The International Journal of Aerospace Psychology

ISSN: 2472-1840 (Print) 2472-1832 (Online) Journal homepage: http://www.tandfonline.com/loi/hiap21

Stirring the Pot: Comparing Stick Input Patterns and Flight-Path Control Strategies in Airline Pilots

Andreas Haslbeck, Hans-Juergen Hoermann & Patrick Gontar

THE INTERNATIONAL JOURNAL OF AEROSPACE PSYCHOLOGY
https://doi.org/10.1080/24721840.2018.1481343

Stirring the Pot: Comparing Stick Input Patterns and Flight-Path Control Strategies in Airline Pilots

Andreas Haslbeck [a], Hans-Juergen Hoermann [b], and Patrick Gontar [a]

[a]Institute of Ergonomics, Technical University of Munich, Munich, Germany; [b]Institute of Aerospace Medicine, German Aerospace Center, Hamburg, Germany

ABSTRACT

Objective: This study evaluated airline pilots' inceptor input patterns and flight-path control strategies during a manual instrument approach as a function of recent flight practice.

Background: Manual flying skills erode due to an extensive use of automation and rare opportunity to practice these skills.

Method: One hundred and twenty-six randomly selected pilots of a European airline took part in this experiment, performing a simulated manual raw data precision approach. All of the pilots were allocated to 1 of 4 groups according to their fleet and rank: first officers and captains on short haul, as well as first officers and captains on long haul. A new method to analyze flight-path control strategies by differentiating between constant and variable flight-path errors was proposed. Time-domain measures were taken into account to evaluate sidestick inputs.

Results: We distinguished between 2 different flight-path control strategies; both differed in the deviations achieved. In addition, the pilots who predominantly used 1-dimensional sidestick inputs also had smaller deviations from the ideal flight-path.

Conclusion: Pilots showed a relationship between manual fine-motor flying skills and recent flight practice, especially in long-haul fleets.

The extent to which automation supports pilots is high in today's aviation. Consequently, manual skills are used less frequently but remain important, especially in the case of malfunctioning systems. Analyses of recent aircraft accidents such as Air France Flight 447 (Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile [BEA], 2012), Asiana Flight 218 (National Transportation Safety Board [NTSB], 2014), and Colgan Air Flight 3407 (NTSB, 2010) have shed light on inadequate manual flying skills due to a lack of recent practice and training.

According to a simplified model, analyses of *manual control* on fly-by-wire (FBW) aircraft involve two interdependent approaches. The first approach addresses the outcome of the human–machine system, where the flight-path as actually flown is compared to the intended flight-path. This reflects what can be called flight-path control strategies (FCSs) or an *outer control loop*, which drives the pilot's control inputs to the flight control system (or directly to the control surfaces). The operational quality of the outer control loop is often referred to as *manual flying performance* (Haslbeck & Hoermann, 2016). The second approach addresses the inputs on all flight control elements, predominantly the primary flight control or inceptor such as the control yoke or sidestick. Sidestick input patterns (SIPs) are influenced by automatized behavior routines, which are shaped by training

CONTACT Andreas Haslbeck haslbeck@tum.de Institute of Ergonomics, Technical University of Munich, Institute of Ergonomics, Boltzmannstrasse 15, Munich 85747, Germany.

Supplemental data for this article can be accessed on the publisher's website.

drills, individual experience, proficiency, and the current mental state of the pilot. According to our model, such SIPs reflect the so-called *inner control loop*. In this article, characteristic strategies of flight-path control and sidestick inputs (inner loop) are examined as potential explanatory factors for manual flight performance (outer loop). Efficient control patterns could serve as training goals in simulator refresher trainings. If characteristic individual patterns could be identified, for example, by computerized monitoring systems, they could serve as objective feedback to shape the pilots' control behavior. Objective evaluation is preferable, as instructor ratings can suffer from low interrater reliability (Gontar & Hoermann, 2015).

The Flight-Path Control Strategy: Flying the Needles

We previously compared flight-path performance data (Haslbeck & Hoermann, 2016) of pilots from long- to short-haul fleets. It was found that short-haul pilots with 16 or 17 landings a month maintain a higher level of fine-motor flying skills than long-haul pilots with only two or three landings a month. These results confirm previous findings by Veillette (1995), Gillen (2008), and Ebbatson (2009).

Pilots' experience, expressed by their cockpit position (captain vs. first officer), and age as a covariate, has also shown minor effects on fine-motor skills (Haslbeck & Hoermann, 2016). Consequently, we assume that these factors will also affect SIPs. There is one important limitation for this conclusion: It is important to be aware of the potentially confounding effects of factors such as the level of flying practice and of the aircraft type itself (short-haul Airbus A320 vs. long-haul A340). However, the FBW flight control system was designed with the idea of highest applicable commonality (Vadrot & Aubry, 1994) providing the same flying and handling qualities (Bissonnette & Culet, 2013; Brière & Traverse, 1993; Favre, 1994; Joint Aviation Authorities, 2004). Hence, we assume that both aircraft types can be controlled with similar precision, but we do not expect that both types can be hand-flown by applying identical control patterns. Consequently, this potentially confounding effect between factors such as the level of practice and the aircraft type needs to be kept in mind, particularly for the analysis of SIP.

Following Hollnagel's (2009) concept of the efficiency–thoroughness trade-off, we propose two generic FCSs: (a) an active, effortful strategy (*optimizer*) aiming to minimize flight-path deviations at the cost of greater workload, and (b) a passive, effortless strategy (*steady path*) allowing flight-path deviations within certain *windows of performance* (Morris & Miller, 1996), thereby saving resources. The former behavior would be characterized by a higher variability (variable error) in flight-path deviations (McClernon & Miller, 2011), and the latter behavior by a constant displacement from the ideal track (constant error). We assume that these FCSs are generally executed with only little conscious control. However, this does not mean that they cannot be switched or modified in the case of specific instructions, situational demands (e.g., terrain or traffic avoidance) or mental states (e.g., goals, workload, fatigue).

Sidestick Inputs: Stirring the Pot

Three methodologies can be distinguished with respect to performance measurement in manual control systems: time domain, frequency domain, and pilot model related (Baron, 1988). Baron suggested using time domain measures for flight-path control measurement and applying frequency domain parameters for inceptor inputs. This was picked up by Ebbatson, Harris, Huddlestone, and Sears (2008, 2012), who analyzed commercial pilots' manual flight performance by comparing flight-path data and control inputs with the results of power spectral density analyses. In their studies, participants were required to fly a manual precision approach with either symmetric or asymmetric thrust conditions. Flight-path performance showed no significant differences for the two conditions, but the participants developed different control patterns for the inceptor. In the asymmetric condition, participants showed lower frequencies but higher amplitudes in lateral control and higher amplitude for the rudder inputs.

Ruediger (2014) analyzed the control stick inputs of military pilots in an experiment with different levels of workload. She investigated the concept of *pilot gain* (Niewind, 2011) as the aggressiveness with which pilots operate the control stick, and discussed several approaches and metrics to quantify this aspect of pilots' manual control behavior.

Haslbeck, Gontar, and Schubert (2012) analyzed preliminary data from the study presented here with a focus on sidestick handling. The results showed that (1) about one third of pilots use more than one kind of grasp to handle the sidestick, and (2) participating long-haul captains (CPTs) exhibit more stick inputs as well as more two-dimensional roll and pitch inputs than short-haul first officers (FOs).

Compared to Ebbatson et al. (2008), our study focuses on time domain measures. Time domain measures are easier to interpret because they refer to the pilot's inceptor deflections per time. Pilots and flight instructors can relate time domain measures more easily to certain flight-path deviations than frequency domain measures. Therefore, they reveal strengths and weaknesses of operator manual control behavior more clearly. With the time-based approach, for example, the relative time period during which the pilot has actively manipulated the inceptor during a specific manual flight phase can be compared to the relative time period without any inputs. According to Kantowitz and Casper (1988) and Niewind (2012), active control inputs are closely correlated with workload. Consequently, flying with a minimum of control inputs can conserve mental and physical resources and prevent pilot-induced oscillations.

Another point is the dimensionality of sidestick inputs: one-dimensional (sequential = roll or pitch) inputs on one axis only, also denoted as an alternation pattern, compared to two-dimensional (simultaneous = roll and pitch) inputs on both axes. Flight instructors recommend the alternation input pattern for FBW aircraft, because simultaneous inputs on two dimensions increase the workload and can mask feedback information (see Kantowitz & Casper, 1988). Therefore, it is considered more error-prone and has already been demonstrated as less precise (Poulton, 1974; Ziegler, 1968). In non-FBW aircraft, the bank angle during turns causes a loss of vertical lift and consequently of altitude, which has to be compensated for by corresponding elevator inputs. Here, pilots have to control both roll and pitch axes simultaneously. During normal operation, the Airbus FBW flight control system compensates for this loss of lift when the aircraft turns around its roll axis (Airbus SAS, 2008). Consequently, pilots can separate the control of both axes on Airbus aircraft to attain greater accuracy on their intended flight-path during turns.

Pilots who participated in our study were asked beforehand whether they would immediately correct flight-path deviations or tolerate them until they exceeded a certain threshold, thereby applying *windows of performance* (McClernon & Miller, 2011). These subjective data have shown a significant altitude effect (above vs. below 1,000 ft above ground) on pilots' behavior (Haslbeck & Gontar, 2014). Pilots prefer immediate corrective actions in the lower altitude segment as compared to the higher altitude segment.

Research Questions and Hypotheses

To examine the FCSs and SIPs of pilots the following research questions (RQs) and corresponding hypotheses (H) have been derived:

RQ1: Do airline pilots develop distinct FCSs?

H1: There are two opposing FCSs (McClernon & Miller, 2011; Morris & Miller, 1996): *optimizer* and *steady path*. Pilots' control behaviors can be categorized to a large extent into these two strategies.

RQ2: If distinct FCSs can be identified, do they differ in terms of efficiency?

H2: Average flight-path deviations are generally smaller for pilots using the optimizer strategy.

RQ3: Do pilots show persisting SIPs, or is their behavior adaptive?

H3: There are two different SIPs, predominantly one-dimensional control and predominantly two-dimensional control.

RQ4: Are all of those control patterns related to the pilot's level of practice and experience?

H4: Based on previous work (Haslbeck & Hoermann, 2016), we expect that recent flight practice will have a greater effect on control strategies than the level of experience.

RQ5: Do pilots adjust their FCSs and SIPs as functions of altitude or flight-path deviations?

H5: Pilots adjust their control patterns in relation to altitude due to different flying tasks.

RQ6: Are FCSs and SIPs systematically related to each other?

H6: FCSs correlate with SIPs.

Method

The data analyzed in this article were gathered in two complementary experiments in 2011 and 2013. These data have previously been analyzed (Haslbeck & Hoermann, 2016) for flight-path deviations only, whereas this article adopts the same data set to identify distinct FCSs and SIPs and to examine their relations to manual flying performance.

Participants

Pilots' level of practice and training was the independent variable of this experiment. For the experiment, the crew roster department of one cooperating European airline randomly selected professional airline pilots holding valid Airline Transport Pilot licenses: Airbus A320 FOs (2F) and CPTs (2C) as well as Airbus A340 FOs (4F) and CPTs (4C). Based on a power analysis, we were aiming for 30 pilots in each group. Table 1 shows the demographic information for all 126 participating pilots assigned to the four groups. Age and overall flight hours reflect increasing long-term flight experience, and the number of landings performed as the pilot flying (PF) in the 30 days prior to the experiment represents a measure of daily flight practice. The number of years since flight school addresses long-term skill retention.

Scenario and Instruction

All participants were required to fly a 10-min manual flight scenario approaching Munich Airport (26R EDDM) under raw data conditions (without autopilot, without flight director), with the

Table 1. Demographic data for participants (Table 1 from Haslbeck & Hoermann, 2016).

				Age		Flight Hours: Overall/On Type		Landings in Past 30 Days		Years Since Flight School	
	Rank	Fleet	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
2F	FO	A320	39	30.1	2.8	3,438/2,415	1,848/1,266	16.1	6.3	5.8	2.8
4F	FO	A340	28	36.4	3.3	7,204/3,469	1,987/1,812	2.4	1.5	12.2	2.9
2C	CPT	A320	30	43.0	4.3	11,276/3,847	1,931/2,355	16.6	10.2	18.1	3.3
4C	CPT	A340	29	49.8	3.6	14,969/2,909	2,951/1,818	3.5	2.1	24.4	4.1

Note. FO = first officer; CPT = captain.

exception of auto thrust being available in normal law mode. Weather parameters were set to 1,200 m visibility, constant but gusty wind of 220°/17 to 22 kt, and a 270 ft ceiling with light rain. These conditions are well within the capabilities of a proficient pilot with more than 1,000 flight hours on the respective aircraft type, even if the weather conditions imply a medium to high task load during the approach. The experimental scenario started with a malfunctioning flight director and autopilot 8 min before touchdown. Most pilots also executed manual power control (i.e., auto thrust disengaged). The pilots were instructed to perform the approach according to licensing standards and company rules as accurately as possible (i.e., minimum deviations from localizer, glideslope, and airspeed).

Apparatus and Dependent Measures

The experiments were conducted in two certified, full-motion (FFS Level D) flight simulators, one Airbus A320 and one Airbus A346. For analysis, we divided the approach phase into three segments based on different altitude levels, also representing three different levels of difficulty. The first segment between 3,000 and 1,000 ft above ground level (AGL) represents the preparation of a *stabilized approach* (SKYbrary, 2016). This means acclimatization to manual flight with more tolerances to deviations from the ideal flight-path and medium difficulty (mean duration on A320 M = 158 s; on A340 M = 144 s). According to company rules, the approach had to be stabilized at 1,000 ft AGL, which subsequently means a more difficult task under instrument meteorological conditions until reaching 270 ft AGL (M = 58 s/52 s). During this second altitude segment the criteria for a stabilized approach apply; therefore, it is the most demanding part for the pilots as well as the most relevant one for the evaluation of manual skills. Passing the cloud layer at 270 ft AGL, the runway became visible, defining a visual approach of medium difficulty. This third segment ended at 50 ft above ground (M = 21 s/20 s) shortly before the landing flare was initiated.

To analyze FCS, the flight-path deviations were calculated for both localizer and glideslope. The mean error (ME) of these two scores shows the deviation from the instrument landing system (ILS); this was compared with the corresponding standard deviation (SD), a measure of variability (McClernon & Miller, 2011). For visualization, the results were plotted in a two-quadrant diagram (Figure 1) using Cartesian coordinates. The same data can be considered as polar coordinates: The vector's length from the origin to the data point is equal to the root mean square error (RMSE) of the respective deviation, and the angle α represents the ratio between the aforementioned strategies. $ME > SD$ (0°–40° and 140°–180°) represents a

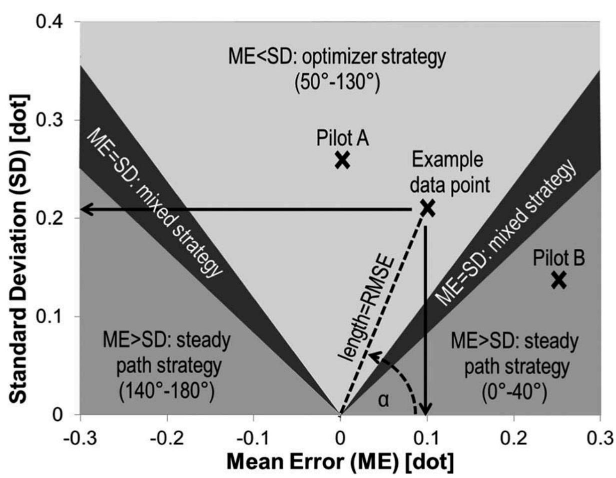

Figure 1. Generic diagram for the analysis of flight-path control strategies showing one example data point representing a specific flight-path as well as two examples that were used to explain the results, Pilot A (optimizer) and Pilot B (steady path).

generally smooth flight-path. The corresponding FCS was called *steady path*. $ME < SD$ (50°–130°) represents a more fluctuating flight-path, denoted as the *optimizer* strategy. A pilot using the steady path strategy (Pilot B, Figure 1) would fly a generally smoother flight-path, leaving minor deviations uncorrected. The optimizer strategy (Pilot A, Figure 1) involves several corrections resulting in an oscillating flight-path. At the same time, both pilots (A and B) could have the same RMSE. For data points with more or less equal mean and standard deviation, none of these strategies dominated and the strategy was denoted as *mixed*. The mixed strategy was assigned to data points within a cone of ±5° about the $ME = SD$ bisectrix. This value was selected as a trade-off between the need for relevant sample sizes for both distinct strategies and the need for detectable differences among the groups.

Regarding the SIP, recorded sidestick deflections for pitch and roll were counted if they were larger than a threshold of 0.15°. The percentage of stick inputs below the threshold is denoted as *no input*, as it might have resulted from unintended slight deflections while the pilot's hand was resting on the stick. For the one- or two-dimensional patterns, stick inputs on either one or both axes had to be above the threshold. The RMSE for ILS deviation and sidestick deflection was calculated as a typical measure for tracking performance analysis (Baron, 1988; McClernon & Miller, 2011; Rantanen, Johnson, & Talleur, 2004).

Results

Flight-Path Control Strategies

All effects are reported as significant at $p < .05$ and where possible η_p^2 as well as r are given as measures for effect size. For every participant, the predominant FCS ($ME < SD$, $ME \approx SD$, $ME > SD$) was determined separately for all three altitude segments on localizer and glideslope (see Figure 2 for one data set and online supplement for remaining individual data).

Regarding the localizer (Figure 3), both distinct strategies occurred with approximately equal frequency when fleet and altitude were not distinguished (46.0% optimizer, 40.2% steady path, 13.8% mixed). However, chi-square tests of independence for lateral control showed that the distribution of the three control strategies was not independent of fleet, $\chi^2(2) = 6.04$, $p = .049$; and altitude, $\chi^2(4) = 37.65$, $p < .001$. In these chi-square analyses, the CPTs and FOs of each fleet had to be collapsed into a single group to ensure the expected frequency counts of five and greater. Generally, more A340 pilots (46.8%) than A320 pilots

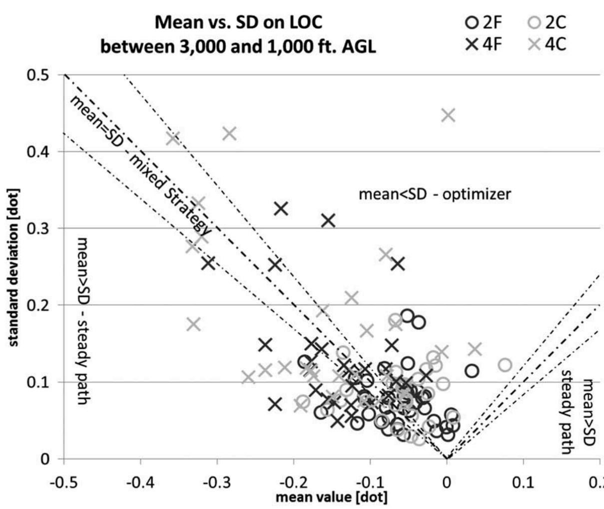

Figure 2. Exemplary comparison of means and standard deviations of flight-path deviations from localizer for the altitude segment between 3,000 and 1,000 ft above ground level (AGL).

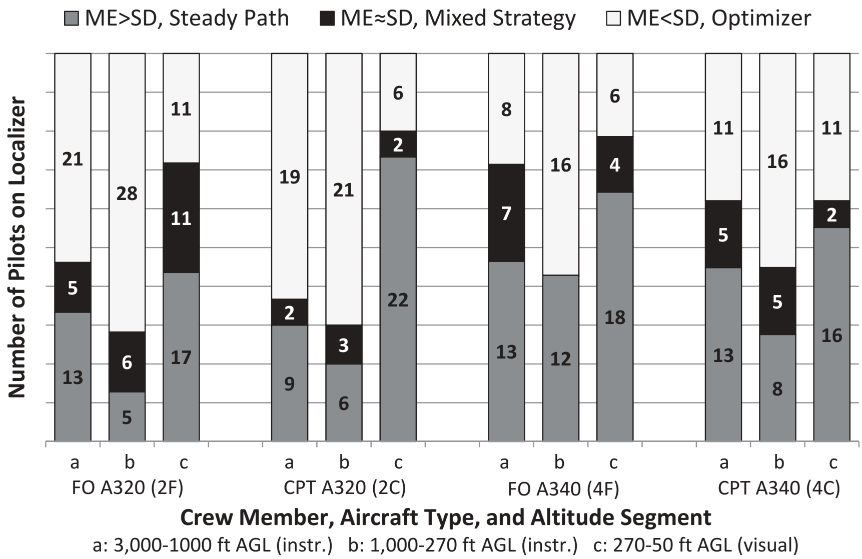

Figure 3. Flight-path control strategies on localizer. Each set of columns shows one group of pilots and the control strategies used for the three altitude segments. This data confirms Hypothesis 1 for the localizer.

(34.8%) followed the steady path strategy. In the middle-altitude segment there was a tendency toward the optimizer strategy (64.3% optimizer vs. 24.6% steady path), whereas a tendency toward the steady path strategy was observable in the lower altitude segment (27.0% optimizer vs. 57.9% steady path). Both of the latter two tendencies were more pronounced in the A320 fleet than in the A340 fleet.

For the glideslope, however, the optimizer was clearly the predominant strategy (Figure 4) in almost all cases (68.3% optimizer vs. 20.6% steady path). Chi-square analyses of independence of the vertical control strategies from fleet and altitude were conducted. Both effects were significant with $p < .01$. More short-haul (77.8%) than long-haul pilots (56.7%) preferred the optimizer strategy, $\chi^2(2) = 22.70$, $p < .001$. In the lower altitude segment, vertical control increasingly followed the steady path strategy

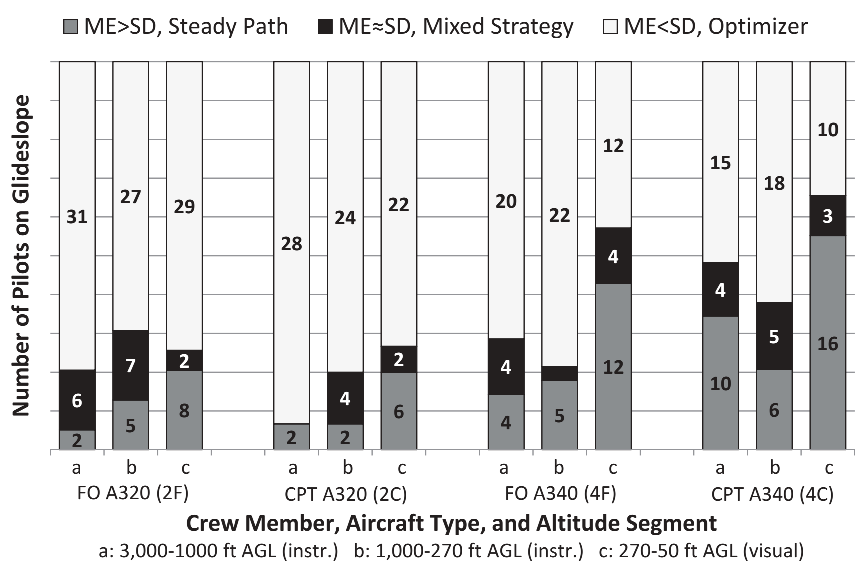

Figure 4. Flight path control strategies on glideslope. Each set of columns shows one group of pilots and the control strategies used for the three altitude segments. This data confirms Hypothsis 1 for the glideslope.

(57.9% optimizer vs. 33.3% steady path) across the fleets, $\chi^2(4) = 19.05$, $p < .001$. In particular, several long-haul pilots changed to the steady path strategy in the lower altitude segment (49.1% A340 pilots with steady path vs. 20.3% A320 pilots with steady path below 270 ft).

Table 2 summarizes the transition frequencies for the three altitude segments. The largest number of pilots applied and retained the optimizer strategy (32.5%); these were mostly short-haul crew members. The second largest group started with the optimizer strategy and changed to steady path during the approach (15.1%). Long-haul pilots seem to maintain the steady path strategy especially for the localizer (14.0%).

With one exception (A340, 3,000–1,000 ft AGL, localizer), all optimizer strategies were associated with smaller ILS deviations. Table 3 shows mean scores of the RMSE values for every altitude segment on localizer and glideslope separately for both control strategies. Three Bonferroni-corrected independent *t* tests indicated significant differences in ILS deviations between both control strategies in the lower altitude segment, but no significant differences in the upper and middle altitude segments. Overall, the results confirm that the accuracy of the ILS tracking was significantly related to the applied FCS when approaching the runway.

As shown in Table 4, the optimizer strategy did not necessarily require more stick deflections. None of the *t* tests was significant for both axes. A340 pilots used even fewer stick inputs when applying the optimizer strategy for pitch control in the visual segment. Only during the stabilized

Table 2. Frequencies of transitions between flight-path control strategies.

	2F	2C	4F	4C	Total
Only steady path in at least two altitude segments on both ILS dimensions, no strategy change	1	0	2	6	9
Only optimizer in at least two altitude segments on both ILS dimensions, no strategy change	19	13	4	5	41
Transition from optimizer on both ILS dimensions to steady path on both ILS dimensions	5	3	7	4	19
Transition from steady path on both ILS dimensions to optimizer on both ILS dimensions	2	0	1	2	5
Always opposite strategies on both ILS dimensions	3	4	3	0	10

Note. The existence of different flight-path control strategies (Hypothesis 1) was confirmed and several pilots also interchanged between them. ILS = instrument landing systems.

Table 3. Comparison of instrument landing system deviations (root mean square error) between optimizer and steady path flight-path control strategies for short- (A320) and long-haul (A340) pilots.

	Localizer (dot)						Glideslope (dot)					
	3,000–1,000		1,000–270		270–50		3,000–1,000		1,000–270		270–50	
Altitude Segment (ft AGL)	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
A320 optimizer	0.10	0.04	0.12	0.06	0.11[a]	0.06	0.13	0.05	0.17	0.07	0.62	0.40
A320 steady path	0.12	0.05	0.15	0.07	0.20[a]	0.11	0.17	0.05	0.18	0.11	0.76	0.42
A340 optimizer	0.28	0.24	0.24	0.09	0.17[b]	0.11	0.27	0.12	0.31	0.13	0.73[c]	0.53
A340 steady path	0.24	0.14	0.28	0.14	0.29[b]	0.15	0.39	0.25	0.41	0.24	1.30[c]	0.80

Note. One-tailed independent *t* tests indicate statistically significant differences: [a]$t(51.63) = 3.63$, $p < .012$, $r = .45$; [b]$t(49) = 2.80$, $p = .048$, $r = .37$; [c]$t(46.69) = 3.05$, $p = .024$, $r = .98$. Optimizers achieved smaller flight-path deflections (significant for the lower altitude segment) confirming Hypothesis 2. For the *t* tests, the *t* statistic (*t*), degrees of freedom (*df*), Bonferroni-corrected *p* values (*p*), and effect sizes (*r*) are reported. AGL = above ground level.

Table 4. Comparison of sidestick inputs (root mean square error) between optimizer and steady path flight-path control strategies.

	Roll Axis (Degree)						Pitch Axis (Degree)					
	3,000–1,000		1,000–270		270–50		3,000–1,000		1,000–270		270–50	
Altitude segment (ft AGL)	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
A320 optimizer	1.91	0.57	2.38	0.90	4.53	1.57	0.89	0.29	1.04	0.46	2.02	0.73
A320 steady path	1.87	0.55	2.33	0.57	4.54	1.51	0.86	0.29	0.70	0.23	2.04	0.45
A340 optimizer	2.77	0.64	3.67	1.21	5.69	2.13	1.06	0.25	1.45	0.63	2.54	0.56
A340 steady path	2.94	1.09	3.33	0.90	5.32	1.76	1.35	0.62	1.29	0.63	3.08	0.79

approach phase did the A320 pilots have more stick inputs when using the optimizer strategy for pitch control. In summary, Table 3 and Table 4 demonstrate that the optimizer strategy corresponded to smaller ILS deviations, which were not at the cost of applying more frequent inputs.

Sidestick Input Patterns: Amount and Dimensionality of Stick Inputs

Sidestick inputs were analyzed for the roll axis (Figure 5) and the pitch axis (Figure 6) by a 2 × 2 × 3 (between-factor fleet [A320, A340] × between-factor rank [FO, CPT] × within-factor altitude [3,000–1,000, 1,000–270, 270–50]) multivariate analysis of variance (MANOVA; Table 5). Sidestick input data were not normally distributed, but positively skewed. Hence, these data were log10-transformed for the statistical analysis. The analysis revealed significant effects of fleet, rank, and altitude.

The extent and dimensionality of sidestick inputs are presented in Figure 7 as percentages for the SIP (and in the online supplement for individuals). A statistical analysis of the absolute numbers of inputs was done by three separate 2 × 2 × 3 (between-subjects fleet [A320, A340] × between-subjects rank [FO, CPT] × within-subjects altitude [3,000–1,000, 1,000–270, 270–50]) mixed analyses of variance (ANOVAs; Table 6). Separate ANOVAs were chosen instead of one MANOVA, because the different variables were not independent of each other. Univariate tests found significant effects of fleet, altitude, and the interaction between fleet and altitude. The largest effect sizes were found for altitude. The results showed that long-haul pilots always applied significantly more inputs and more combined inputs than their short-haul colleagues. Two different and clear tendencies with respect to altitude are visible: With decreasing altitude, (a) the number of inputs and (b) the number of combined inputs increased in both fleets.

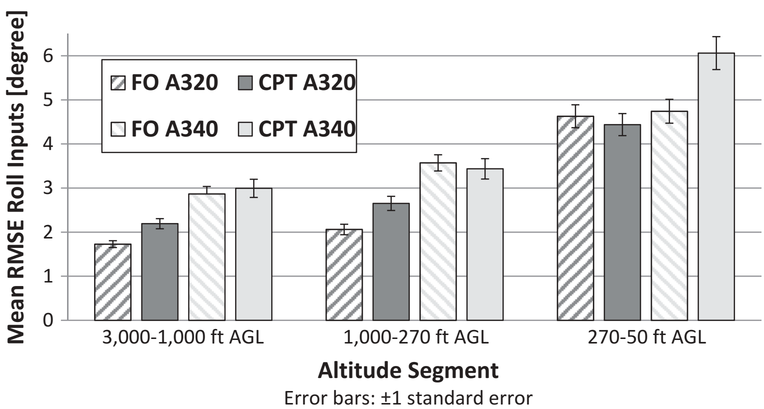

Figure 5. Mean (root mean squared error) amplitude of sidestick inputs on roll axis.

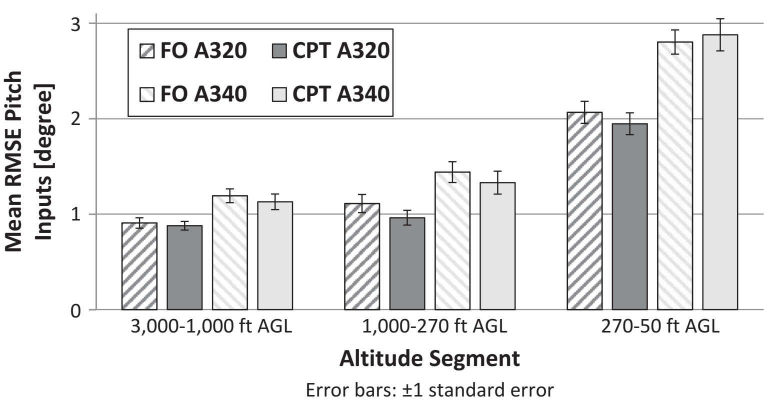

Figure 6. Mean (root mean squared error) amplitude of sidestick inputs on pitch axis.

Table 5. Statistical analysis of log10-transformed sidestick input data.

Source	Multivariate Tests	Univariate Tests
Between subjects		
Fleet	V = .29, F(2, 121) = 24.23, p < .001, η_p^2 = .29	Roll: F(1, 122) = 46.77, p < .001, η_p^2 = .28 Pitch: F(1, 122) = 30.95, p < .001, η_p^2 = .20
Rank	V = .12, F(2, 121) = 7.95, p = .001, η_p^2 = .12	Roll: F(1, 122) = 5.59, p = .020, η_p^2 = .04 Pitch: F(1, 122) = 0.81, p = .369, η_p^2 = .00, *ns*
Fleet * Rank	V = .01, F(2, 121) = 0.82, p = .444, η_p^2 = .01, *ns*	
Within subjects		
Altitude	V = .90, F(4, 119) = 253.68, p < .001, η_p^2 = .90	Roll: ε = .87, F(1.74, 211.75) = 328.74, p < .001, η_p^2 = .73 Pitch: F(2, 244) = 377.87, p < .001, η_p^2 = .76
Altitude * Fleet	V = .18, F(4, 119) = 6.57, p < .001, η_p^2 = .18	Roll: ε = .87, F(1.74, 211.75) = 10.19, p < .001, η_p^2 = .08 Pitch: F(2, 244) = 1.29, p < .276, η_p^2 = .01, *ns*
Altitude * Rank	V = .01, F(4, 119) = .41, p = .802, η_p^2 = .01, *ns*	
Altitude * Fleet * Rank	V = .16, F(4, 119) = 5.83, p < .001, η_p^2 = .16	Roll: ε = .87, F(1.74, 211.75) = 14.78, p < .001, η_p^2 = .11 Pitch: F(2, 244) = 0.39, p = .681, η_p^2 = .00, *ns*

Note. For these multivariate analyses of variance (Pillai's statistic) log10-transformed sidestick input data were taken to maintain normality. Normality was not met (1 out of 24): pitch, 3,000–1,000 ft above ground level, captains on A320. Equality of covariance matrices and error variances was not met (p = .049); however, due to only two dependent variables, we considered this violation as minor (Tabachnick & Fidell, 2007, p. 252). Greenhouse–Geisser estimates of sphericity (ε) were applied when the assumption of sphericity was not met. These findings support Hypothesis 4 and Hypothesis 5.

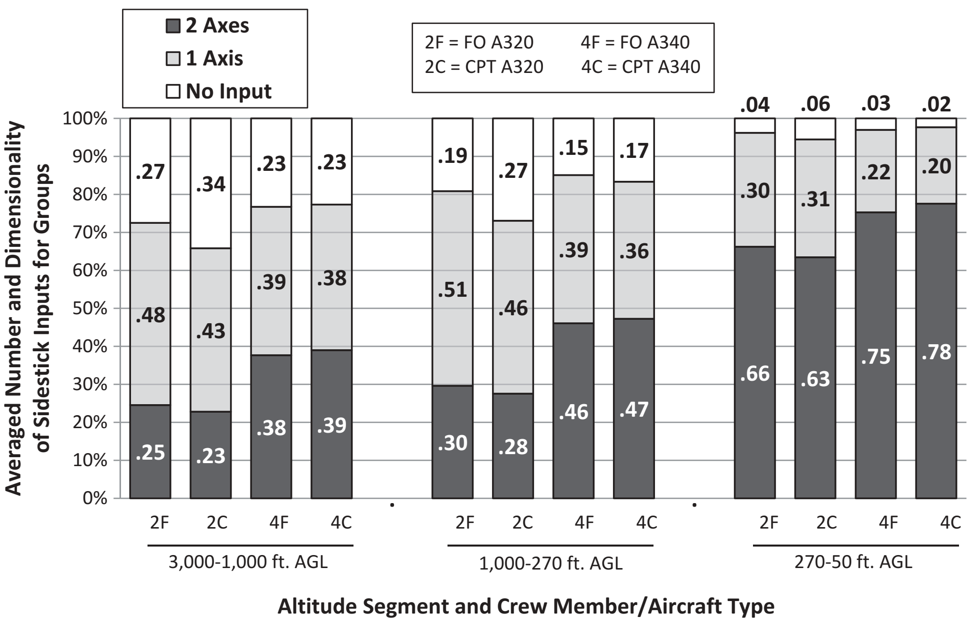

Figure 7. Averaged relative number and dimensionality of sidestick inputs for three altitude segments normalized to 100%, supporting Hypothesis 3.

Correlational Analysis Between Flight-Path and Inceptor Inputs

The RMSEs of ILS deviations and sidestick deflections were log10-transformed to ensure the applicability of the normality assumption. The transformed values were compared using two-tailed Pearson's

Table 6. Statistical analysis of the number and dimensionality of sidestick inputs.

Source	Univariate Tests
Between subjects	
Fleet	No inputs: $F(1, 122) = 10.04$, $p = .002$, $\eta_p^2 = .08$ °One axis: $F(1, 122) = 48.63$, $p < .001$, $\eta_p^2 = .29$ ^Two axes: $F(1, 122) = 44.24$, $p < .001$, $\eta_p^2 = .27$
Rank	No inputs: $F(1, 122) = 2.39$, $p = .125$, $\eta_p^2 = .02$, *ns* °One axis: $F(1, 122) = 3.63$, $p = .059$, $\eta_p^2 = .03$, *ns* ^Two axes: $F(1, 122) = .02$, $p = .893$, $\eta_p^2 = .00$, *ns*
Fleet * Rank	No inputs: $F(1, 122) = 2.12$, $p = .148$, $\eta_p^2 = .02$, *ns* °One axis: $F(1, 122) = .31$, $p = .578$, $\eta_p^2 = .00$, *ns* ^Two axes: $F(1, 122) = .73$, $p = .394$, $\eta_p^2 = .01$, *ns*
Within subjects	
Altitude	No inputs: $F(2, 244) = 233.67$, $p < .001$, $\eta_p^2 = .66$ One axis: $\varepsilon = .90$, $F(1.80, 219.34) = 198.96$, $p < .001$, $\eta_p^2 = .62$ Two axes: $F(2, 244) = 656.80$, $p < .001$, $\eta_p^2 = .84$
Altitude * Fleet	No inputs: $F(2, 244) = 4.32$, $p = .014$, $\eta_p^2 = .03$ One axis: $\varepsilon = .90$, $F(1.80, 219.34) = 2.25$, $p = .114$, $\eta_p^2 = .02$, *ns* Two axes: $F(2, 244) = 3.84$, $p = .023$, $\eta_p^2 = .03$
Altitude * Rank	No inputs: $F(2, 244) = 1.88$, $p = .154$, $\eta_p^2 = .02$, *ns* One axis: $\varepsilon = .90$, $F(1.80, 219.34) = 2.19$, $p = .120$, $\eta_p^2 = .02$, *ns* Two axes: $F(2, 244) = .01$, $p = .995$, $\eta_p^2 = .00$, *ns*
Altitude * Fleet * Rank	No inputs: $F(2, 244) = .66$, $p = .520$, $\eta_p^2 = .01$, *ns* One axis: $\varepsilon = .90$, $F(1.80, 219.34) = 1.72$, $p = .185$, $\eta_p^2 = .01$, *ns* Two axes: $F(2, 244) = .11$, $p = .893$, $\eta_p^2 = .00$, *ns*

Note. Greenhouse–Geisser estimates of sphericity (ε) were applied when the assumption of sphericity was not met. ° The assumption of homogeneity of variance was violated for the lowest altitude segment, $F(3, 122) = 2.94$, $p = .036$. ^ The assumption of homogeneity of variance was violated for the lowest altitude segment, $F(3, 122) = 3.10$, $p = .029$. These findings support Hypothesis 4 and Hypothesis 5.

correlations (Table 7). According to the correlation results, larger ILS deviations were related to larger stick deflections, especially for glideslope deviations and pitch inputs. The highest correlations were found for the 1,000 to 270 ft AGL segment, again between glideslope deviations and the pitch axis.

To compare the effects of one- and two-dimensional stick inputs for each fleet and altitude segment separately, the pilot data were ordered into two groups: pilots with a majority of one-dimensional SIPs compared to pilots with a majority of two-dimensional inputs (Table 8). For this analysis, averaged values of localizer and glideslope were calculated for each pilot in each altitude segment.

The results showed that pilots with more one-dimensional inputs (operating roll and pitch axes discretely) achieved overall smaller ILS deviations. This comparison could not be made for the A340 pilots in the lowest altitude segment due to the long-haul pilots' large number of two-dimensional inputs.

Table 7. Correlations between flight-path deviations and sidestick deflections.

	Localizer—Roll Axis			Glideslope—Pitch Axis		
	3,000–1,000	1,000–270	270–50	3,000–1,000	1,000–270	270–50
2F	$r = .21$	$r = .33$*	$r = .45$**	$r = .39$*	$r = .58$**	$r = .48$**
2C	$r = .19$	$r = .19$	$r = .35$	$r = .47$**	$r = .39$*	$r = .36$
4F	$r = .33$	$r = .21$	$r = .35$	$r = .41$*	$r = .64$**	$r = .18$
4C	$r = .56$**	$r = .30$	$r = .04$	$r = .62$**	$r = .71$**	$r = .32$

Note. These correlations support Hypothesis 6.
*Two-tailed Pearson's correlations significant at the .05 level. **Two-tailed Pearson's correlations significant at the .01 level.

Table 8. Comparison of combined instrument landing system (ILS) deviations between pilots with more one-dimensional inputs and pilots with more two-dimensional inputs.

	3,000–1,000 ft AGL				1,000–270 ft AGL				270–50 ft AGL			
Altitude Segment	*M*	*SD*	*n*	*t*	*M*	*SD*	*n*	*t*	*M*	*SD*	*N*	*t*
A320, more one-dimensional inputs	0.12	0.04	61	$t(67) = 2.41$, $p = .050$,	0.15	0.06	53	$t(67) = 1.56$, $p = .310$,	0.29	0.17	6	$t(67) = 1.35$, $p = .455$,
A320, more two-dimensional inputs	0.15	0.04	8	$r = .28$	0.17	0.05	16	$r = .19$	0.43	0.26	63	$r = .16$
A340, more one-dimensional inputs	0.25	0.16	31	$t(55) = 1.81$, $p = 0.190$,	0.25	0.09	22	$t(55) = 2.07$, $p = .110$,	—	—	0	
A340, more two-dimensional inputs	0.33	0.15	26	$r = .24$	0.32	0.13	35	$r = .27$	0.65	0.38	57	

Note. An averaged value (mean) of localizer and glideslope error, the standard deviation (*SD*), and the number of cases (*n*) are given for ILS deviations. For the one-tailed independent *t* tests, the *t* statistic (*t*), degrees of freedom (*df*), Bonferroni-corrected *p* values (*p*), and effect sizes (*r*) are reported. AGL = above ground level.

Discussion

Strategies for the Flight-Path (RQ1 & RQ2)

This study introduced a method to describe and identify different FCSs during a manual ILS approach. This method is based on the mutual proportions of constant error (average deviation from ideal flight-path) and variable error (average variability around the ideal flight-path). By applying the method, we matched the observed strategies to two distinct strategies, confirming H1: (a) the optimizer strategy (Pilot A, Figure 1) with smaller constant error and larger variable error, and (b) the steady path strategy (Pilot B, Figure 1) with larger constant error and smaller variable error. The glideslope is more difficult to control in a straight, stabilized approach, because of (a) a lower angle of ±0.4° in the funnel-shaped ILS approach compared to ±0.8° for the localizer, (b) the fact that two primary control instruments (sidestick, thrust levers) and configuration (flaps, gear) influence glideslope deviations (see Johnson & Pritchett, 2002), and (c) the glideslope refers to a continuously changing altitude, whereas the localizer corresponds to a constant heading (Haslbeck & Bengler, 2016). Consequently, we expected to find the steady path strategy applied for the localizer, and the optimizer strategy applied more often on the glideslope. The optimizer strategy was indeed found more often, especially in the stabilized approach segment (1,000–270 ft AGL), which is difficult and demanding because of stricter tolerances for deviations. The optimizer strategy was predominantly applied by short-haul pilots. This might be partly due to the greater agility of the A320 compared to that of the A340. However, a few exceptions were observed.

Below 270 ft AGL, pilots showed a tendency to switch from the optimizer to the steady path strategy at least for lateral control (Table 2 and Figure 3). In comparison to the long-haul pilots, many short-haul pilots retained the optimizer strategy for vertical control in this phase (Figure 4). With the optimizer strategy, pilots often achieved smaller flight-path deviations (Table 3), confirming H2. In 3 out of 12 comparisons, flight-path deviations differed significantly between both strategies. Thus, more active control behavior leads to better results in terms of fine-motor flight performance. The long-haul pilots' preference for the steady path strategy might thus indicate an efficiency–thoroughness trade-off resulting from limited amount of recent practice and therefore a higher mental load.

Patterns for Sidestick Inputs (RQ3 & RQ4)

The analysis of sidestick inputs also revealed different control patterns for the inceptor. SIPs were observed to be either one- or two-dimensional for both axes. The Airbus FBW flight control system makes it easy to use discrete inputs for roll and pitch. Because of the clearer feedback, training concepts for Airbus fleets contain recommendations for one-dimensional inputs to the stick axes. Combined inputs lead to (a) a reduced roll rate (International Air Traffic Association, 2015, p. 19), and can cause (b) an asymmetric stress locally overstressing the structure (e.g., the engine mount). In very rare cases, a roll input combined with pitch inputs near the maximum angle of attack can also unilaterally exceed the critical angle of attack thus leading to an upset (spin) due to an asymmetric stall (Stowell, 1996, pp. 88–89). More practical arguments for separate inputs are the higher attainable accuracy and reduced monitoring capacity required for one-dimensional control.

Pilots on long- and short-haul flights showed significantly different SIPs. In previous work, we could attribute the flight-path performance to recent flying practice (Haslbeck & Hoermann, 2016). However, for sidestick inputs we cannot definitively separate the influence of flight practice from the influence of the different aircraft types. Short-haul pilots applied more one-dimensional inputs than long-haul crews, who preferred two-dimensional inputs. At first glance this finding supports H4 and might indicate superior fine-motor control behavior on the part of short-haul pilots, who have more daily practice than long-haul pilots. However, it could also be due to an additional influence of agility differences of the aircraft.

Looking at the altitude segments, we found even stronger effects on the characteristics of stick inputs. Pilots increased the amount of control activity, especially the number of combined inputs, as the approach proceeded (Figure 7), which supports H3 as well as H5 and indicates an adaptation to the task. The majority of combined inputs were observed during the last altitude segment: About two thirds (short-haul) and three quarters (long-haul) of the lowest altitude segment was marked by two-dimensional inputs, also denoted as *stirring the pot*. If the aircraft is still deviating from localizer or glideslope shortly before touchdown, the urgency of corrective actions increases, making hectic control behavior more likely. When comparing this finding to Figure 3 and Figure 4, it becomes clear that steady path does not mean a steady hand. The final approach segment can be described by two apparently contradicting facts: (a) the highest number of combined inputs and (b) the highest number of steady path strategies. This indicates that pilots partly overcontrolled their aircraft by too many small superposing sidestick inputs (see Gray, 2007). Performing more flight-path corrections, meaning more control activities and observing lingering ILS deviations, increases workload and hence reduces available resources for other piloting tasks (see Ruediger, 2014). The quality of manual flight training could be enhanced by application of metrics as reported in this study to provide timely feedback on a trainee's control activities. The method of grip used to hold the sidestick needs to be addressed first (see Haslbeck et al., 2012) to establish more consistent, homogeneous ways to handle the inceptor; second, input dimensionality needs to be reduced by favoring more discrete and fewer combined inputs (Poulton, 1974; Ziegler, 1968) for this compensatory tracking task. Direct feedback to the trainee can thereby improve the training progress.

Correlation Between Flight-Path Control and Sidestick Inputs (RQ5 & RQ6)

FCSs and SIPs obviously share the strong influence of altitude, confirming H5. Pilots are required to adjust their control behavior during the different approach phases due to changes in the flying task's requirements. Through training and experience, pilots acquire a repertoire of different patterns, which they apply in relation to situational demands. Confirming H6, we found several positive correlations between the number of pitch inputs and the glideslope deviations as well as between the number of roll inputs and localizer deviations (Table 7). These correlations make sense because larger deviations normally require larger compensating inputs. Further findings showed that pilots preferring the alternation pattern (one-dimensional sidestick inputs) achieved smaller flight-path deviations than pilots using two-dimensional inputs (Table 8), which is in accordance with our expectations (Poulton, 1974).

Limitations

Our analyses included only pilots from one airline with the corresponding training scheme and specific career model. Replication within another airline would be desirable. In addition, we only considered Airbus aircraft with FBW and sidestick technology, which might affect pilots' SIPs. Even though Airbus maximizes the commonalities of control characteristics across aircraft types, we cannot assume that the A320 and the A340 can be identically hand flown (Haslbeck & Hoermann, 2016). There is also a difference in the final approach speeds between both types (A320 = 137 kt; A340 = 154 kt), resulting in differing durations of the three altitude segments.

Recommendations

Based on our findings, we see several implications for the training of pilots' manual flying skills. The methods described here allow for detailed analyses of desirable FCSs. Pilots' demonstrated fine-motor flight performance deficits should be counteracted by more frequent performance-based training opportunities. For example, the advantages of discrete inputs to both sidestick axes in terms of higher precision and reduced workload should be demonstrated to maintain or regain the ability to control the aircraft with one-dimensional inceptor inputs. Without additional practice, the risk of further skill deterioration increases, especially for long-haul pilots. If a company chooses to constrain manual flying by strict use-of-automation-in-flight policies as a mitigation strategy, their flight crews' performance would progressively depend on failproof automated control systems of the aircraft, an assumption that cannot be taken for granted.

Acknowledgments

The authors acknowledge the support of Ekkehart Schubert in both flight simulator experiments and the support of Tanja Herfert and Katharina Hasselmann in data analyses.

Funding

This research was partially funded by the German Federal Ministry for Economic Affairs and Energy via Federal Aeronautical Research Program LuFo IV-2 (20V0803).

ORCID

Andreas Haslbeck http://orcid.org/0000-0001-6486-0817
Hans-Juergen Hoermann http://orcid.org/0000-0002-6358-7027
Patrick Gontar http://orcid.org/0000-0003-0345-2566

References

Airbus, S.A.S. (2008). *A318/A319/A320/A321 Flight Crew Training Manual*. Blagnac, France: Airbus S.A.S.

Baron, S. (1988). Pilot Control. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 347–386). San Diego, CA: Academic Press.

Bissonnette, N., & Culet, J. P. (2013). *Airbus A318/A319/A320/A321/A330/A340* (Flight Standardization Board (FSB) Report). Renton. Retrieved from http://fsims.faa.gov/wdocs/fsb/a320%20rev%204.htm

Brière, D., & Traverse, P. (1993). AIRBUS A320/A330/A340 Electrical Flight Controls: A Family of Fault-Tolerant Systems. *Proceedings of The Twenty-Third International Symposium on Fault-Tolerant Computing*, 616–623. doi:10.1109/FTCS.1993.627364

BEA. (2012). *Final Report on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF 447 Rio de Janeiro - Paris*. Le Bourget, France.

Ebbatson, M. (2009). *The Loss of Manual Flying Skills in Pilots of Highly Automated Airliners* (PhD Thesis). Cranfield University, Cranfield.

Ebbatson, M., Harris, D., Huddlestone, J., & Sears, R. (2008). Combining Control Input with Flight-Path Data to Evaluate Pilot Performance in Transport Aircraft. *Aviation, Space, and Environmental Medicine*, *79*(11), 1061–1064. doi:10.3357/ASEM.2304.2008

Ebbatson, M., Harris, D., Huddlestone, J., & Sears, R. (2012). Manual Flying Skill Decay: Evaluation Objective Performance Measures. In A. J. D. Voogt & T. C. D'Oliveira (Eds.), *Mechanisms in the chain of safety. Research and operational experiences in aviation psychology*. Surrey, England, Burlington, VT: Ashgate.

Favre, C. (1994). Fly-by-wire for commercial aircraft: The Airbus experience. *International Journal of Control*, *59*(1), 139–157. doi:10.1080/00207179408923072

Gillen, M. (2008). *Degradation of piloting skills* (Master's Thesis). Grand Forks, ND: University of North Dakota.

Gontar, P., & Hoermann, H.-J. (2015). Interrater reliability at the top end: Measures of pilots' nontechnical performance. *The International Journal of Aviation Psychology*, *25*(3–4), 171–190. doi:10.1080/10508414.2015.1162636

Gray, W. R., III (2007). A Boundary Aviodance Tracking Flight Test Technique for Performance and Workload Assessment. In Society of Experimental Test Pilots (Ed.), *SETP Annual Symposium Proceedings*.

Haslbeck, A., & Bengler, K. (2016). Pilots' gaze strategies and manual control performance using occlusion as a measurement technique during a simulated manual flight task. *Cognition, Technology & Work*, *18*(3), 529–540. doi:10.1007/s10111-016-0382-2

Haslbeck, A., & Gontar, P. (2014). How pilots believe to act - and how they really do: Insights from flight simulator studies. In A. Droog (Ed.), *Aviation Psychology: Facilitating change(s). Proceedings of the 31st EAAP Conference* (pp. 317–324). Groningen, The Netherlands: European Association for Aviation Psychology.

Haslbeck, A., Gontar, P., & Schubert, E. (2012). The way pilots handle their control stick - effects shown in a flight simulator study. In C. Lemmer (Ed.): *Vol. 19. Reports of the DLR-Institute of Transportation Systems, 30th European Annual Conference on Human Decision-Making and Manual Control. Proceedings* (pp. 21–26). Braunschweig, Germany: DLR-Institut für Verkehrssystemtechnik.

Haslbeck, A., & Hoermann, H.-J. (2016). Flying the needles: Flight deck automation erodes fine-motor flying skills among airline pilots. *Human Factors*, *58*(4), 533–545. doi:10.1177/0018720816640394

Hollnagel, E. (2009). *The ETTO principle: efficiency-thoroughness trade-off; why things that go right sometimes go wrong*. Farnham, UK: Ashgate.

International Air Transport Association. (2015). *Guidance material and best practices for the implementation of upset prevention and recovery training*. Montréal. Retrieved from https://www.iata.org/whatwedo/ops-infra/itqi/Documents/gmbp_uprt_2015.06.23.pdf

Johnson, E. N., & Pritchett, A. R. (2002). Generic pilot and flight control model for use in simulation studies. In American Institute of Aeronautics and Astronautics. (Ed), *Proceedings of the Modeling and Simulation Technologies Conference and Exhibit*. Virginia, USA: American Institute of Aeronautics and Astronautics.

Joint Aviation Authorities. (2004). *Airbus A320 - A330 - A340 Cross Crew Qualification & Mixed Fleet Flying*. Hoofddorp, The Netherlands: Joint Aviation Authorities.

Kantowitz, B. H., & Casper, P. A. (1988). Human Workload in Aviation. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 157–187). San Diego, CA: Academic Press.

McClernon, C. K., & Miller, J. C. (2011). Variance as a measure of performance in an aviation context. *The International Journal of Aviation Psychology*, *21*(4), 397–412. doi:10.1080/10508414.2011.606765

Morris, T. L., & Miller, J. C. (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology*, *42*(3), 343–360. doi:10.1016/0301-0511(95)05166-X

National Transportation Safety Board. (2014). *Crash of Asiana Flight 214 Accident Report Summary: Descent Below Visual Glidepath and Impact With Seawall*. Public Meeting of June 24, 2014. Washington, DC. Retrieved from http://www.ntsb.gov/news/events/2014/asiana214/abstract.html

Niewind, I. (2011). Investigations on boundary avoidance tracking and pilot inceptor workload. *CEAS Aeronautical Journal*, *2*(1–4), 147–156. doi:10.1007/s13272-011-0037-1

Niewind, I. (2012). *Pilot gain and the workload buildup flight test technique: a closer investigation of pilot inceptor workload* (Institute Report No. IB 111-2012/74). Braunschweig, Germany.

NTSB. (2010). *Loss of Control on Approach Colgan Air, Inc.: Operating as Continental Connection Flight 3407 Bombardier DHC-8-400, N200WQ Clarence Center, New York February 12, 2009*. Aircraft Accident Report (NTSB/AAR No. 10/01). Washington. Retrieved from http://www.ntsb.gov/publictn/2010/AAR1001.pdf

Poulton, E. C. (1974). *Tracking skill and manual control*. New York, NY: Academic Press.

Rantanen, E. M., Johnson, N. R., & Talleur, D. A. (2004). *The effectiveness of personal computer aviation training device, a flight training device, and an airplane in conducting instrument proficiency checks: volume 2: objective pilot performance measures*. Final Technical Report AHFD-04-16/FAA-04-6. Oklahoma City, USA.

Ruediger, I. (2014). *Pilot Gain and the Workload Buildup Flight Test Technique* (Dissertation). Technische Universität Braunschweig, Braunschweig, Germany.

SKYbrary. (2016). *Stabilised Approach*. Retrieved from http://www.skybrary.aero/index.php/Stabilised_Approach

Stowell, R. (1996). *Emergency maneuver training: Controlling your airplane during a crisis* (1st ed.). New Jersey, NJ: Rich Stowell Consulting.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson/Allyn & Bacon.

Vadrot, R., & Aubry, C. (1994). Cross crew qualification and mixed fleet flying: The airbus familiy concept. *FAST*, *(17)*, 21–26.
Veillette, P. R. (1995). Differences in aircrew manual skills in automated and conventional flight decks. *Transportation Research Record*, (1480), 43–50.
Ziegler, P. N. (1968). Single and dual axis tracking as a function of system dynamics. *Human Factors*, *10*(3), 273–276. doi:10.1177/001872086801000310