Absence of Pilot Monitoring Affects Scanning Behavior of Pilot Flying: Implications for the Design of Single-Pilot Cockpits

Anja K. Faulhaber, TU Braunschweig, Germany, **Maik Friedrich**, and **Tatjana Kapol**, German Aerospace Center, Braunschweig, Germany

Objective: This study examines whether the pilot flying's (PF) scanning behavior is affected by the absence of the pilot monitoring (PM) and aims at deriving implications for the design of single-pilot cockpits for commercial aviation.

Background: Due to technological progress, a crew reduction from two-crew to single-pilot operations (SPO) might be feasible. This requires a redesign of the cockpit to support the pilot adequately, especially during high workload phases such as approach and landing. In these phases, the continuous scanning of flight parameters is of particular importance.

Method: Experienced pilots flew various approach and landing scenarios with or without the support of the PM in a fixed-base Airbus A320 simulator. A within-subject design was used and eye-tracking data were collected to analyze scanning behavior.

Results: The results confirm that the absence of the PM affects the PF's scanning behavior. Participants spent significantly more time scanning secondary instruments at the expense of primary instruments when flying alone. Moreover, the frequency of transitions between the cockpit instruments and the external view increased while mean dwell durations on the external view decreased.

Conclusion: The findings suggest that the PM supports the PF to achieve efficient scanning behavior. Information should be presented differently in commercial SPO to compensate for the PM's absence and to avoid visual overload.

Application: This research will help inform the design of commercial SPO flight decks providing adequate support for the pilot particularly in terms of efficient scanning behavior.

Keywords: single pilot, reduced crew, commercial aviation, eye tracking, simulator study

Address correspondence to Anja K. Faulhaber, TU Braunschweig, Institute of Flight Guidance, Hermann-Blenk-Str. 27, 38108 Braunschweig, Germany; e-mail: a.faulhaber@tu-braunschweig.de

HUMAN FACTORS
Vol. 00, No. 0, Month XXXX, pp. 1-13
DOI:10.1177/0018720820939691
Article reuse guidelines: sagepub.com/journals-permissions

INTRODUCTION

In the early years of commercial aviation, the cockpit crew consisted of five members including two pilots, a flight engineer, a navigator, and a radio operator. The latter three roles became redundant due to technological advancements which resulted in a "progressive de-crewing" (Harris, 2007). Nowadays, two-crew operations (TCO) are common practice in commercial aviation with a captain and a first officer assuming the roles of pilot flying (PF) and pilot monitoring (PM; U.S. Department of Transportation, Federal Aviation Administration, 2015). The PF is responsible for flying the aircraft, managing the flight path, and delegating tasks to the PM (e.g., handling of flaps and landing gear). The PM monitors the flight path and aircraft systems, handles radio communication, and supports the PF by reading aloud checklists, calling out deviations from the optimal path, and executing the PF's instructions. These roles of PF and PM can be interchanged between captain and first officer throughout the flight. Additionally, the role of cockpit automation is increasingly important and pilots can hand over more and more tasks to automated systems. Hence, a further de-crewing to single-pilot operations (SPO) might become feasible and seems desirable from an economic perspective to reduce costs, improve operational flexibility, and counteract an impending pilot shortage (Bilimoria et al., 2014; Comerford et al., 2013).

Pilots can already operate an aircraft single handedly from either cockpit seat due to technical redundancy. This is a mandatory prerequisite to provide safe operations even if one pilot has to leave the cockpit or in case of pilot incapacitation. SPO is therefore theoretically feasible with the present-day flight deck but simply removing one pilot is not the solution. Several concepts of operation have already been proposed for SPO in the literature (for an overview see Neis et al., 2018; Vu et al., 2018). Most of these concepts include additional automation in the cockpit to support the single pilot, a ground-based human operator with varying functions depending on the respective concept, or a combination of both. In general, the literature suggests that any further crew reduction requires considerable changes in cockpit and automation design (e.g., Etherington et al., 2016).

Among other reasons, SPO has been proposed because workload is low during the cruise phase of flight and could easily be handled by a single pilot (Lachter et al., 2017). Workload is higher, though, during operations in terminal areas such as departure and arrival (Lee & Liu, 2003). Particularly the arrival is critical with 49% of fatal accidents occurring during final approach and landing (Boeing, 2018). Final approach refers to the last part of the landing approach when the aircraft is aligned with the centerline of the runway and descending to land. This is a particularly demanding flight phase for the pilots because they need to maneuver close to the ground. Hence, final approach and landing require a thorough investigation to identify adequate and safe solutions for SPO.

Approach and landing are commonly flown with the instrument landing system (ILS) that provides guidance both vertically via the glideslope and horizontally via the localizer. Pilots need to monitor deviations from the optimal path closely on the primary flight display (PFD) while configuring the aircraft for landing and keeping an eye on the external view. Their scanning behavior is therefore of utmost importance and is understood here as the amount of time spent on specified areas of interest (AoI) and the frequency of transitioning between them. Efficient scanning behavior has been linked to better performance previously and the role of the PM was found to be essential in the detection of deviations (Dehais et al., 2017; Lefrancois et al., 2016). But how will scanning behavior change if the PM is absent? Studying the PF's eye movements may yield valuable insights into the influence of the PM on the PF and may provide indications on how to support the pilot during SPO.

There are only few studies that report results regarding the pilot's eye movements during SPO. Most of these studies use eye tracking to evaluate new tools or systems for SPO. Examples are tools for human–autonomy teaming (Strybel, Keeler, Barakezyan et al., 2018; Strybel, Keeler, Mattoon et al., 2018), the use of augmented reality (AR) glasses in SPO (Tran et al., 2018), and synthetic vision displays to support the pilot during approach and landing (Ellis et al., 2011). The literature review indicates that no study has analyzed pilot scanning behavior during SPO in a present-day flight deck without additional support.

The Present Study

The objective of the present study is to fill the gap in the literature by studying differences in pilot scanning behavior during final approach and landing when flying SPO in contrast to TCO. Problems with SPO can be expected particularly during high workload phases, for example, due to bad weather or abnormal events such as system failures (e.g., Bailey et al., 2017; Etherington et al., 2016; Koltz et al., 2015). Thus, we investigate scanning behavior during different scenarios designed to evoke varying workload levels. The aim of the study is thereby to reveal how the absence of the PM influences the PF's scanning behavior under different circumstances and to provide implications for SPO flight deck design.

Our analysis of pilot scanning behavior focuses on three cockpit areas: The cockpit instruments are divided into *primary instruments* and *secondary instruments*. Primary instruments refer to the instruments containing essential flight information which require continuous monitoring such as altitude, speed, and course. In commercial aircraft, the PFD contains all this information. The secondary instruments include all further cockpit instruments not monitored continuously, for example, the flaps, the landing gear, and the electronic centralized aircraft monitor (ECAM). The third cockpit area is the *external view*. In the context of these areas, we developed three research questions (RQs) which will be explained in the following.

RQ1: Is scanning behavior on primary instruments influenced by the absence of the PM?

Regarding the primary instruments, research has shown that pilots spend most of the time scanning the PFD (e.g., Mumaw et al., 2001; Reynal et al., 2016; Sarter et al., 2007). In SPO with a present-day flight deck, the pilot can be expected to take over all of the PM's tasks. We test whether the scanning behavior on primary instruments is affected by this additional task load and if the effect is stronger for specific primary instruments.

RQ2: How successfully does the pilot incorporate secondary instruments into his/her scanning behavior in SPO?

Most of the additional tasks for the pilot in SPO regard the monitoring and configuring of secondary instruments. Such additional tasks require the pilot to borrow attention from other AoIs (Wickens et al., 2002), in our case from the primary instruments or from the external view. Hence, we expect that pilots spend more time scanning secondary instruments in SPO and investigate how they compensate for this additional temporal demand. Moreover, within the secondary instruments, the ECAM is of particular interest. During final approach and landing, pilots have to complete a landing checklist displayed on the ECAM. As mentioned previously, checklists are usually read aloud by the PM but in SPO, pilots have to read and check the items on their own. Hence, auditory information is reduced and pilots rely heavily upon their visual attention (Wickens et al., 2003). This visual load might be critical particularly during abnormal events when pilots have to complete an additional abnormal procedure displayed on the ECAM. Thus, we investigate how efficiently pilots scan the ECAM in terms of time spent looking at it and frequency of visits.

RQ3: Are the pilot's head-up time and transition behavior between cockpit instruments and external view affected during SPO?

A task analysis by Schutte (2017) showed that 70% of the PM's tasks during arrival are head-down on primary and secondary instruments. Especially during abnormal events, the PM almost exclusively works on the abnormal procedure drawing even more attention away from the external view. Hence, we assume that head-up time on the external view will decrease in SPO. This effect might be even stronger during abnormal scenarios. Additionally, Ellis et al. (2011) found that pilots transitioned more frequently between the cockpit instruments and the external view during SPO. Besides SPO, they also tested synthetic vision displays and it is not clear what caused the effect. Our study will show whether transition behavior between the cockpit instruments and the external view is affected by the absence of the PM. Therefore, we investigate the transition frequency and the average duration of each visit to the external view.

METHODS

Participants

Fourteen pilots participated in the study. Datasets of four participants were excluded from analysis either due to technical issues or valid eye data below 60% caused by glasses. All 10 participants were male and aged between 26 and 55 years (M = 39.9, SD = 9.1). Their ranks were captain (n = 3), first officer (n = 4), and no rank due to no airline affiliation (n = 3). All participants had experience with fly-by-wire sidestick-controlled aircraft and their flight hours ranged from 300 to 10,260 flight hours (M = 4,556, SD = 3,699). Seven participants had a type rating for the Airbus A320. Participation was voluntary and not compensated in any way.

Experiment Design

The experiment consisted of various ILS approach and landing scenarios at Frankfurt Airport runway 25 left. The task was to fly the scenarios manually either with a PM in TCO or alone in SPO. The initial configuration of the aircraft was the same for each scenario: The aircraft was in a freeze mode aligned for final approach at a distance of 8 nm from the runway, at an altitude of 2,600 ft, and with a heading of 249°. The initial airspeed was set to 180 kt and the flaps were already extended to 15° (indication 2). The spoilers were still disarmed and the landing gear retracted. Autothrust and autopilot were engaged so that the aircraft was in a stable position when participants took over to land manually.

We designed three scenarios with different workload levels for the experiment: baseline, turbulence, and abnormal. The baseline scenario was designed to simulate optimal conditions with a clear view, no clouds, and calm wind. In the turbulence scenario, the only difference was that we added moderate turbulence to increase the workload. The abnormal scenario was the same as the baseline scenario but a higher workload level was induced by an engine fire triggered at an altitude of 1,800 ft. This included an engine fire alarm followed by an abnormal procedure checklist appearing on the ECAM. The experiment was in a 2 × 3 factorial within-subject design with the factors crew configuration (TCO and SPO) and scenario (baseline, turbulence, abnormal). Hence, the participants flew each of the three scenarios twice as PF, once in TCO and once in SPO. We chose this design for better comparability between the experimental conditions. Learning effects due to repeating scenarios and effects of order were accounted for by counterbalancing the order of the scenarios. Additionally, half of the 10 participants started the experiment in the SPO crew configuration and the other half started the experiment in the TCO crew configuration.

Apparatus

The study was conducted in a fixed-base flight simulator at the Institute of Flight Guidance, TU Braunschweig. It is configured according to the Airbus A320 cockpit and used for teaching and research purposes. SMI Eye-Tracking Glasses (SensoMotoric Instruments, Germany) served to collect eye-tracking data. The eye-tracking system records gaze position, eye direction, and head position binocularly with a sampling rate of 60 Hz. Before the experiment, the eye tracker was calibrated for each participant with a one-point calibration.

Procedure

Before each session, the participants were randomly divided into teams of two. The experiment had a total runtime of approximately 2 hr per team and started with a briefing to inform the participants about the simulator configuration, the weather conditions, and their task. The participants received the required approach charts, checklists, and a Quick Reference Handbook. They were additionally instructed that they were already cleared for landing and communication with air traffic control (ATC) was only required in extraordinary situations. In such situations, the experimenter sitting in the back of the cockpit responded as ATC. Participants gave informed consent and each participant was allowed two training trials as PF. If they felt comfortable after the first training trial, the second one was skipped.

After the training trials, the eye tracker was calibrated and the experimental runs started. These runs were divided into four blocks (Table 1). The first block consisted of the three scenarios in the SPO crew configuration with the first participant randomly selected as PF. The experimenter started each scenario and instructed the participants to disengage autothrust and autopilot immediately in order to land manually. After touchdown, the simulator was reset to the initial configuration for the next run and participants completed the NASA Task Load Index (TLX; Hart & Staveland, 1988) used to assess subjective workload ratings. The second participant, who was not required in the cockpit during the SPO block, completed a demographic questionnaire in the briefing room. In block two, the second participant joined in as PM for the TCO crew configuration. In block three, the participants changed seats and the second participant then started as PF in the TCO crew configuration. The experimental runs ended after the fourth block with the second participant in the SPO crew configuration. Thus, the participants experienced each scenario three times—as PF in the SPO and in the TCO

TABLE 1: Roles of the Two Participants Throughout the Four Blocks of an Experimental Session With an Example Order of Scenarios

Block	Crew Configuration	Scenarios	First Participant	Second Participant
1	SPO	Baseline, turbulence, abnormal	PF	-
2	TCO	Turbulence, baseline, abnormal	PF	PM
3	TCO	Abnormal, turbulence, baseline	PM	PF
4	SPO	Baseline, abnormal, turbulence	-	PF

Note. SPO = single-pilot operations; TCO = two-crew operations; PF = pilot flying; PM = pilot monitoring.

crew configurations and once more in TCO as PM. The session ended with a team debriefing interview.

Data Collection and Analysis

Besides the demographic questionnaire, the NASA TLX workload ratings, and the debriefing interviews, we collected simulator parameters, eye-tracking data, and video recordings. This paper focuses on the analysis of the eye-tracking data to assess differences in pilot scanning behavior. Results concerning workload and pilot performance have already been reported previously (Faulhaber & Friedrich, 2019; Faulhaber, 2019). Eye-tracking data and workload ratings were only collected for the PF who was always sitting in the left seat for better comparability. Only eye-tracking data obtained between an altitude of 2,500 ft and touchdown were analyzed when autopilot and autothrust were disengaged. We used the software Eye-Tracking Analyser by DLR (Friedrich et al., 2017) to analyze eye-tracking data. A velocity-based detection algorithm was applied to identify saccades and fixations (Salvucci & Goldberg, 2000). We defined 19 AoIs to analyze the eye-tracking data (Figure 1). Only the valid data (M = 85.76%, SD = 18.6) collected within these AoIs were used. Due to the small sample size and because assumptions for parametric tests were not clearly met, we conducted the nonparametric Wilcoxon signed-rank test for significance testing with an alpha level of 0.05. Moreover, we calculated effect sizes r as described by Field (2009).

RESULTS

Following our RQs, the AoIs were classified into primary instruments, secondary instruments, and external view as described in Figure 1. The results showed that the scanning patterns were, in general, more dispersed across the secondary instruments in the SPO crew configuration. Averaged across all three scenarios, participants looked at a higher number of AoIs during SPO (M = 12.43, SD = .82) as opposed to TCO (M = 9.57, SD = .6). This effect was statistically significant ($z = -2.81$, $p = .002$, $r = .63$). Areas looked at during SPO but not TCO were exclusively located within the secondary instruments, such as the radio management panel and the gear and flap levers. These are instruments which are commonly operated by the PM in TCO. In the following, further results will be reported for primary instruments, secondary instruments, and external view separately.

Scanning Behavior on Primary Instruments

With RQ1, we wanted to test whether the pilot's scanning behavior on primary instruments would be influenced by the absence of the PM. Hence, we analyzed the dwell time percentages on primary instruments whereby a dwell is defined as one visit to an AoI from the moment of entering until the exit (Holmqvist et al., 2015). Dwell time percentage is consequently understood as the sum of dwell durations on a respective AoI relative to the total dwell time on all AoIs.

Figure 1. Cockpit of the A320 fixed-base flight simulator with the AoIs highlighted. *Primary instruments*: (1) attitude indicator, (2) airspeed indicator, (3) altimeter (including variometer), (4) course indicator, (5) mode indicator; *Secondary instruments*: (6) navigation display (ND), (7) ECAM, (8) multifunction display (MFD), (9) gear lever, (10) multifunction control display unit (MCDU), (11) radio management panel, (12) thrust levers, (13) flap lever and speed brakes, (14) attention getter panel, (15) flight control unit (FCU), (16) overhead panel; *External view*: (17) left window, (18) front window, and (19) right window. AoI = area of interest; ECAM = electronic centralized aircraft monitor.

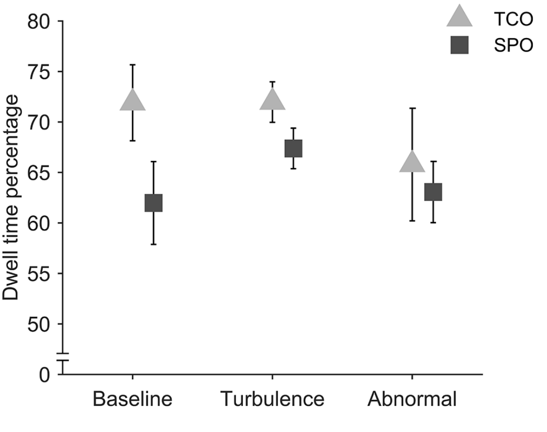

Figure 2. Mean dwell time percentages on primary instruments for the three scenarios in TCO and SPO crew configurations. Error bars show the standard error of the mean. TCO = two-crew operations; SPO = single-pilot operations.

Figure 2 shows the results for the mean dwell time percentages on primary instruments which revealed that participants spent on average less time scanning primary instruments in the SPO baseline scenario (M = 61.98%, SD = 12.97) as compared to the TCO baseline scenario (M = 71.9%, SD = 11.9). The Wilcoxon signed-rank test showed that this effect was statistically significant (z = 2.19, p = .03, r = .49). In the turbulence scenario, a similar trend could be observed with a higher mean dwell time percentage during TCO (M = 71.97%, SD = 6.34) than during SPO (M = 67.38%, SD = 6.35) but the effect did not reach significance (z = 1.78, p = .08, r = .4). Also in the abnormal scenario, participants spent on average slightly less time looking at primary instruments during SPO than during TCO but the effect was not significant (z = 0.87, p = .43, r = .19).

TABLE 2: Mean (*SD*) Dwell Time Percentages for Primary Instrument AoIs

	Baseline		Turbulence		Abnormal	
AoI	TCO	SPO	TCO	SPO	TCO	SPO
Attitude indicator	42.95 (13.06)	35.09 (8.15)	41.39 (8.05)	40.79 (7.92)	37.29 (14.03)	33.25 (11.43)
Airspeed indicator	15.14 (3.42)	14.27 (4.25)	17.12 (4.92)	15 (2.79)	17.13 (4.52)	16.46 (5.35)
Altimeter	10.82 (5.28)	9.93 (6.38)	12.08 (4.71)	10.68 (5.45)	10.18 (7.17)	11.15 (6.85)
Course indicator	2.99 (3.63)	2.55 (3.05)	1.38 (1.64)	0.89 (0.62)	1.16 (1.67)	2.2 (3.16)

Note. TCO = two-crew operations; SPO = single-pilot operations; AoI = area of interest.

To identify whether the effect was stronger for specific primary instruments, we additionally analyzed the mean dwell time percentages for the separate AoIs. Results showed that participants looked on average less at each one of them during SPO throughout all scenarios except for the altimeter and course indicator in the abnormal scenario (Table 2). The mode indicator was excluded here as it reached only marginal and zero values.

Scanning Behavior on Secondary Instruments

In the context of RQ2, we wanted to investigate how successfully pilots incorporate the secondary instruments into their scanning behavior during SPO. We expected that pilots would spend considerably more time scanning secondary instruments during SPO and our results support this assumption (Figure 3). In the SPO baseline scenario, participants took nearly twice the amount of time looking at secondary instruments (M = 14.16%, SD = 4.94) as opposed to the TCO baseline scenario (M = 7.29%, SD = 4.07). The Wilcoxon signed-rank test indicated that this effect was statistically significant (z = −2.7, p = .004, r = .6). Moreover, in the turbulence scenario, the mean dwell time percentage on secondary instruments was higher during SPO (M = 11.39%, SD = 5.71) than during TCO (M = 7.93%, SD = 4.07) and the effect was statistically significant (z = −2.09, p = .04, r = .47). The results were also significant in the abnormal scenario with means of 16.98% (SD = 10.09) for SPO and 8.21% (SD = 5.26) for TCO (z = −2.8, p = .002, r = .63).

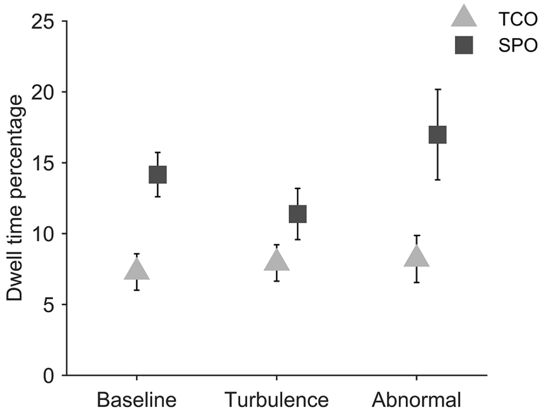

Figure 3. Mean dwell time percentages on secondary instruments for the three scenarios in TCO and SPO crew configurations. Error bars show the standard error of the mean. TCO = two-crew operations; SPO = single-pilot operations.

Furthermore, within the secondary instruments, we were particularly interested in the scanning of the ECAM as it displays the landing checklist and abnormal procedure. To evaluate the efficiency of the pilot's scanning behavior, we tested how much time they spent looking at the ECAM and how frequently they dwelled there. The results for the mean dwell time percentages on the ECAM (Figure 4, left) showed that in the abnormal scenario, participants spent significantly more time scanning the ECAM during SPO than during TCO (z = −2.29, p = .02, r = .51). In the other scenarios, the mean dwell time percentages on the ECAM also trended slightly higher during SPO as opposed to TCO

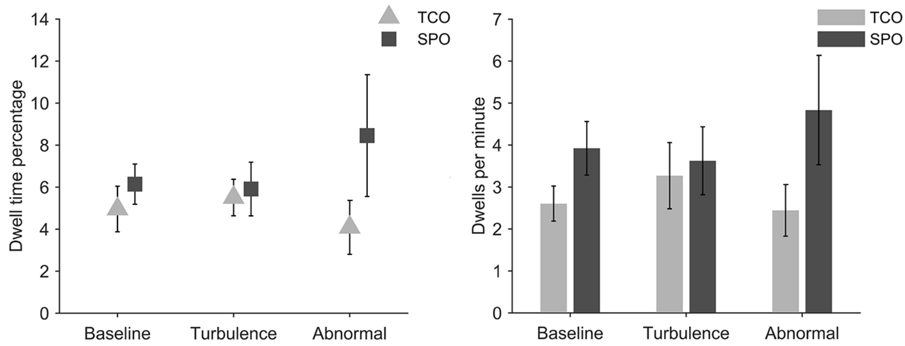

Figure 4. Dwell data for the ECAM for the three scenarios in TCO and SPO crew configurations: mean dwell time percentages on the left and mean number of dwells per minute on the right. Error bars show the standard error of the mean. ECAM = electronic centralized aircraft monitor; TCO = two-crew operations; SPO = single-pilot operations.

but the results were not significant either in the baseline scenario ($z = -0.82$, $p = .45$, $r = .18$) or in the turbulence scenario ($z = -0.15$, $p = .92$, $r = .03$).

Regarding the dwell frequencies on the ECAM (Figure 4, right), the results revealed that the mean number of dwells per minute was significantly higher during SPO than TCO in the baseline scenario ($z = -2.09$, $p = .04$, $r = .47$) as well as in the abnormal scenario ($z = -2.5$, $p = .01$, $r = .56$). Similarly, in the turbulence scenario, the mean dwell frequency trended higher in the SPO crew configuration but not significantly ($z = -0.66$, $p = .56$, $r = .15$). Hence, these results suggest that the increase in the dwell time percentage on the ECAM during SPO was not due to longer dwells but to more frequent dwells.

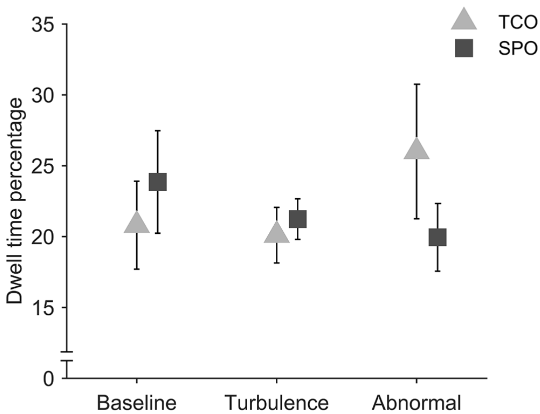

Figure 5. Mean dwell time percentages on the external view for the three scenarios in TCO and SPO crew configurations. Error bars show the standard error of the mean. TCO = two-crew operations; SPO = single-pilot operations.

Scanning Behavior on the External View

Regarding the scanning behavior on the external view, the first part of RQ3 aimed at investigating if the head-up times in terms of dwell time percentages on the external view decreased during SPO. Contrary to our expectations, the results showed that mean dwell time percentages on the external view were slightly higher in SPO baseline and turbulence scenarios compared to the respective scenarios in TCO (Figure 5). In the abnormal scenario, however, participants looked on average less outside during SPO than during TCO. None of these effects reached statistical significance.

For the second part of RQ3, we investigated the transition behavior between the cockpit instruments and the external view in terms of transitions per minute and mean dwell durations on the external view. The results for the mean number of transitions per minute (Figure 6, left) showed that transitions were significantly more frequent during SPO in the baseline scenario ($z = -2.8$, $p = .002$, $r = .63$) and in the turbulence

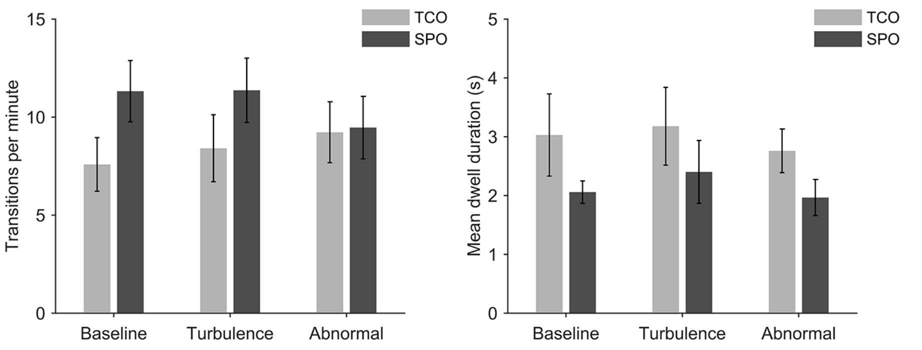

Figure 6. Mean number of transitions per minute between the cockpit instruments and the external view on the left and mean dwell durations on the external view in seconds on the right. Error bars show the standard error of the mean.

scenario ($z = -2.4$, $p = .01$, $r = .54$). The mean transition frequency also trended slightly higher in the SPO abnormal scenario compared to the scenario in TCO but not significantly ($z = -0.56$, $p = .63$, $r = .13$).

Regarding the mean dwell durations on the external view (Figure 6, right), the results indicated that dwells were significantly shorter during SPO than during TCO in the baseline scenario ($z = 2.5$, $p = .01$, $r = .56$) as well as in the abnormal scenario ($z = 2.5$, $p = .01$, $r = .56$). Also in the turbulence scenario, the mean dwell duration was shorter during SPO as compared to TCO but the effect did not reach significance ($z = 1.58$, $p = .13$, $r = .35$).

DISCUSSION

The objective of this study was to investigate if and how the scanning behavior of the PF would be affected by the absence of the PM and what this might entail for potential future SPO. The results showed that scanning patterns were, in general, more dispersed across a higher number of secondary instrument AoIs in the SPO crew configuration. Moreover, participants had to spend more time on secondary instrument AoIs such as the ECAM to extract relevant information during SPO, particularly in the abnormal scenario. They compensated for this additional temporal demand by looking less at primary instruments. Participants also transitioned more frequently between the cockpit instruments and the external view during SPO and dwells on the external view were significantly shorter in duration. Additionally, head-up times decreased significantly in the SPO abnormal scenario.

The interpretation and further discussion of these results should take the limitations of the study into account. First of all, the number of participants was relatively small; a study with more participants might provide more robust results. Second, some of the pilots complained about the eye-tracking glasses because they felt that the glasses constricted their visual field. In the abnormal scenario, there was no consensus on how to handle the engine fire and not all of the participants performed the recommended procedure as displayed on the ECAM. Finally, due to our experiment design, participants experienced the same scenarios several times. We chose this design for better comparability but of course this may lead to learning effects in general and to reduced surprise effects in the abnormal scenario. We tried to keep these to a minimum via counterbalancing but the repetitive design might still have affected effect size and significance of the results. Hence, the effects observed should be even more pronounced in a real-life setting. Nevertheless, the results provide indications regarding differences in scanning behavior during SPO as opposed to TCO and implications for the design of commercial SPO cockpits.

In general, the results showed that pilots' scanning patterns were more dispersed across the cockpit instruments during SPO. Areas looked at during SPO but not TCO were exclusively secondary instruments such as gear and flap levers which are operated by the PM in TCO. The single pilot in SPO had to take over these tasks but was not trained to do so. Further training adjusted to SPO could help the pilot incorporate these tasks more efficiently and improve scanning behavior. Nowadays, pilot training focuses to a great extent on teamwork by means of crew resource management. This might still be relevant in SPO, especially in the context of human–autonomy teaming (Mosier & Fischer, 2014; Shively, Lachter et al., 2018). Nevertheless, new procedures are required for SPO so that pilots can be trained to complete checklists and abnormal procedures in an efficient and safe manner on their own. Moreover, our results motivate changes in automation and flight deck design that support efficient scanning behavior during SPO, as will be discussed in the following.

Implications for the Design of SPO Flight Decks

Due to the additional task load, participants spent significantly more time looking at secondary instruments during SPO and consequently less time on primary instruments with the essential flight information. Several pilots suggested in the debriefing interviews that certain tasks related to the secondary instruments could be automated, for example, flaps, spoilers, and landing gear. However, the role of automation in SPO is controversial and some participants were strongly opposed to higher degrees of automation. Pilots do not want to become supervisors of cockpit automation which would not be adequate task allocation either (Schutte, 2015). If implemented correctly, for example, by taking research on human–autonomy teaming for SPO into account (e.g., Shively, Brandt et al., 2016; Shively, Lachter et al., 2018; Shively, Brandt et al., 2016, Cover et al., 2018; Strybel, Keeler, Barakezyan et al., 2018; Strybel, Keeler, Mattoon et al., 2018), automating secondary tasks could allow the pilots more time and less distractions to focus on the essential information.

Moreover, our findings showed that participants transitioned more frequently between the cockpit instruments and the external view and dwells on the external view were shorter during SPO. We assumed that without the PM pilots could not detect deviations from the optimal flight path as displayed on the PFD while looking at the external view. In order not to miss anything, pilots had to adjust their transition behavior and reduce the dwell durations on the external view. From these results we concluded that SPO flight decks need to make essential instrument information available even when the pilot is looking at the external view. Which kind of information is essential depends of course on the requirements of the respective flight phase. During approach and landing, for example, crucial information concerns deviations from glideslope and localizer as well as speed. A redesign of the flight deck could allow pilots to extract such information via their peripheral vision (Johnston et al., 2015; Nikolic & Sarter, 2001) thereby reducing the number of transitions required. Head-up displays (HUDs) are a possible solution in this context as they support more efficient transition behavior between essential flight parameters and the external view (Ellis et al., 2011; Martin-Emerson & Wickens, 1997).

Within the secondary instruments, our results showed that pilots spent more time scanning the ECAM and dwelled there more frequently during SPO. We concluded that the presentation of checklists is not optimal for SPO because pilots have to dwell on the ECAM frequently to extract the relevant information. Checklists should instead be presented in a way that is easier and faster to process for a single pilot. In this context, Arango Pérez and Behrend (2019) suggested the interactive use of AR glasses which prompt the pilot toward checklist items. Moreover, AR glasses are of particular interest in extraordinary situations to support the efficient completion of abnormal procedures (Tran et al., 2018). In conclusion, our results indicate that pilots had to dwell more frequently on the ECAM which could be reduced by superimposing information via AR glasses to improve scanning behavior. Potential negative consequences related to the use of both HUDs and AR glasses, such as cognitive and attentional tunneling (Crawford & Neal, 2006; Wickens & Alexander, 2009), need to be taken into account in the development of adequate tools.

Furthermore, the fact that pilots dwelled more frequently on the ECAM and transitioned more often between the cockpit instruments and the external view during SPO can be interpreted as a form of visual overload. Due to the absence of the PM, there is no auditory feedback regarding deviations from the flight path, checklists, and abnormal procedures. These results are in line with Wickens et al. (2003) who stated that pilots rely more on their visual attention during SPO due to the reduction of oral communication. Moreover, this is also supported by our findings regarding workload. Our results from the NASA TLX showed that the pilots' workload was not generally increased during SPO but particularly the temporal demand increased significantly. Also the eye-tracking data showed faster eye movements reflecting a certain time pressure and stress during SPO (for details see Faulhaber & Friedrich, 2019; Faulhaber, 2019). Thus, the higher temporal demand seems to be connected to the visual overload which is an issue of critical importance for SPO (Wickens et al., 2002).

In the context of visual overload, the need for multisensory interfaces emerged (Sarter, 2000). The use of tactile feedback has already shown promising results (e.g., Sklar & Sarter, 1999). Regarding the auditory-oral modality, SPO cockpit systems should possess the capability for natural language processing (Bailey et al., 2017; Cummings et al., 2016). This allows the pilots to communicate with cockpit systems the same way as they are communicating with the PM nowadays. They could receive information via natural language communication but also control systems via voice commands similarly as giving instructions to the PM, for example, to extend the landing gear or set the flaps. Our results showed that these tasks put an additional visual load on the pilot as these were areas that pilots did not even need to look at during TCO. Visual load could hence be reduced by allowing natural language communication with the cockpit systems. In summary, we concluded from our results that we need to avoid visual overload during SPO, for example, by outsourcing to other sensory modalities.

Future Work and Conclusions

Our results showed that participants spent significantly more time scanning secondary instruments and transitioned more frequently between cockpit instruments and the external view during SPO. We discussed HUDs to reduce transition frequencies but further research is required regarding HUDs in the context of commercial SPO taking into account previously discovered issues (for an overview see Crawford & Neal, 2006; Fadden et al., 1998). Moreover, we discussed the use of AR glasses for more efficient scanning behavior during SPO. AR glasses could even replace HUDs while providing further benefits (Moehle & Clauss, 2015). Future work should tackle major challenges regarding the wearing comfort of AR glasses, technical reliability, specifics of implementation in SPO, and potential negative consequences such as attentional tunneling. Additionally, we showed that visual overload has to be avoided in SPO and suggested the use of multisensory interfaces. Further research should focus on how to effectively implement multisensory interfaces in SPO and which sensory modality to use for specific types of information. Our results thereby provided indications for the design of commercial SPO flight decks supporting efficient scanning behavior and revealed several topics for future work.

ACKNOWLEDGMENTS

This research was conducted within the graduate program "Gendered configurations of humans and machines. Interdisciplinary analyses of technology" (KoMMa.G) funded by the federal state of Lower Saxony, Germany. The authors would like to thank Peter Hecker, Thomas Feuerle, and Stefan Seydel for their support throughout the project.

KEY POINTS

- When the PM was absent, pilots spent significantly more time scanning secondary instruments which they compensated for by looking less at primary instruments.

- Within the secondary instruments, pilots dwelled more frequently on the ECAM to extract the relevant information when flying alone.
- Pilots transitioned more frequently between the cockpit instruments and the external view and dwells on the external view were shorter during SPO.
- These findings are interpreted in terms of visual overload. HUDs, AR glasses, and multisensory interfaces are discussed as potential solutions to support efficient scanning behavior in SPO.

ORCID iDs

Anja K. Faulhaber https://orcid.org/0000-0003-1819-2373
Maik Friedrich https://orcid.org/0000-0003-3742-2322

REFERENCES

Arango Pérez, A., & Behrend, F. (2019). *A holographic checklist assistant for the single pilot* [Conference session]. 2019 Integrated Communications, Navigation and Surveillance Conference (ICNS), Herndon, VA.

Bailey, R. E., Kramer, L. J., Kennedy, K. D., Stephens, C. L., & Etherington, T. J. (2017). *An assessment of reduced crew and single pilot operations in commercial transport aircraft operation* [Conference session]. IEEE/AIAA 36th Digital Avionics Systems Conference (DASC), St. Petersburg, FL.

Bilimoria, K. D., Johnson, W. W., & Schutte, P. C. (2014). *Conceptual framework for single pilot operations* [Conference session]. Proceedings of the International Conference on Human-Computer Interaction in Aerospace (HCI-Aero '14), Santa Clara, CA (Article 4). ACM

Boeing. (2018). *Statistical summary of commercial jet airplane accidents: Worldwide operations 1959-2017*. Boeing. www.boeing.com/news/techissues/pdf/statsum.pdf.

Comerford, D., Brandt, S. L., Lachter, J., Wu, S.–C., Mogford, R., Battiste, V., & Johnson, W. W. (2013). *NASA's single-pilot operations technical interchange meeting: Proceedings and findings*. NASA/CP-2013-216513. NASA.

Cover, M., Reichlen, C., Matessa, M., & Schnell, T. (2018). Analysis of airline pilots subjective feedback to human autonomy teaming in a reduced crew environment. In S. Yamamoto & H. Mori (Eds.), *Human interface and the management of information: Information in applications and services. HIMI 2018. Lecture notes in computer science* (Vol. 10905, pp. 359–368). Springer.

Crawford, J., & Neal, A. (2006). A review of the perceptual and cognitive issues associated with the use of head-up displays in commercial aviation. *The International Journal of Aviation Psychology*, *16*, 1–19. https://doi.org/10.1207/s15327108ijap1601_1

Cummings, M. L., Stimpson, A. J., & Clamann, M. (2016). *Functional requirements for onboard intelligent automation in single pilot operations* [Conference session]. AIAA Aerospace Conference, San Diego, CA.

Dehais, F., Behrend, J., Peysakhovich, V., Causse, M., & Wickens, C. D. (2017). Pilot flying and pilot monitoring's aircraft state awareness during go-around execution in aviation: A behavioral and eye tracking study. *The International Journal of Aerospace Psychology*, *27*, 15–28. https://doi.org/10.1080/10508414.2017.1366269

Ellis, K. K. E., Kramer, L. J., Shelton, K. J., Arthur, J. J., & Prinzel, L. J. (2011). Transition of attention in terminal area NextGen operations using synthetic vision systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *55*, 46–50. https://doi.org/10.1177/1071181311551010

Etherington, T. J., Kramer, L. J., Bailey, R. E., Kennedy, K. D., & Stephens, C. L. (2016). *Quantifying pilot contribution to flight safety for normal and non-normal airline operations* [Conference session]. IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), Sacramento, CA.

Fadden, S., Ververs, P. M., & Wickens, C. D. (1998). Costs and benefits of head-up display use: A meta-analytic approach. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *42*, 16–20. https://doi.org/10.1177/154193129804200105

Faulhaber, A. K. (2019). *From crewed to single-pilot operations: Pilot performance and workload management* [Symposium]. Proceeding of the 20th International Symposium on Aviation Psychology, Dayton, OH.

Faulhaber, A. K., & Friedrich, M. (2019). Eye-tracking metrics as an indicator of workload in commercial single-pilot operations. In L. Longo & C. M. Leva (Eds.), *Human mental workload: Models and applications. H-WORKLOAD 2019. communications in computer and information science* (Vol. 1107, pp. 213–225). Springer.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Sage.

Friedrich, M., Rußwinkel, N., & Möhlenbrink, C. (2017). A guideline for integrating dynamic areas of interests in existing set-up for capturing eye movement: Looking at moving aircraft. *Behavior Research Methods*, *49*, 822–834. https://doi.org/10.3758/s13428-016-0745-x

Harris, D. (2007). A human-centred design agenda for the development of single crew operated commercial aircraft. *Aircraft Engineering and Aerospace Technology*, *79*, 518–526. https://doi.org/10.1108/00022660710780650

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload. Advances in Psychology* (Vol. 52, pp. 139–183). North Holland Press.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2015). *Eye tracking: A comprehensive guide to methods and measures*. Oxford.

Johnston, J. C., Ruthruff, E., & Lien, M.-C. (2015). Visual information processing from multiple displays. *Human Factors*, *57*, 276–297. https://doi.org/10.1177/0018720814545974

Koltz, M. T., Roberts, Z. S., Sweet, J., Battiste, H., Cunningham, J., Battiste, V., Vu, K.-P. L., & Strybel, T. Z. (2015). An investigation of the harbor pilot concept for single pilot operations. *Procedia Manufacturing*, *3*, 2937–2944. https://doi.org/10.1016/j.promfg.2015.07.948

Lachter, J., Brandt, S. L., Battiste, V., Matessa, M., & Johnson, W. W. (2017). Enhanced ground support: Lessons from work on reduced crew operations. *Cognition, Technology & Work*, *19*, 279–288. https://doi.org/10.1007/s10111-017-0422-6

Lee, Y.-H., & Liu, B.-S. (2003). Inflight workload assessment: Comparison of subjective and physiological measurements. *Aviation, Space, and Environmental Medicine*, *74*, 1078–1084.

Lefrancois, O., Matton, N., Gourinat, Y., Peysakhovich, V., & Causse, M. (2016). *The role of pilots' monitoring strategies in flight performance* [Conference session]. Proceedings of the 32nd Conference of the European Association for Aviation Psychology (EAAP32), Cascais, Portugal.

Martin-Emerson, R., & Wickens, C. D. (1997). Superimposition, symbology, visual attention, and the head-up display. *Human Factors*, *39*, 581–601. https://doi.org/10.1518/001872097778667933

Moehle, R., & Clauss, J. (2015). Wearable technologies as a path to single-pilot part 121 operations. *SAE International Journal of Aerospace*, *8*, 81–88. https://doi.org/10.4271/2015-01-2440

Mosier, K. L., & Fischer, U. M. (2014). *CRM principles and practices for SPO*. Final Report. San Francisco State University & Georgia Institute of Technology.

Mumaw, R. J., Sarter, N. B., & Wickens, C. D. (2001). *Analysis of pilots' monitoring and performance on an automated flight deck* [Symposium]. Proceeding of the 11th International Symposium on Aviation Psychology, Columbus, OH.

Neis, S. M., Klingauf, U., & Schiefele, J. (2018). *Classification and review of conceptual frameworks for commercial single pilot operations* [Conference session]. IEEE/AIAA 37th Digital Avionics Systems Conference (DASC), London.

Nikolic, M. I., & Sarter, N. B. (2001). Peripheral visual feedback: A powerful means of supporting effective attention allocation in event-driven, data-rich environments. *Human Factors*, *43*, 30–38. https://doi.org/10.1518/001872001775992525

Reynal, M., Colineaux, Y., Vernay, A., & Dehais, F. (2016). *Pilot flying vs. pilot monitoring during the approach phase: An eye-tracking study* [Conference session]. Proceedings of the International Conference on Human-Computer Interaction in Aerospace (HCI-Aero '16), Paris, France (Article 7). ACM

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. Proceedings of the 2000 Symposium on Eye Tracking Research and Applications (ETRA '00), Palm Beach Gardens, FL (pp. 71-78). ACM.

Sarter, N. B. (2000). The need for multisensory interfaces in support of effective attention allocation in highly dynamic event-driven domains: The case of cockpit automation. *The International Journal of Aviation Psychology*, *10*, 231–245. https://doi.org/10.1207/S15327108IJAP1003_02

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors*, *49*, 347–357. https://doi.org/10.1518/001872007X196685

Schutte, P. C. (2015). How to make the most of your human: Design considerations for single pilot operations. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics. EPCE 2015. Lecture notes in computer science* (Vol. 9174, pp. 480–491). Springer.

Schutte, P. C. (2017). *Task analysis of two crew operations in the flight deck: Investigating the feasibility of using single pilot* [Symposium]. Proceeding of the 19th International Symposium on Aviation Psychology, Dayton, OH.

Shively, R. J., Brandt, S. L., Lachter, J., Matessa, M., Sadler, G., & Battiste, H. (2016). Application of human-autonomy teaming (HAT) patterns to reduced crew operations (RCO). In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics. EPCE 2016. Lecture notes in computer science* (Vol. 9736, pp. 244–255). Springer.

Shively, R. J., Lachter, J., Koteskey, R., & Brandt, S. L. (2018). Crew resource management for automated teammates (CRM-A). In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics. EPCE 2018. Lecture notes in computer science* (Vol. 10906, pp. 215–229). Springer.

Sklar, A. E., & Sarter, N. B. (1999). Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. *Human Factors*, *41*, 543–552. https://doi.org/10.1518/001872099779656716

Strybel, T. Z., Keeler, J., Barakezyan, V., Alvarez, A., Mattoon, N., Vu, K.–P., & Battiste, V. (2018). Effectiveness of human autonomy teaming in cockpit applications. In S. Yamamoto & H. Mori (Eds.), *Human Interface and the Management of Information: Information in Applications and Services. HIMI 2018. Lecture Notes in Computer Science* (Vol. 10905, pp. 465–476). Springer.

Strybel, T. Z., Keeler, J., Mattoon, N., Alvarez, A., Barakezyan, V., Barraza, E., & Battiste, V. (2018). Measuring the effectiveness of human autonomy teaming. In C. Baldwin (Ed.), *Advances in Neuroergonomics and cognitive engineering. AHFE 2017. advances in intelligent systems and computing* (Vol. 586, pp. 23–33). Springer.

Tran, T. H., Behrend, F., Fünning, N., & Arango Pérez, A. (2018). *Single pilot operations with AR-glasses using Microsoft HoloLens* [Conference session]. IEEE/AIAA 37th Digital Avionics Systems Conference (DASC), London.

U.S. Department of Transportation, Federal Aviation Administration. (2015). *Roles and responsibilities for pilot flying (PF) and pilot monitoring (PM): Safety alert for operators 15011*. Flight Standards Service. https://www.faa.gov/other_visit/aviation_industry/airline_operators/airline_safety/safo/all_safos/media/2015/safo15011.pdf.

Vu, K. -P. L., Lachter, J., Battiste, V., & Strybel, T. Z. (2018). Single pilot operations in domestic commercial aviation. *Human Factors*, *60*, 755–762. https://doi.org/10.1177/0018720818791372

Wickens, C. D., & Alexander, A. L. (2009). Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology*, *19*, 182–199. https://doi.org/10.1080/10508410902766549

Wickens, C. D., Goh, J., Helleberg, J. R., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human Factors*, *45*, 360–380. https://doi.org/10.1518/hfes.45.3.360.27250

Wickens, C. D., Helleberg, J. R., & Xu, X. (2002). Pilot maneuver choice and workload in free flight. *Human Factors*, *44*, 171–188. https://doi.org/10.1518/0018720024497943

Anja K. Faulhaber is a PhD student at the Institute of Flight Guidance, TU Braunschweig. She received her master's degree in cognitive science from Osnabrueck University in 2017. Her research interests include human–machine interaction and human factors in aviation.

Maik Friedrich works at the German Aerospace Center in Braunschweig. He received a PhD from the Institute of Psychology, TU Chemnitz in 2019. His research focuses on eye-tracking analyses at an individual level in relation to task performance.

Tatjana Kapol works as a simulation specialist at the German Aerospace Center in Braunschweig since 1997. She is responsible for eye-tracking measurements in the cockpit simulator and was involved in several validation campaigns in the areas of cockpit research and airport management research.

Date received: June 12, 2020
Date accepted: June 9, 2020