fischer & Orasanu, 2000; Jentsch et al., 1999; Limor & Borowsky, 2020; Milanovich, Driskell, Stout, &

Salas, 1998; Orlady, 1982; Tarnow, 2000), there is a clear case to introduce measures to boost the efficacy and status of monitoring and intervention on aviation flight decks and indeed the FAA has recently introduced further guidance on how to train monitoring and intervention by the PM more effectively (FAA, 2022). This initiative, which has advisory character only and is currently only targeted at aircraft operators in the US, may well have a positive effect, even globally, as airlines become aware of it, as, for the first time, it now clearly demands the training of every pilot, irrespective of rank or level of experience, in actively taking-over control when in the PM role and reacting appropriately when in the PF role (FAA, 2022, secs. 3–11). Currently, existing regulations, e.g. in Europe (EASA, 2022, sec. 4), require teaching of practical intervention techniques such as taking over control for training captains as PM only, with the effect that such training is still not mandated except for those few captains who have a training role.

Interestingly, neither the FAA advisory circular nor the related NASA study on monitoring (Mumaw, Billman, & Feary, 2020) explains or acknowledges the inconsistency inherent in standard operating procedures: when using the auto-pilot these procedures require that the PF, in certain critical flight phases, must "guard and follow through"[13] the flight controls in order, in the event of auto-pilot malfunctions, to be able to immediately take-over manual control; and yet, when the PF is flying manually, the same guarding of controls by the PM, in the event of piloting errors by the PF, is not required. The potential benefits stemming from such guarding by the PM, e.g. for a quicker and more effective intervention by the PM during critical flight phases such as take-off, landing or go-arounds, would warrant further research to establish the effectiveness of such guarding and following through by the PM and its influence on the teamwork between PF and PM.

---

[13] Guarding or following through the controls means to have the hands and feet loosely on the controls but not physically acting on them, however being ready to do so immediately, if required.

Moreover, these papers also do not reflect on the fact, mentioned above, that some states or operators still require that the PF be the PIC for certain complex or demanding operations despite the opposite scientific recommendation (Jentsch et al., 1999; Sexton, 2004). These requirements may even reinforce the pernicious fallacy that the safest role assignment is to have the PIC at the controls, which in turn might explain why we found that most changes of control involve PICs taking control from SICs rather than vice-versa. Such thinking might be instilled by a reliance on flight experience as the primary factor for pilot competency and a disregard of the implications for effective teamwork. However, our analysis shows that the policy for what is termed “captain-only” operation should be critically reviewed taking into account the paramount role of the PM for safe flight operation.

Although it might be hoped that enhanced training of crew teamwork - and monitoring and intervention in particular - would alleviate the crew assignment effect, the evidence presented here does not encourage the view that more assiduous crew teamwork could resolve the issue. Firstly, cognitive overload of the PIC when PF is unlikely to be reduced by more efficient teamwork and secondly, despite efforts to improve CRM training, the crew assignment effect has grown over the twenty-one-year period that we studied. Some consideration of the organization of the distribution of responsibilities by the flight crew is plainly called for.

One might well question the suitability of the hierarchical design of the current flight crew team setting which allows that only one pilot is PIC-rated, meaning that only one pilot is fully trained and assessed in leadership, judgment and decision-making. Currently the qualification and training requirements for SICs (Co-pilots) are generally lower, demanding less knowledge, less experience and respectively less training, especially in leadership, judgment and decision-making, in comparison to the PIC-role (IATA, 2020). An alternative scheme requiring both pilots in the PF and PM role to be trained to the same standard in all aspects including leadership, judgment and decision-making, and so be PIC-rated, should mitigate pernicious status hierarchy effects but also enable flight crews to regularly switch roles and at the same time ensure that the PM-role was always occupied by the PIC ensuring the required currency for all pilots in both functional roles. Both the roles (PF and PM) and the command (PIC and SIC) could be alternated on each flight leg.

A teamwork method that moves in this direction has already been designed, and even incorporated in existing aviation legislation, known as *PIC under Supervision* (PICuS) that allows the SIC to routinely perform the duties of the PIC albeit under the supervision of the PIC (EASA, 2022, para. FCL.035). This method is used during training in order to introduce co-pilots to command responsibilities and further their development to PIC status. Given that both crew members were PIC rated there would be some value in investigating the efficacy of this method for more safely distributing the leadership and decision-making tasks among the crew with both functional roles (PF/PM) being routinely alternated by every flight leg. Moreover, further evolution of the label from the passive term *Pilot Monitoring* (PM) to the more active term *Pilot Supervising* (PS) might better transmit the notion that monitoring also requires effective intervention.

Conclusion

The analyses and findings regarding the effects of flight crew role assignment presented here show clear evidence of predictable variation in aviation risks that indicate strong potential for introducing reforms to improve safety in aviation. In contrast to the aviation industry's current focus on flight crew training in teamwork, including monitoring and intervention, our analyses suggest that the underlying teamwork role settings are dysfunctional and require reform. In particular while current standard practice for crew role assignment and human factor and simulator training programs assume that the regular combination of command and control (PIC as PF) is safe and acceptable our analyses indicate that this assumption is fatally wrong.

Identifying and implementing effective reforms will of course require careful further research. Continuing calls for improved CRM training may address ineffective monitoring and detrimental status hierarchy effects (cf. FAA, 2022; Noort, Reader, & Gillespie, 2021a; Perkins et al. 2022). Nevertheless we submit that there is a clear case for considering a structural redesign of the relative status of flight crew and their roles. In the same way that prevention is better than cure, a new crew assignment design – one that eliminated status hierarchy effects and unburdened the pilot in command from the concurrent task of operating the aircraft controls - could obviate the need for an uncertain search for ways to inculcate actions to counter the deleterious behaviors prompted by the current arrangements.

Our research has also revealed that there are opportunities for the aviation industry to better learn from data gained from safety relevant events by routinely including the analysis of factors such as the role assignment and the mode of operation in investigation reports and by building a single global and freely accessible database allowing comprehensive research.

References

Airbus. (2023). A Statistical Analysis of Commercial Aviation Accidents 1958-2022. In *Airbus*. Retrieved from https://accidentstats.airbus.com

Barshi, I., & Bienefeld, N. (2018). When Silence Is Not Golden - How Could This Happen? Managing Errors in Organizations. In J. U. Hagen (Ed.), *How Could This Happen?: Managing Errors in Organizations* (pp. 45–57). https://doi.org/10.1007/978-3-319-76403-0_3

Beaty, D. (1969). *The human factor in aircraft accidents*. New York, NY: Stein and Day.

Behrend, J., & Dehais, F. (2020). How role assignment impacts decision-making in high-risk environments: Evidence from eye-tracking in aviation. *Safety Science*, *127*(March 2019), 104738. https://doi.org/10.1016/j.ssci.2020.104738

Besco, R. O. (1995). Releasing the hook on the Copilot's catch 22. *Proceedings of the Human Factors and Ergonomics Society*, *1*, 20–24. https://doi.org/10.1177/154193129503900106

Beveridge, S., Henderson, S., Martin, W., & Lamb, J. (2018). Command and Control: The Influence of Flight Crew Role Assignment on Flight Safety in Air Transport Operations. *Aviation Psychology and Applied Human Factors*, *8*(1), 1–10. https://doi.org/10.1027/2192-0923/a000130

Bienefeld, N., & Grote, G. (2012). Silence That May Kill. *Aviation Psychology and Applied Human Factors*, *2*(1), 1–10. https://doi.org/10.1027/2192-0923/a000021

Boeing. (2022). *Statistical Summary of Commercial Jet Airplane Accidets - Worldwide Operations - 1959 – 2021*. Retrieved from https://www.boeing.com/resources/boeingdotcom/company/about_bca/pdf/statsum.pdf

Boss, K. K. (2012). *Characteristics and Analysis of Major U.S. Air Carrier Accidents between 1991 and 2010* (Oklahoma State University). Retrieved from https://shareok.org/handle/11244/7178

Broome, D. (2011). Accident Reduction through Crew Resource Management. *Journal of Aviation/Aerospace Education & Research*, *20*(2). https://doi.org/10.15394/jaaer.2011.1345

Civil Aviation Authority (CAA). (2013). *Monitoring matters: Guidance on the development of pilot monitoring skills (CAA Paper 2013/02)*. Retrieved from https://publicapps.caa.co.uk/modalapplication.aspx?appid=11&mode=detail&id=5447

Cooper, G. E., White, M. D., & Lauber, J. K. (1980). Resource management on the flight deck. *NASA Conference Publication 2120*, (March), 31–58. Retrieved from https://ntrs.nasa.gov/citations/19800013796

Dismukes, R. K., & Berman, B. A. (2010). Checklists and Monitoring in the Cockpit: Why Crucial Defenses Sometimes Fail. *NASA Technical Memorandum*, (July), 1–62. Retrieved from https://ntrs.nasa.gov/citations/20110011145

Driskell, J. E., & Salas, E. (1992). Collective behavior and team performance. *Human Factors*, *34*(3), 277–288. https://doi.org/10.1177/001872089203400303

European Union Aviation Safety Agency (EASA). (2020). *Annual Safety Review 2020*. https://doi.org/10.2822/147804

European Union Aviation Safety Agency (EASA). (2022). *Easy Access Rules for Flight Crew Licensing (Part-FCL) (1178/2011-Annex-1)*. Retrieved from https://www.easa.europa.eu/en/document-library/easy-access-rules/online-publications/easy-access-rules-aircrew-regulation-eu-no?page=1&kw=FCL.510.A

European Union Aviation Safety Agency (EASA). (2023). *Easy Access Rules for Air Operations (Regulation (EU) No 965/2012)*. Retrieved from https://www.easa.europa.eu/en/document-library/easy-access-rules/online-publications/easy-access-rules-air-operations?page=1&kw=NOTECHS

Federal Aviation Administration (FAA). (2000). *Adivsory Circular 120-71 - STANDARD OPERATING PROCEDURES FOR FLIGHT DECK CREWMEMBERS*. Retrieved from https://www.icao.int/safety/fsix/library/ac 120 71.pdf

Federal Aviation Administration (FAA). (2003). *Advisory Circular No 120-71A: Standard Operating Procedures for Flight Deck Crewmembers*. Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/23216

Federal Aviation Administration (FAA). (2017). *Adivisory Circular No 120-71B: Standard Operating Procedures and Pilot Monitoring Duties for Flight Deck Crewmembers*. Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentid/1030486

Federal Aviation Administration (FAA). (2022). *Advisory Circular 120-123 - Flight Path Management*. Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/1041433

Federal Aviation Administration (FAA). (2023). *Code of Federal Regulations (CFR) 14 - Aeronautics and Space - Chapter I - FAA - Airmen - Certification: Pilots, Flight Instructor, Ground Instructors*. Retrieved from https://www.ecfr.gov/current/title-14/chapter-I/subchapter-D/part-61/subpart-G/section-61.159

Fischer, U., & Orasanu, J. (2000). Error-Challenging Strategies: Their Role in Preventing and Correcting Errors. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *44*(1), 30–33. https://doi.org/10.1177/154193120004400109

Flight Safety Foundation (FSF). (2014). *A practical guide for improving Flight Path Monitoring*. Retrieved from https://flightsafety.org/files/flightpath/EPMG.pdf

Flight Safety Foundation (FSF). (2021). *Global Action Plan for the Prevention of Runway Excursions (GAPPRE)*. 1–130. Retrieved from https://skybrary.aero/articles/global-action-plan-prevention-runway-excursions-gappre

Flight Safety Foundation (FSF). (2022). 2021 Safety Report. In *Flight Safety Foundation*. Retrieved from https://flightsafety.org/wp-content/uploads/2022/02/FSF-2021-Safety-Report_rev5.pdf

Flin, R., Martin, L., Goeters, K.-M., Hörmann, H.-J., Amalberti, R., Valot, C., & Nijhuis, H. (2003). Development of the NOTECHS (non-technical skills) system for assessing pilots' CRM skills. In *Human Factors and Aerospace Safety* (Vol. 3). https://doi.org/10.4324/9781315194035-1

Foushee, H. C. (1984). Dyads and triads at 35,000 feet: Factors affecting group process and aircrew performance. *American Psychologist*, *39*(8), 885–893. https://doi.org/10.1037/0003-066X.39.8.885

Helmreich, R. L. (2006). Red Alert - The development of CRM. *Flight Safety Australia*, (October), 24–31. Retrieved from https://www.skybrary.aero/bookshelf/flight-safety-australia-article-red-alert-sep-oct-2006

Helmreich, R. L., & Foushee, H. C. (2010). Why CRM? Empirical and Theoretical Bases of Human Factors Training. In B. G. Kanki, R. L. Helmreich, & J. B. T.-C. R. M. (Second E. Anca (Eds.), *Crew Resource Management* (Second, pp. 3–57). https://doi.org/10.1016/B978-0-12-374946-8.10001-9

Hripcsak, G., Duke, J. D., Shah, N. H., Reich, C. G., Huser, V., Schuemie, M. J., … Ryan, P. B. (2015). Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. *Studies in Health Technology and Informatics*, *216*, 574–578. https://doi.org/10.3233/978-1-61499-564-7-574

International Air Transport Association (IATA). (2020). *Command Training Edition 1 Guidance Material and Best Practices* (One). Retrieved from https://de.scribd.com/document/472001998/guidance-material-and-best-practices-for-command-training

International Air Transport Association (IATA). (2022). *2022- Safety Report*. Retrieved from https://www.iata.org/en/publications/safety-report/executive-summary/

International Civil Aviation Organization (ICAO). (1998). *Human Factors Training Manual - ICAO Doc 9683*. Retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjdkOzixKmAAxUmVPEDHYVbC2gQFnoECBIQAQ&url=https%3A%2F%2Fwww.bazl.admin.ch%2Fdam%2Fbazl%2Fde%2Fdokumente%2FFachleute%2FFlugplaetze%2FICAO%2Ficao_doc_9683_human_factors_training_manual

International Civil Aviation Organization (ICAO). (2020a). Annex 13 - Aircraft Accident and Incident Investigation. In *International Standards and Recommended Practices* (Twelfth). Retrieved from http://www.iprr.org/manuals/Annex13.html

International Civil Aviation Organization (ICAO). (2020b). *Procedures for Air Navigation Services - Training - ICAO Doc 9868* (THIRD). Retrieved from https://www.bazl.admin.ch/bazl/de/home/themen/rechtliche-grundlagen/anhaenge-icao.html

International Civil Aviation Organization (ICAO). (2020c). *Safety Report 2020 Edition*. Retrieved from https://www.icao.int/safety/pages/safety-report.aspx

International Civil Aviation Organization (ICAO). (2022). *ANNEX 6 to the Convention on International Civil Aviation - Operation of Aircraft - Part 1 - International Commercial Air Transport - Aeroplanes* (Twelve). Retrieved from https://www.bazl.admin.ch/bazl/de/home/themen/rechtliche-grundlagen/anhaenge-icao.html

Jentsch, F., Barnett, J., Bowers, C. A., & Salas, E. (1999). Who is flying this plane anyway? What mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors*, *41*(1), 1–14. https://doi.org/10.1518/001872099779577237

Khatwa, R.; Roelen, A. L. C. (1997). Controlled Flight into Terrain (CFIT) accidents of air taxi, regional and major operators. *FSF 9th European Aviation Safety Seminar*. Retrieved from https://reports.nlr.nl/items/ed91fd4e-3db9-4c7b-b333-422e6f79e258

Khatwa, R., & Helmreich, R. L. (1999). Analysis of critical factors during approach and landing in accidents and normal flight. *Proceedings of the Corporate Aviation Safety Seminar*, (December 1998), 141–259. Retrieved from https://www.skybrary.aero/bookshelf/killers-aviation

Klinect, J. (2005). Line Operations Safety Audit : A Cockpit Observation Methodology for Monitoring Commercial Airline Safety Performance (University of Texas). Retrieved from https://search.lib.utexas.edu/permalink/01UTAU_INST/be14ds/alma991010228509706011

Limor, J., & Borowsky, A. (2020). Does Specific Flight Experience Matter? The Relations Between Flight Experience of Commercial Aviation Aircrews and Missed Approach Incidents. *International Journal of Aerospace Psychology*, *30*(1–2), 38–53. https://doi.org/10.1080/24721840.2020.1715803

Merritt, A., & Klinect, J. (2006). *Defensive Flying for Pilots: An Introduction to Threat and Error Management*. Retrieved from https://www.skybrary.aero/sites/default/files/bookshelf/1982.pdf

Milanovich, D. M., Driskell, J. E., Stout, R. J., & Salas, E. (1998). Status and cockpit dynamics: A review and empirical study. *Group Dynamics*, *2*(3), 155–167. https://doi.org/10.1037/1089-2699.2.3.155

Mosier, K., Fischer, U., & Orasanu, J. (2011). *Flight Crew Decision Making : Now and NextGen*. Retrieved from https://bpb-us-w2.wpmucdn.com/sites.gatech.edu/dist/d/917/files/2018/10/Mosier_Fischer_Orasanu-Crew-Decision-Making-Now-and-in-NextGen-2011.pdf

Mumaw, R. J., Billman, D., & Feary, M. S. (2020). *Analysis of Pilot Monitoring Skills and a Review of Training Effectiveness*. (December). Retrieved from http://www.sti.nasa.gov

National Transportation Safety Board (NTSB). (1994). *Review of Flightcrew-Involved, Major Accidents of U.S. Air Carriers, 1978 - 1990*. Retrieved from https://www.ntsb.gov/safety/safety-studies/Documents/SS9401.pdf

Noort, M. C. (2020). *The behavioural nature of safety voice: advancing concepts and measures to enable the prevention of harm* (The London School of Economics and Political Science). Retrieved from http://etheses.lse.ac.uk/4205/1/Noort__Behavioural-nature-safety-voice.pdf

Noort, M. C., Reader, T. W., & Gillespie, A. (2019). Speaking up to prevent harm: A systematic review of the safety voice literature. *Safety Science*, *117*(January), 375–387. https://doi.org/10.1016/j.ssci.2019.04.039

Noort, M. C., Reader, T. W., & Gillespie, A. (2021a). Safety voice and safety listening during aviation accidents: Cockpit voice recordings reveal that speaking-up to power is not enough. *Safety Science*, *139*(March), 105260. https://doi.org/10.1016/j.ssci.2021.105260

Noort, M. C., Reader, T. W., & Gillespie, A. (2021b). The sounds of safety silence: Interventions and temporal patterns unmute unique safety voice content in speech. *Safety Science*, *140*(June 2020), 105289. https://doi.org/10.1016/j.ssci.2021.105289

NTSB. (1994). *Review of Flightcrew-Involved, Major Accidents of U.S. Air Carriers, 1978 - 1990* (pp. 1–107). pp. 1–107.

Orlady, H. W. (1982). Flight Crew Performance when PF and PNF duties are exchanged. *Proceedings of the Human Factors Society - 20th Annual Meeting*, 59–63. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=4a41fa6c895ff307b34639cd71f2cbd516207ce6

Palmer, M. T., Lack, A. M., & Lynch, J. C. (1995). Communication Conflicts of Status and Authority in Dyadic, Task-Based Interactions: Status Generalization in Airplane Cockpits. *Journal of Language and Social Psychology*, *14*, 85–101.

https://doi.org/10.1177/0261927X95141005

Paris, C. R., Cannon-Bowers, J. A., & Salas, E. (2000). Teamwork in multi-person systems: A review and analysis. *Ergonomics*, *43*(8), 1052–1075. https://doi.org/10.1080/00140130050084879

Perkins, K., Ghosh, S., Vera, J., Aragon, C., & Hyland, A. (2022). The Persistence of Safety Silence: How Flight Deck Microcultures Influence the Efficacy of Crew Resource Management. *International Journal of Aviation, Aeronautics and Aerospace, 9(3)*, 6. https://doi.org/10.15394/ijaaa.2022.1728

Salas, E., Burke, C. S., Bowers, C. A., & Wilson, K. A. (2001). Team Training in the Skies: Does Crew Resource Management (CRM) Training Work? *Human Factors*, *43*(4), 35. Retrieved from https://doi.org/10.1518/001872001775870386

Salas, E., Wilson, K. A., Burke, C. S., Wightman, D. C., & Howse, W. R. (2006). Crew Resource Management Training Research, Practice, and Lessons Learned. *Reviews of Human Factors and Ergonomics*, *2*(1), 35–73. https://doi.org/10.1177/1557234x0600200103

Scott, S. (2013). *The influence of culture on decision-making (unpublished thesis)*.

Sexton, J. B. (Ed.). (2004). *The Golden Rules of Group Interaction in High Risk Environments*. Retrieved from http://high-reliability.org/GIHRE_White_Book.pdf

Stamolampros, P. (2022). Economic pressures on airlines' safety performance. *Safety Science*, *148*(January), 105626. https://doi.org/10.1016/j.ssci.2021.105626

Stewart, S. (1992). *Flying the big jets* (Third). Shrewsbury, UK: Airlife.

Sumwalt, R., Cross, D., & Lessard, D. (2015). Examining how breakdowns in pilot monitoring of the aircraft flight pathc. *International Journal of Aviation, Aeronautics, and Aerospace*, *2*(3). https://doi.org/10.15394/ijaaa.2015.1063

Sumwalt, R. L., & Morrison, R. (1997). *What Asrs Data Tell About Inadequate Flight Crew Monitoring*. Retrieved from http://asrs.arc.nasa.gov/docs/rs/56_What_ASRS_Data_Tell_About_Inadequate.pdf

Sumwalt, R. L., Thomas, R. J., & Dismukes, K. (2002). Enhancing Flight-crew Monitoring Skills Can Increase Flight Safety. *55th International Air Safety Seminar, 4-7 November 2002, Dublin, Ireland*. Retrieved from https://www.skybrary.aero/sites/default/files/bookshelf/928.pdf

Tarnow, E. (2000). Towards the Zero Accident Goal: Assisting the First Officer: Monitor and Challenge Captain Errors. *Journal of Aviation/Aerospace Education & Research*. https://doi.org/10.15394/jaaer.2000.1269

Zhang, X., Pérez-Stable, E. J., Bourne, P. E., Peprah, E., Duru, O. K., Breen, N., … Denny, J. (2017). Big data science: Opportunities and challenges to address minority health and health disparities in the 21st century. *Ethnicity and Disease*, *27*(2), 95–106. https://doi.org/10.18865/ed.27.2.95