Supplementing Haptic Feedback in Flight Envelope Protection Through Visual Display Indications

de Rooij, G.; Van Baelen, D.; Borst, C.; van Paassen, M.M.; Mulder, Max

DOI
10.2514/1.I011191

Publication date
2023

Document Version
Final published version

Published in
Journal of Aerospace Information Systems (online)

JOURNAL OF AEROSPACE INFORMATION SYSTEMS
Vol. 20, No. 6, June 2023

Supplementing Haptic Feedback in Flight Envelope Protection Through Visual Display Indications

Gijs de Rooij,* Dirk Van Baelen,† Clark Borst,‡ Marinus M. van Paassen,§ and Max Mulder¶
Delft University of Technology, 2629 HS Delft, The Netherlands

https://doi.org/10.2514/1.I011191

Haptic cues on the side stick are a promising method to reduce loss of control in-flight incidents. They can be intuitively interpreted and provide immediate support, leading to a shared control system. However, haptic interfaces are limited in providing information, and the reason for cues may not always be clear to pilots. This study presents the results of the conceptual development of visual display symbology that supports haptic feedback on the side stick in communicating flight envelope boundaries to pilots. Novel indications for the limits of airspeed, load factor, angle of attack, and angle of bank, which for the first time simultaneously indicate magnitude and direction of the haptic cues, were integrated in an Airbus primary flight display. The symbology was tested in a pilot-in-the-loop experiment with professional Airbus pilots ($N = 16$) flying several approaches in alternate law with haptic feedback. Objective results do not show clear improvements, although the time spent outside the flight envelope is slightly reduced. Subjective results indicate a preference, however, for the new display and an increased understanding of the haptic feedback. Further research is recommended to improve the interface design, remove unused indications, and test a bank scenario using current operational bank limits.

I. Introduction

LOSS of control in-flight (LOC-I) has been the primary cause of fatal commercial jet airplane accidents in the last decade [1]. There are multiple definitions of LOC-I, but a common factor is that it involves flying outside the flight envelope (FE), with the potential of making it impossible for pilots to control the aircraft [2]. Modern fly-by-wire aircraft are protected from such FE excursions, but when automation degrades or fails, these protections are reduced and pilots find themselves in a situation where it is not clear how to keep the aircraft within the envelope.

An example of such an occurrence is Air Asia flight 8501 in 2014 [3]. Due to a fault in the rudder travel limiter unit (RTLU) of the Airbus A320 and subsequent actions by the crew, the aircraft switched from normal to alternate control law, losing most of its protections and disconnecting the autopilot. The RTLU fault made the aircraft bank to 54°. Startled by this, the crew responded incorrectly, banking the aircraft to even extremer angles and eventually pulling the aircraft into an unrecovered prolonged stall. All 162 people on board perished when the aircraft crashed into the Java Sea. This and other incidents, like the Air France 447 [4], show that once protections are lost, pilots lack clear cues on their state with respect to FE boundaries and on how to return to flight within the envelope.

Previous research on haptic feedback, as a way to communicate information to human operators, has shown that haptic cues might close this information gap and decrease LOC-I incidents [5]. Nevertheless, it was found that pilots were sometimes unsure as to what triggered the haptic feedback and what corrective action to take. It was recommended that a visual representation of the haptic cuing and FE is developed to help pilots understand what the haptic feedback is telling them [6]. In combination with haptic feedback, this may assist pilots in recognizing the edges of the FE and act accordingly. Research on an unmanned aerial vehicle collision avoidance system indeed suggests an increase in user acceptance when adding visualizations to a haptic system [7].

Several research projects have investigated the design of visual displays that can show (more) FE information. In airliners the primary flight display (PFD) appears to be the preferred location to integrate such information, although some researchers proposed a separate display. A common factor in most existing solutions is the separation of output and input space into separate displays, that is, showing either the limits of the envelope (the output space) [8–12] or the limits in control inputs (the input space) that would otherwise bring the aircraft outside that envelope [12]. No previous research is known about a single aircraft display specifically integrating the limits on the input and output space together with information on associated haptic feedback.

This study builds on the foundations of the aforementioned research [5] by investigating a display design that aimed to integrate the input and output spaces, while also showing the force and direction cues of the specific haptic feedback implemented in that research. For the first time, a single display design is presented that combines all of these. The paper starts with background information on FEs and haptic feedback in Sec. II. Section III presents the display design and explains the rationale behind it. The display was tested in a human-in-the-loop flight simulator experiment involving 16 professional airline pilots, as discussed in Sec. IV, to assess the potential of said display. Section V describes the experimental results, which are then discussed in Sec. VI together with some recommendations for further research. Section VII concludes the paper.

II. Background

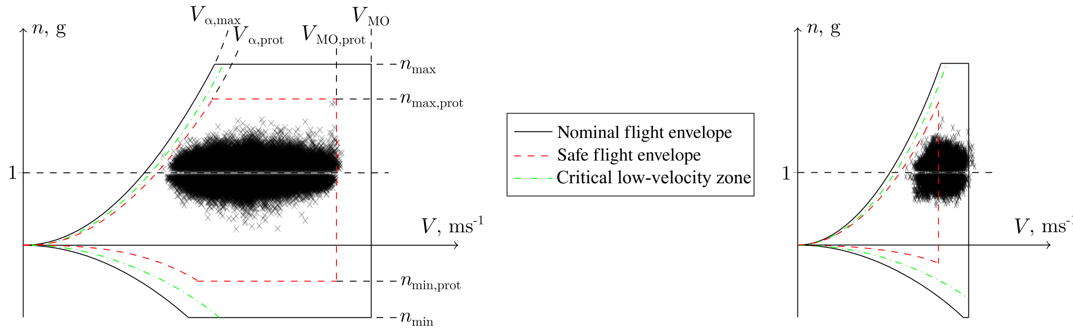
A basic understanding of the FE is required to grasp the working of the haptic feedback and any display design choices. This section provides a short introduction to these concepts, together with a couple of implementation details that are specific to this research.

A. Flight Envelope

The longitudinal performance limits of an aircraft are often captured in a FE that relates velocity V to load factor n. A common FE shape is depicted by the solid line in Fig. 1. The upper velocity limit is dictated by the maximum velocity or V_{MO} that can be attained by the aircraft, respecting aerodynamic and vibration limits. Structural limits, indicated by horizontal lines, put a minimum n_{min} and maximum n_{max} on the load

Presented as Paper 2020-0373 at the AIAA SciTech 2020 Forum, Orlando, FL, January 6–10, 2020; received 25 August 2022; revision received 20 January 2023; accepted for publication 23 February 2023; published online 31 March 2023.

*Ph.D. Student, Section Control & Simulation; g.derooij@tudelft.nl.
†Ph.D. Student, Section Control & Simulation; vanbaelend@gmail.com.
‡Assistant Professor, Section Control & Simulation; c.borst@tudelft.nl.
§Associate Professor, Section Control & Simulation; m.m.vanpaassen@tudelft.nl. Senior Member AIAA.
¶Professor, Section Control & Simulation; m.mulder@tudelft.nl. Associate Fellow AIAA.



a) Flaps UP **b) Flaps 3**

Fig. 1 Typical FEs with velocity (V) versus load factor (n) [5]. Augmented with load factor data for 10,066 A320 flights [14] for illustrative purposes. Actual envelopes depend on aircraft configuration and loading.

factor independent of airspeed. At low speeds, a quadratic relation limits the minimum velocity $V_{\alpha,\max}$. Flying below $V_{\alpha,\max}$ at a too high load factor will stall the aircraft. With extended flaps, both $V_{\alpha,\max}$ and V_{MO} decrease, leading to a much smaller FE (Fig. 1b). Airbus aircraft in addition cap the lower and upper load factor limits to 0 and $2g$, respectively, when the flaps are not up [13], but the model from this study keeps the load factor limits at -1 and $2.5g$ in order to match the haptic feedback from our previous study [5].

In that study, safety margins were added to the FE to create a so-called safe flight envelope (SFE), indicated by the red dashed line in Fig. 1. The associated protection margins were chosen such that pilots have sufficient time to steer the aircraft away from the boundaries after being alerted of leaving the SFE. The load factor margins are $0.5g$, lower speed margins vary along the envelope, and the high-speed margin is fixed at 20 kts below V_{MO}. Another margin can be distinguished near the lower velocity regime indicated by the dashed green line, and showing critically low velocity close to a stall.

The envelopes in Fig. 1 are overlaid with maximum and minimum load factors encountered in 10,066 Airbus A320 flights [14]. Note that the envelopes shown here are for illustration purposes only and do not precisely match the actual envelope corresponding to those flights. In flaps up, aircraft in general stay well away from the boundaries, with only some flights nearing the SFE. On the contrary, with the significantly smaller FE corresponding to a flaps 3 configuration, the majority seems to operate to the right of the SFE. This can be explained by the lack of the fixed 20 kts overspeed margin in real-life operations.

B. Haptic Feedback

For a full description of the haptic feedback, the readers are referred to our previous paper [5]. We will explain the basic working principles by means of Fig. 2, illustrating the haptic profile, that is, the stick deflection δ and force required F. Break-out zone δ_{br} and associated spring coefficient k_{br} give the pilot a haptic feeling of the neutral point δ_{np}. Outside this break-out zone, the spring coefficients are related to the negative (k^-) or positive (k^+) deflection of the stick. While only longitudinal haptic feedback was considered

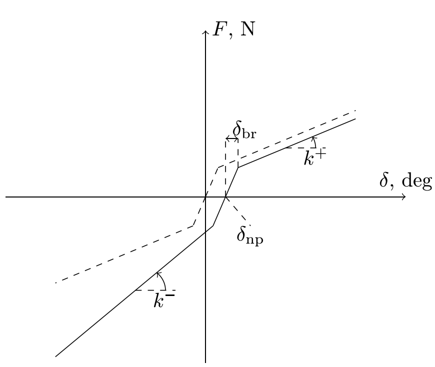


Fig. 2 Haptic side stick profile.

there, lateral feedback, based on the same principles, has since been implemented and both were used in the present research.

The system can be summarized with five haptic cues to communicate the FE to the pilot. First, when the aircraft leaves the SFE, a discrete force cue warns the pilot. Second, continuing to steer the aircraft out of the SFE results in a progressively increased stiffness, as shown by the asymmetric profile in Fig. 2. Third, when zero stick input is insufficient to return to the SFE for low velocities, the neutral point moves. Fourth, the stick shaker activates when crossing the critical lower velocity indicated on the FE in Fig. 1a by the dashed green line. Fifth, the neutral point of the stick shifts during an overspeed situation to indicate the automatic pitch up command. A more detailed visualization of these steps is given in Sec. III.C.

III. Display Design

The haptic feedback system from Sec. II.B was tested with professional airline pilots in previous research [6], yielding an overall positive evaluation but also a recommendation to investigate adding a visual display to complement the haptic cues. Indeed, combining haptic feedback with a visual display could fulfill the guideline of using multiple resources when presenting important information [15]. To address the shortcomings of existing displays, such as the lack of integration of input and output space, a new display was designed. It should show the pilots which envelope limit is triggering the haptics, where the aircraft is with respect to the (S)FE and what forces are acting on the stick. This section first elaborates on the principle behind a design that fulfills all of these requirements, and then explains the look and feel of the various new display elements.

A. Design Principle

To support the haptic system, the indications on the display have to match the forces felt through the side stick in both magnitude and direction. From the cues discussed in Sec. II.B, the discrete cue and changing stiffness can be visualized by an ordinary spring (upper part of Fig. 3) that is positioned next to the side stick. When the aircraft approaches the edge of the SFE, as discussed in Sec. II.A, the spring moves toward the stick. Upon leaving the SFE the free end of the spring—visualized by the left-most vertical line—barely touches the stick. At this point the haptic feedback gives a discrete impulse (pulse) on the stick to grab the pilot's attention. When the aircraft gets further into the protection zone, the spring is progressively compressed, its width increases, and so does the force exerted by the spring. This force acts in the direction opposite to the movement of the stick, making it harder for the pilot to maintain a stick input in that direction. If the compression is relaxed, the spring lengthens again while its thickness and force decrease. Like any spring, the force is only felt when the spring starts getting compressed. Maximum compression is reached when the two vertical lines touch each other, after which the spring coefficient does not change any further.

To ease implementation in the display, improve clarity and reduce clutter, the spring can instead be visualized in the form of a piston cylinder whose thickness is similar to the width of the spring (lower

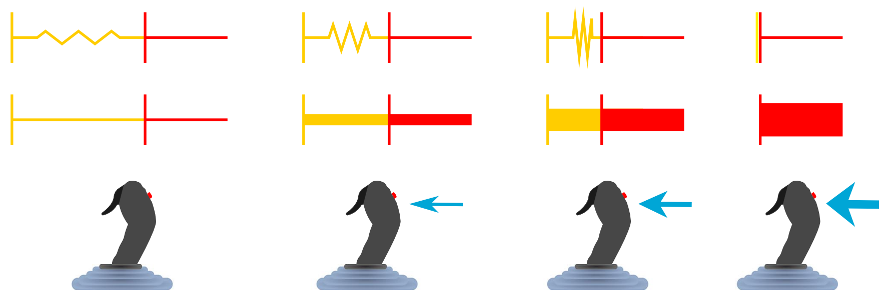


Fig. 3 Spring (top) and piston (bottom) symbols with increasing levels of compression, corresponding to an increasing haptic force on the deflected side stick (blue arrow) in the direction of the stick's neutral position.

part of Fig. 3). Apart from visualizing the "feel" from the haptics in both magnitude and direction, these indications also show the pilot in which direction he or she should provide control inputs to alleviate the required force and return the aircraft to the SFE. All of this is known to help pilots understand and consequently appreciate haptic feedback better [7]. No indication is added for the stick shaker, as this cue is a trigger to bring the pilot's attention to the low speed, rather than an actual limit; the neutral position shift is neither explicitly visualized, as it comes in combination with an increased stick stiffness and thus another display indication.

The piston analogy is used throughout the enhanced display. The symbols and colors are kept uniform over the various indications to adhere to Wickens's design principle of consistency [15]. In line with industry-recommended color coding [16], yellow is used to indicate the protection limit, beyond which the aircraft is outside the SFE but still within the FE. The actual FE limits are indicated in red.

To help pilots quickly determine what flight parameter is driving the haptic feedback on their control inputs, the various axes (bank, load factor, angle of attack (AoA), and airspeed) are displayed separately. Where possible, the new indications are placed on parts of the display that are already showing the related parameter(s) according to the proximity compatibility principle [15]. Figure 4 shows the PFD with all the FE indications in place. The various elements are discussed in greater detail below.

B. Novel Display Indications

1. Load Factor

The first addition is a load factor indicator to the left of the airspeed tape (Fig. 5). The new indicator consists of a vertical tape showing the load factor currently acting on the aircraft. Similar to the existing speed and altitude tapes, the indicator is of the inside-out style: the aircraft is fixed, and the reference scale is moving. The reference scale has tick marks every 1*g*. The FE limits are indicated by horizontal lines that attach to vertical lines running away from the fixed

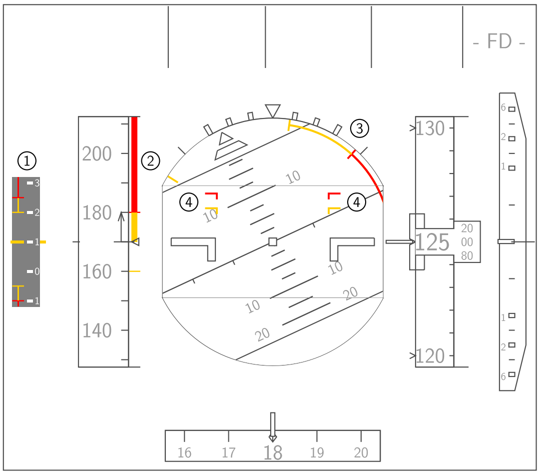


Fig. 4 Wireframe view of the Airbus A320 PFD with additional load factor indicator ① and FE limits for airspeed ②, bank angle ③, and angle of attack ④.

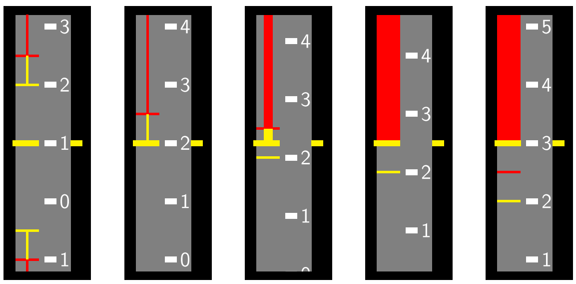


Fig. 5 Load factor indicator progressively reaching and eventually exceeding the upper limit.

reference line. The FE limit is indicated in red, while the SFE limit is shown in yellow. When the aircraft leaves the SFE, the thickness of the vertical line on the associated side increases linearly according to the piston principle. The horizontal yellow and red lines stay fixed at their positions on the moving scale to provide a quick indication of the distances to the FE boundaries. An example of an excessive load factor maneuver is shown in the sequence of Fig. 5. The big red line at the top of the rightmost figure gives a clear "pitch down" cue to the pilot. Approaching and crossing the lower limit exhibits a similar but mirrored sequence on the lower part of the scale.

2. Airspeed

The haptic system provides speed cues on the pitch axis of the side stick, because pitching up or down is an effective method to control airspeed (next to controlling the throttle). To make it clear to the pilots that the pitch cue is actually a speed cue, an indication is added to the speed tape rather than the pitch ladder (Fig. 6). For the overspeed protection, the standard overspeed barber pole at V_{MO} is replaced by a protection and maximum limit indication similar to that of the load factor. The protection is always 20 knots below the maximum speed. Once the aircraft crosses the protection limit, a gentle nose up command is encouraged by the haptics.

A similar indication on the lower side of the speed tape corresponds to the low-speed part of the FE, where the haptics will eventually encourage a nose down command. Midway between the yellow and red limit, the stick shaker will activate to alert the pilot of an impending stall.

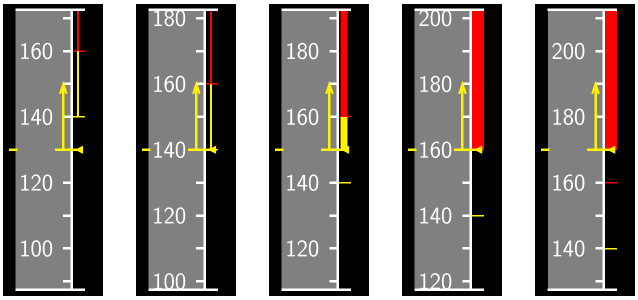


Fig. 6 Overspeed indicator progressively reaching and eventually exceeding the limit.

One potential issue with this representation is that the aircraft nose must go up for the speed to go down, and vice versa. The way the speed tape is oriented leads to indications that are not adhering to the “principle of the moving part” [17,18]. A big red line at the top of the speed tape might be interpreted as a nose down cue, while the proper thing to do is to pull the nose up. The other indications (bank, load factor, and AoA) do give cues in the correct direction. However, since the speed tape on the A320’s current PFD already has an indication for overspeed that is similar in direction to this new piston-symbol, it can be considered an acceptable design.

3. Bank Angle

For the bank angle protection, the piston-like indications are added below the bank indicator scale (Fig. 7). The limits move with the horizon—in line with the inside-out design of the PFD—while the reference aircraft symbol stays fixed. When the aircraft approaches a bank limit, this gives the pilot the sensation that the limits move toward the center of the display from the side that the aircraft is banking to. According to Roscoe’s principle of the moving part, this helps pilots interpret the direction of the limit that matches the directional cue given by the side stick [15]. In the example from Fig. 7, the pilot should roll left to lower the bank angle.

4. Angle of Attack

An indication for margin to stall AoA is added to the PFD as shown in Fig. 8. The distance from the “whisker” indications to the fixed aircraft symbol equals the margin of the current AoA to the stall AoA, similar to Boeing’s pitch limit indication (PLI) [19]. At the red whiskers, the aircraft is flying at its maximum AoA. A vertical line in the center of the display grows in width, analogous to the piston indication from the design principle. To put additional emphasis on the importance of unloading the wing by pitching down, the lower end of the piston progressively changes to an arrow as it grows wider. The indications do not rotate with bank, to ensure that the indications are always visible and always match a pitch down command. As such, the display presentation is also compatible with upset recovery techniques, where unloading must be performed first, before any bank angle corrections. The whiskers are placed besides the pitch ladder to not obstruct the ladder.

C. Typical Windshear Recovery

To illustrate the synergy between the FE, display, and haptic feedback, Fig. 9 shows the display indications during a typical windshear escape procedure side by side with the FE and haptic profiles. The series of four frames follows the actions a pilot would typically perform:

Frame 1: The windshear is triggered, indicated by a red windshear text on the PFD and a synthetic voice repeating “windshear” three times. The pilot initiates the standard windshear procedure by applying full thrust and pitching the aircraft to 17.5° of pitch [13].

Frame 2: The pilot receives a pulse on the stick’s pitch axis, as well as an increased stick-back stiffness, to alert him or her that the speed is decreasing outside the SFE. On the speed tape, this is shown by the current speed protruding into the yellow part of the low-speed piston. At the same time, the load factor indication shows that the aircraft is above the safe load factor limit for the current airspeed. And finally, the AoA indication on the pitch scale starts growing in width, as the AoA approaches its maximum.

Frame 3: When the aircraft continues the deceleration, the stick shaker is enabled as an additional low velocity warning. The aircraft is now very close to a stall, and the big red arrow on the pitch ladder of the PFD urges the pilot to push the nose down. This is also felt in the stick by an increase in stiffness on the nose-up side. Additionally, the stick’s neutral point is shifted forward at this point to help the pilot lower the nose.

Frame 4: After the initial windshear recovery, the aircraft is now accelerating. When approaching the maximum velocity limit, as shown here, a pulse warns the pilot of an imminent excursion and the stick moves backward to help the pilot bleed off airspeed. The spring stiffness of the stick is increased to inform the pilot of the distance to the ultimate FE limit, as visualized by the widening of the piston on the speed tape.

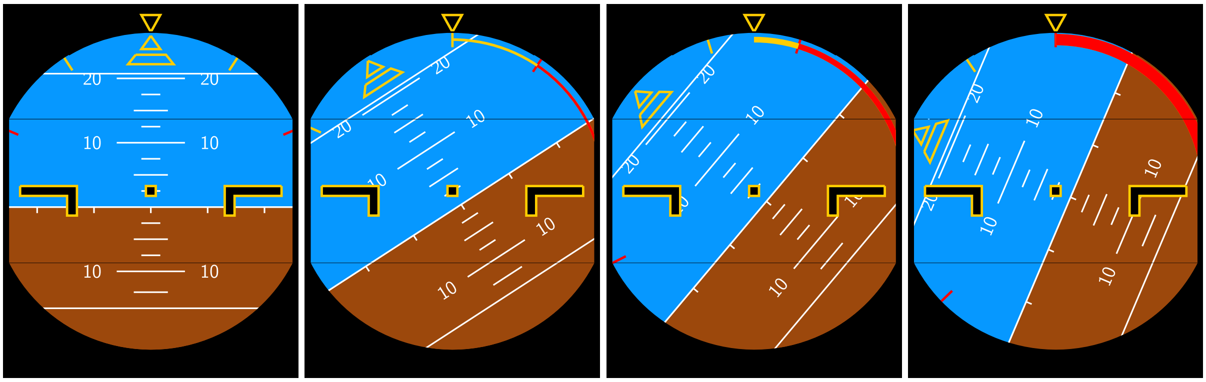


Fig. 7 Bank angle indicator progressively reaching the limit.

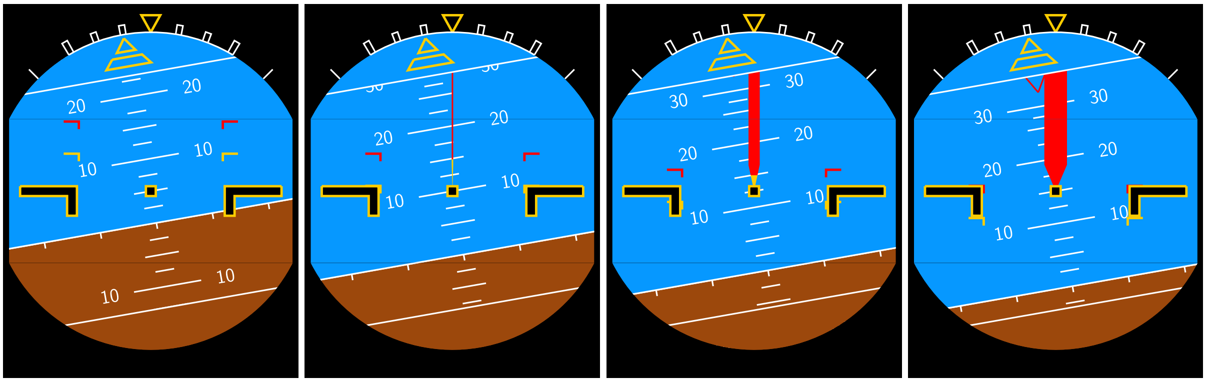


Fig. 8 AoA indicators relative to the fixed aircraft symbol progressively reaching the AoA limit.

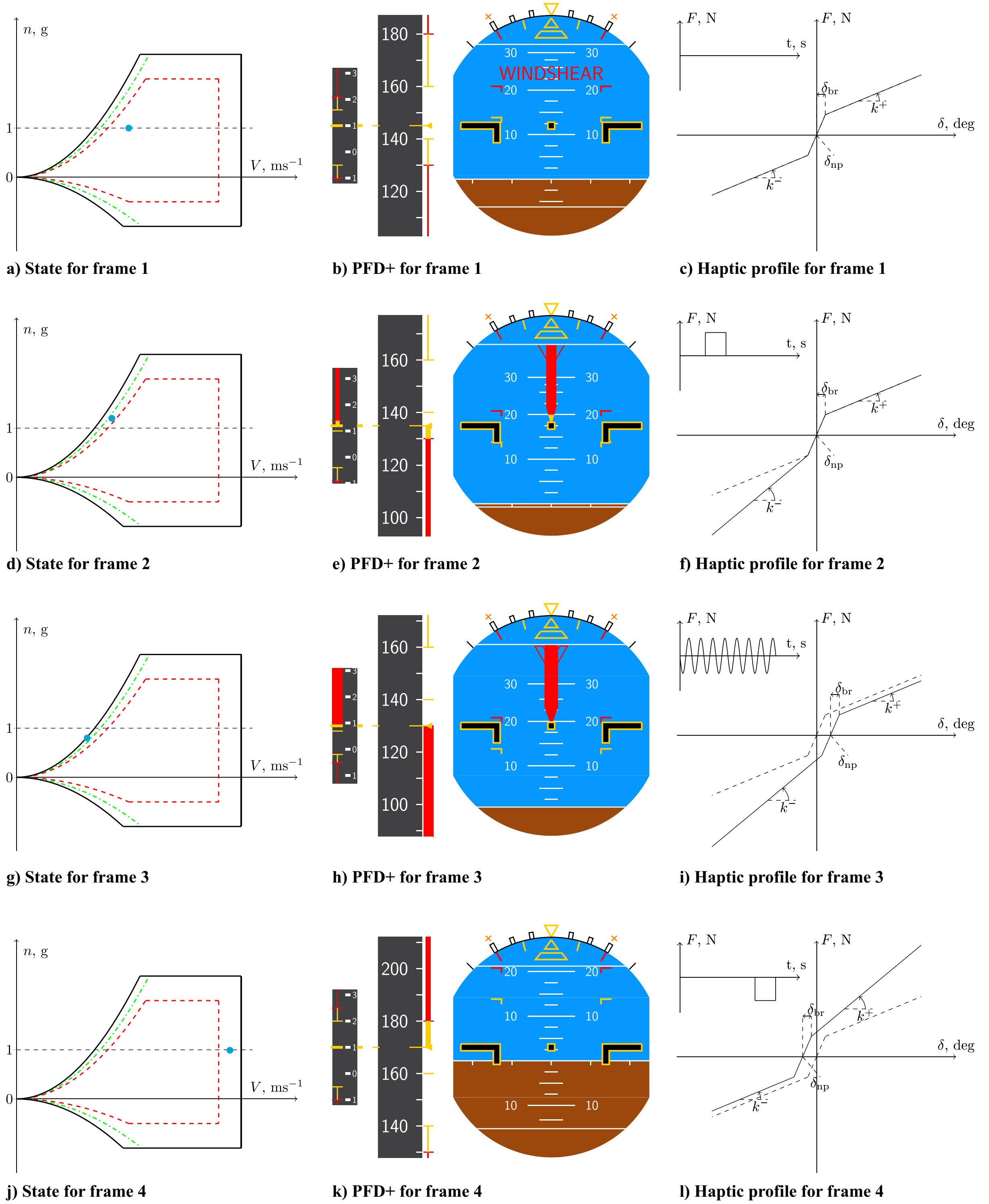


Fig. 9 Typical windshear recovery procedure. The left column shows the FE, the center column shows an excerpt of the PFD, and the right column shows the associated haptic pitch profile.

IV. Method

Since pilots are expected to interact with the display, its design was tested in a human-in-the-loop simulator experiment. The goals of the experiment were to evaluate the interaction of pilots with the display, to investigate possible changes in their control strategy and to study whether it changes their opinion of the haptic feedback.

A. Participants

Fifteen male professional Airbus pilots from four airlines and one male Airbus A320 synthetic flight instructor (SFI) voluntarily participated in the experiment. All participants provided written consent before the experiment. They were divided over two groups (A and B, experiencing a different display order), while balancing for age, experience, and aircraft types as much as possible, as shown in Table 1. Two pilots in each group had previously participated in our haptic feedback evaluation [6]. It is worth noting that the second officers—while not certified to operate the aircraft below 20,000 ft—did receive complete flight training and all had first-officer Boeing experience from previous positions. Of the pilots, 14 had experienced windshear on a real aircraft, of which 9 in an Airbus. All pilots had received upset recovery and prevention training (UPRT) and had experienced alternate law in simulator training.

Table 1 Characteristics of the participants

	Group, M (SD)			
Parameter	A	B	Δ	p
Age, years	42.5 (13.6)	42.1 (8.0)	0.4	0.948
Airbus flight hours	3,575 (3,288)	4,335 (2,306)	760	0.601
Flight hours	8,475 (5,602)	11,092 (4,841)	2,617	0.334
	Group, n			
	A	B		
Aircraft type				
A320	3	3		
A330	5	5		
Rank				
Captains	3	7		
First officers	2	0		
Second officers	2	1		
SFI	1	0		

B. Apparatus

The experiment took place in the Simulation, Motion, and Navigation (SIMONA) Research Simulator at Delft University of Technology. The simulator's exterior and interior are shown in Figs. 10 and 11, respectively. SIMONA is a six-degree-of-freedom motion simulator with a full-fledged cockpit shell. The interior can be configured to resemble any modern glass cockpit transport aircraft. For this particular experiment, the motion system was not used.

An electrically controlled Moog FCS Ecol-8000 side stick with force feedback capabilities as described in [5] was located on the right-hand side of the pilot, who was seated in the right seat. The pedals were not used. A Boeing 777 pedestal with throttle quadrant and flaps lever, and a Boeing 737 Mode Control Panel (Flight Control Unit in Airbus terminology) complemented the interior. The outside visuals were provided by FlightGear** and showed the airport infrastructure, terrain, and important buildings at the airport. A proprietary A320 flight dynamics model including control laws from the German Aerospace Center (DLR) was used as the simulated aircraft [20]. Since the model did not include a landing gear, all flights were automatically stopped upon reaching 50 ft above ground level (AGL).

Fig. 10 Exterior of SIMONA at TU Delft.

Fig. 11 Interior of SIMONA at TU Delft.

The entire simulation was run using the Delft University Environment for Communication and Activation (DUECA) software. DUECA is a framework written in C++ allowing for easy real-time distributed simulations [21]. The PFD and Navigation Display (ND) were drawn using OpenGL (Fig. 11) and very closely resembled the real Airbus displays.

C. Procedure

All participants engaged in the procedure outlined in Table 2. Group A first used the original PFD and then the new PFD, denoted as PFD+, while the order was reversed for group B. The complete experiment took approximately 5 h per pilot.

1) *Briefing:* At the start of the day, the pilots received a short introduction and were asked to fill in a pre-experiment questionnaire on their flying and previous research experience.
Inside the simulator, the pilots were seated in the right seat. After a safety briefing, the various controls and standard displays were explained, as some of these did not perfectly resemble their Airbus counterparts. For instructional purposes, the original PFD was temporarily moved to the left screen—the normal location of the ND—while the right screen showed the haptic profile and the FE.
Without the model in the loop, hence by the simulator operator changing the state of the aircraft directly, all haptic cues were explained. The pilots were asked to close their eyes while experiencing all cues once again to check whether they had understood the explanation of the various cues. Next, the PFD+ was shown, and all cues were thoroughly presented and experienced once again.

2) *Familiarization:* For familiarization with the model and controls, a right-hand circuit to runway 36L at Amsterdam Airport Schiphol (EHAM) was flown twice with the baseline PFD. Note that this is a nonstandard approach, and therefore no instrument landing system or precision approach path indicator was provided.
Next, the pilots performed the following exercises over the North Sea to experience the haptic cues:

i) Pilot induced stall by maintaining altitude with idle throttle.
ii) Pilot reached overspeed by full throttle and pitching down.
iii) Pilot performed nose-dive followed by a strong back stick input to reach the high-g region.
iv) Pilot performed rolling to the left and right.
v) Pilot performed pitching up as far as possible with closed eyes, while keeping the aircraft at the onset of stall.

Upon completion of these exercises, the same circuit as before was flown once more, this time with the PFD+. The haptics were left unchanged with respect to the previous circuit. After one circuit, the same maneuvers were flown over the North Sea as before, apart from the closed-eyes exercise.

3) *Training:* The training phase was set up to more closely resemble operational flights and prepared the pilot for the actual measurement flights. Four approaches were flown toward runway 16R of Seattle Airport (KSEA), for which the layout is shown in Fig. 12. The baseline PFD was used on the first two approaches, while the novel PFD+ was present on the latter two approaches. The conditions per flight are shown in Table 3. After each flight, the pilots were asked to fill in a questionnaire, identical to those used in the measurement runs.

4) *Measurement runs:* For the measurement runs, the pilots were divided into two groups. Group A flew the first set of measurements with the old PFD, followed by the PFD+, and group B vice versa.
At the start of each block of six measurement runs, a go-around

**Open source flight simulator available at https://flightgear.org.

Table 2 Experiment procedure

Group	45 min	40 min	30 min	30 min	60 min	60 min	20 min
A B	Briefing	Familiarization	Training flights	Lunch	PFD flights PFD+ flights	PFD+ flights PFD flights	Debriefing

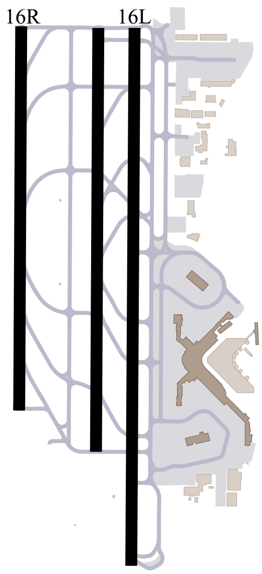


Fig. 12 Airport diagram of KSEA [22].

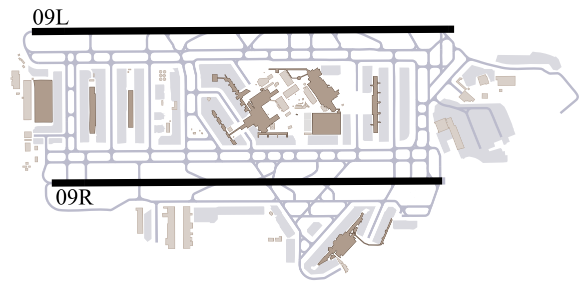


Fig. 13 Airport diagram of EGLL [22].

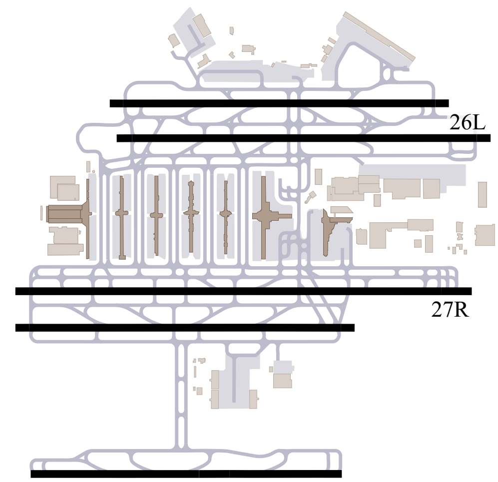


Fig. 14 Airport diagram of KATL [22].

Table 3 Training phase flights

Run	Airport	Scenario	Display
1	KSEA	Windshear	PFD
2	KSEA	Runway sidestep	PFD
3	KSEA	Runway sidestep	PFD+
4	KSEA	Windshear	PFD+

scenario was flown into KSEA with the PFD variant corresponding to that block of flights. This "refreshment" run was used to give the participants a chance to refamiliarize themselves with the flight model, haptic feedback, and, when applicable, PFD+ after a (lunch) break.

Then the six measurement runs were flown. Each ended with a questionnaire, followed by the presentation of a score. The airports and scenarios for these flights were assigned according to a balanced Latin square distribution. After the six flights, the pilots were asked to fill in another questionnaire about the complete set of six flights.

5) *Debriefing:* At the end of the experiment, the pilots received one more questionnaire about their overall experience throughout the day as well as the realism of the simulator. Once the questionnaire had been filled in, the pilots were debriefed. The research question was revealed, and any remaining questions that could not be answered before in order to not influence the experiment outcome were discussed at this point.

D. Secondary Task

Apart from flying the approach, the pilots were given a secondary task in the form of Air Traffic Control (ATC) calls that they had to reply to. Each pilot's call sign reflected the company that the pilot was employed at: "{Company} 107." To ensure that the pilots had to pay attention to the ATC calls, two other aircraft from the same company were introduced with flight numbers 685 and 713. ATC could ask to "report heading," "report speed," and "report altitude." Random realizations were made for each condition to ensure that all pilots received the same ATC commands in the same condition. A call sign and command were selected from a uniform distribution. These were then triggered at a delay after the previous command, determined by a normal distribution ($\mu = 20$ and $\sigma = 2.5$ s). The texts were read out loud by a female American-English accent from the Festival[††] text-to-speech generation library.

E. Independent Variables

Three independent variables were used in the experiment: the airport (two levels), the scenario (three levels), and the display (two levels). In total there were therefore 12 different conditions. To reduce variance in the data, all pilots experienced the same conditions. To mitigate order effects, a randomized balanced Latin square was used. The airport and scenario were varied constantly, while the display variant was fixed during a series of six consecutive flights to prevent pilots from having to re-adapt to the available cues all the time.

1. Airport

Approaches were varied between runway 26L at Hartsfield-Jackson Atlanta International Airport (KATL) and runway 09L at London Heathrow (EGLL). Both airports have runways on either side of the terminals, with comparable spacing (KATL: 1340 m; EGLL: 1420 m) and more or less adjacent thresholds. Their layouts are shown in Figs. 13 and 14. An instrument landing system (ILS) was available on the approach runway, with corresponding indications on the PFD. The pilots were provided with approach charts including a schematic of the runway layouts.

[††]University of Edinburgh, available at http://www.cstr.ed.ac.uk/projects/festival/.

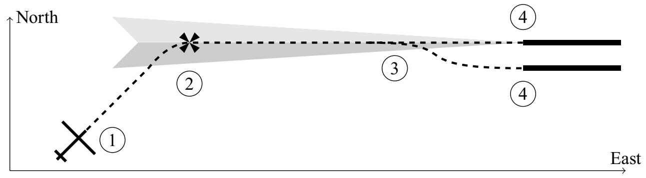


Fig. 15 Flight plan for EGLL (not to scale). Start ①, localizer interception at FAF ②, scenario triggering point ③, and end of flight ④.

Each flight started circa 12 NM from the airport in trimmed flaps up condition at 215 kts and an intercept heading of circa 45°, toward the final approach fix (FAF) on the localizer. At EGLL the starting position was circa 3 NM right of the localizer, while it was circa 4 NM left of the localizer at KATL. Figure 15 shows a typical trajectory toward EGLL.

2. Scenario

The pilots were subjected to three scenarios. These were automatically triggered upon descending through a predetermined altitude given in Table 4. In all scenarios, a stable and variable wind was introduced according to the distribution used in [6]. This wind was identical for all pilots.

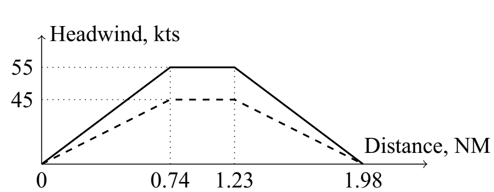
1) *Windshear:* The windshear was implemented using the standard takeoff wind model from the FAA [23] with wind components as shown in Fig. 16. An approach windshear model was not used because it was found not to ensure that the aircraft would fly near the limits of the FE. In the training runs, the strength of the windshear was reduced while keeping the same distances, as indicated by the dashed lines in Fig. 16. In accordance with the Flight Crew Operating Manual (FCOM), the pilots were told to not change the configuration of the aircraft, apply full thrust, pitch up to an initial attitude of 17.5°, and adjust pitch as necessary to control altitude loss [13]. The lack of Speed Reference System (SRS) pitch guidance upon windshear encounter was explicitly briefed. When out of the shear, pilots were asked to climb to the missed approach altitude at which the simulation was halted.

2) *Runway sidestep:* ATC would make either of the following calls, depending on the airport: "{Company} 107, sidestep right to runway 09 right, cleared to land" (EGLL) or "{Company} 107, sidestep left to runway 27 right, cleared to land" (KATL). Pilots were briefed to try to line up with the new runway as quickly as possible without using extreme bank angles.

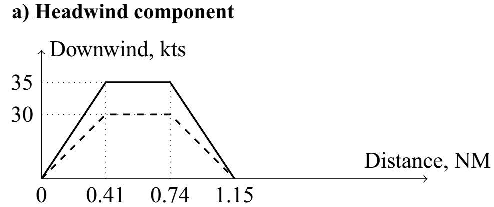
3) *Go-around:* When ATC would make the following call "{Company} 107, go-around," pilots were supposed to climb to the missed approach altitude with a climb rate of 2000 ft/min.

Table 4 Scenario triggering altitudes

		Triggering altitude			
		Above sea level, ft		Above ground level, ft	
Scenario	Airport	PFD	PFD+	PFD	PFD+
Windshear	EGLL	1500	1500	1420	1420
	KATL	2500	2500	1475	1475
	KSEA	1700	1650	1270	1220
Sidestep	EGLL	1200	1200	1120	1120
	KATL	2100	2100	1075	1075
	KSEA	1500	1500	1070	1070
Go-around	EGLL	800	1000	720	920
	KATL	1700	1900	675	875
	KSEA	1200	1300	770	870



a) Headwind component



b) Downwind component

Fig. 16 Windshear model, based on [23]. The dashed profile was used in the training runs.

3. Display

Two variants of the PFD were used in the experiment (Fig. 17): the original PFD was a replica of the PFD on the real A320, while the new PFD+ had several new indications, as discussed in Sec. III. The A320-like ND was the same throughout the experiment and always showed the final approach fix and runway threshold as waypoints (Fig. 17c). This display also showed the current throttle and flap settings to compensate for the absence of their normal indicators in the simulator.

F. Control Variables

The aircraft model and haptic feedback settings were the same in all flights. The aircraft had a total mass of 64,841 kg and was in clean configuration at the start of each flight. All flights took place in alternate law. In terms of haptic feedback, the protection and maximum limits in roll were set to 15 and 30°, respectively, on all flights. These are considerably smaller than the 33 and 67° used by Airbus [13] and have been chosen to ensure that pilots would actually encounter the (artificial) limits, as pilots do not bank beyond circa 30° in normal operation. To ease recognition of these adjusted limits, the crosses on the PFD's bank scale that normally indicate the limit at 67° were moved to 30° for the experiment. Pilots were briefed beforehand on these stricter limits, but also asked to fly like they would normally do.

G. Dependent Measures

Objective and subjective measures are used to assess the display in terms of performance, safety, and pilot appreciation.

1. Objective Measures

Objective data from the simulator were automatically logged at a rate of 100 Hz. Some of the objective measures were afterward computed from this data or handwritten notes on the secondary task.

1) *Control activity:* It is root mean square of the stick deflection angle in degrees.

2) *Margins to FE limits:* Both the FE limits and aircraft states were measured in terms of airspeed, AoA, load factor, and roll angle. The FE margin was computed off-line.

3) *Performance scores*, dependent on the scenario, were used for two reasons. First and foremost to assess whether the pilot's performance changed in the experiment and second to communicate to the pilots in order to encourage them to improve themselves throughout the experiment. The scores were defined as follows:

i) *Windshear:* Total altitude loss in feet from start of windshear till the lowest point during recovery

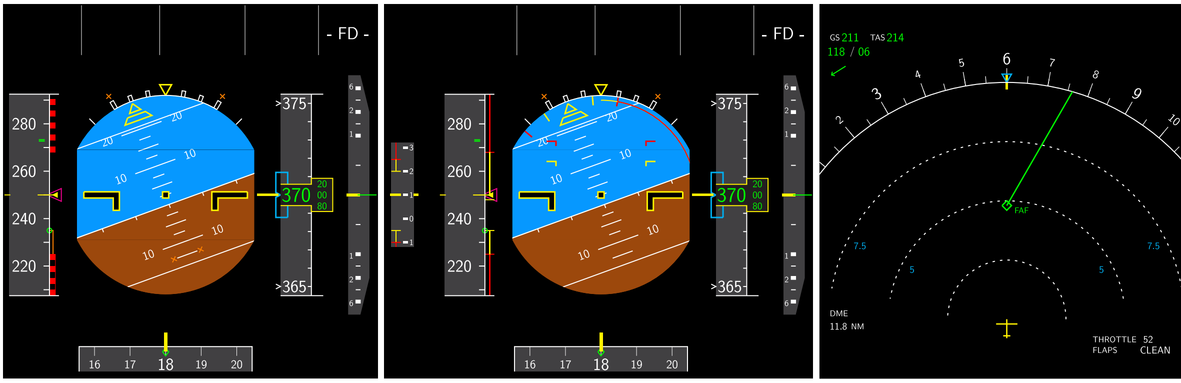
ii) *Sidestep:* The smallest distance in nautical mile to the threshold of the new runway at which the aircraft was more than 300 ft offset to either side of the localizer of that runway

iii) *Go-around:* Ratio of time during climb at which vertical speed was between 1500 and 2500 ft/min, measured from 100 ft above the trigger altitude till 100 ft below the missed approach altitude

iv) *Workload measured with a secondary task:* ratio of correct responses to ATC requests

2. Subjective Measures

Subjective measures were collected through questionnaires at various times throughout the experiment.



a) PFD b) PFD+ c) ND

Fig. 17 Displays used in the experiment.

1) After each flight:
 i) Workload through a Rating Scale Mental Effort (RSME) questionnaire [24]
 ii) Situation awareness through two questions on a linear scale ranging from "Never" (0) to "Always" (100):
  a) Did you have the feeling you were in control?
  b) Did you have the feeling you missed essential information?
 iii) Usefulness of each haptic axis (pitch and roll) and—when flying with the PFD+—each new display element in helping the pilot to stay within the limits of the FE through a five-point Likert scale question per item labeled as *not at all, slightly, moderately, very*, and *extremely*
2) After both consecutive sets of six flights:
 i) System acceptance through Van der Laan rating [25] and Modified Cooper–Harper rating [26]
 ii) Five-point Likert scale questions on three statements, with labels at the minimum (*disagree*), middle (*neutral*), and maximum (*agree*)
 iii) Questions on usefulness of individual haptic and display properties in helping the pilot to stay within the FE limits: five-point Likert scale labeled as *not at all, slightly, moderately, very.* and *extremely*
 iv) Open question on what haptic cue(s) and/or display element(s) to add to the system, if any
 v) Open question on what haptic cue(s) and/or display element(s) to remove from the system, if any
3) At the end of the experiment:
 i) Question on the pilot's display preference (PFD or PFD+) in combination with the haptic system
 ii) Five-point Likert scale statements on the haptics, display, and experiment with a minimum (*disagree*), middle (*neutral*), and maximum (*agree*) label
 iii) Five-point Likert scale question on the safety effect of the system, with a minimum (*unsafer*), middle (*unchanged*), and maximum (*safer*) label
 iv) Five-point Likert scale questions on the realism of various simulation aspects with a minimum (*unrealistic*), middle (*acceptable*), and maximum (*perfect*) label

Apart from the questionnaires, pilots were actively encouraged to verbally communicate any questions, remarks, and thoughts throughout the day. Since all pilots were native Dutch, all questionnaires and instructions were in Dutch.

H. Hypotheses

Based on the dependent measures, the following hypotheses are formulated:

H1 *Workload:* Workload in terms of control activity is expected to be lower with the PFD+ compared to the original display, since the pilot can anticipate the limits. With a lower workload for the primary task, secondary task performance is expected to increase.

H2 *Performance:* In a similar fashion it is also predicted that the addition of a visual display will improve the overall performance of pilots flying with haptic feedback.

H3 *Safety:* Safety metrics are often expected to follow risk homeostasis theory (a tradeoff between performance and perceived level of risk) [27]. However, here it is assumed that pilots consider the edge of the SFE as the maximum allowable risk. It is therefore hypothesized that the margins to the ultimate FE limits will be larger when flying with the PFD+. Additionally, pilots can anticipate the limits in contrast to the haptic feedback.

H4 *Pilot appreciation:* On a subjective level, pilots are expected to show greater appreciation for haptic feedback when combined with the PFD+ as the display should help them understand the haptic cues that they receive, one of the issues raised by pilots in the previous haptic system evaluation.

H5 *Indicator usefulness:* It is expected that the load factor display brings the least improvement compared to the old display, as the respective limits are mostly encountered in combination with other limits. The AoA indication is expected to be most useful as it provides critical information that is currently not directly communicated to the pilot.

V. Results

Several events warranted the selection of data, as some flights could not be used for the main analysis. Section V.A elaborates on this selection. The results are then split into objective results as shown in Sec. V.B and subjective results in Sec. V.C that stem from the questionnaires. Because some of the data did not meet the requirements for an ANAVO (i.e., the data did not follow a normal distribution), nonparametric tests were performed. Since both participant groups experienced all conditions (resulting in paired samples), Wilcoxon signed-rank tests with a 95% confidence interval are used unless explicitly stated otherwise.

A. Data Exclusion

Sixteen pilots participated in the experiment, each flying 12 measurement conditions. Some flights in which a simulator hiccup, before reaching the scenario trigger point, prevented proper execution were restarted. Two pilots from group B crashed their aircraft once during the measurement flights by not recovering from a stall upon windshear occurrence. One pilot crashed on the first measured windshear, while the other crashed on his second windshear. Both were flying with the PFD+ when they crashed, and had already experienced two successful windshears in the training flights. The first pilot indicated after the flight that he did not follow the procedure from the FCOM, but relied on the AoA indication on the PFD+. The other pilot did not provide an explanation, but showed similar behavior. Those flights have been started over without telling the pilots that they would encounter the same condition again. One other PFD+ flight was restarted when the pilot from group A entered a stall while turning to final, before reaching the scenario trigger point. According to his own analysis, he lost his concentration. The crashed and canceled flights are excluded from the results, unless explicitly mentioned.

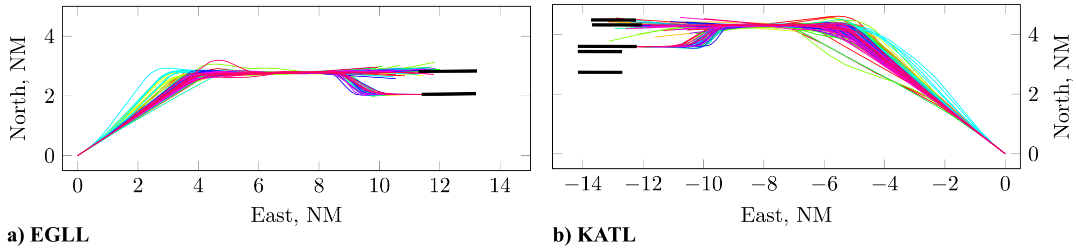


Fig. 18 Flight tracks of all flights combined, colored per pilot.

B. Objective Results

All flown tracks for both airports are shown in Fig. 18. The freedom of the pilots to choose their flight path is clearly visible. Some pilots, regardless of display variant, steered away from the localizer to give themselves a smaller intercept angle, while other pilots steered toward the localizer to overfly the FAF while lined up with the runway. Furthermore, in the go-around and windshear scenarios, many pilots did not maintain runway heading even though that was instructed. Pilots also utilized various flap extension strategies, leading to different approaches in terms of airspeed and corresponding FE limits.

This freedom complicates the comparison of flights. Each flight is therefore cut into two sections for the analysis, based on the following criteria:

Approach: From the start of the flight until the triggering of the scenario, performed in every flight
Windshear: From the onset of the windshear until the aircraft is stable at the missed approach altitude
Runway sidestep: From sending the command to the text-to-speech generator till reaching 50 ft AGL
Go-around: From sending the command to the text-to-speech generator till stable at missed approach altitude

Looking at all the other variables in the data, three more points should be raised. First, not all pilots managed to fly the approach speed of 140 kt when the windshear was triggered. Two pilots from group A have much higher approach speeds than the other pilots, irrespective of the display variant (Fig. 19), which generally corresponds to a smaller loss of altitude. In both groups, the airspeed is (slightly) lower in the second series of flights. Second, two flights stand out with a very high AoA of up to 29°. The pilots of both flights were flying relatively slow and provided full back stick upon encountering the shear. One of the pilots explained that he inadvertently thought he was flying in normal law, where the aircraft would automatically limit the stick inputs. Third, flap extension time is different per pilot, leading to different performance during the initial approach.

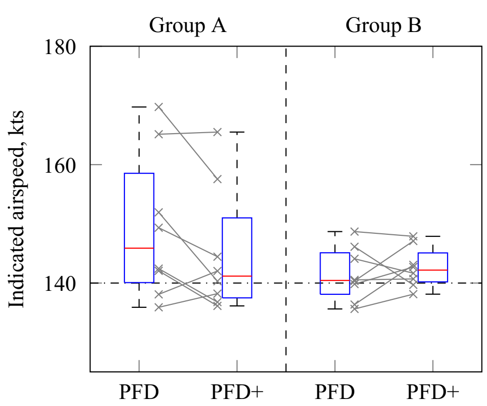


Fig. 19 Mean indicated airspeed per pilot at start of windshear (V_{APP} = 140 kt).

1. Typical Data

Figure 20 shows data for all of the protected variables on a typical flight with windshear scenario. The flap adjustments are clearly reflected in the maximum speed limits, as well as in the maximum permitted AoA. When turning onto the localizer, the pilot exceeded the 15° roll limit activating the haptic feedback on the roll axis. During the windshear the pilot was in the AoA protection zone for circa 5 s and exceeded the maximum AoA limit very briefly. Finally, a small airspeed violation can be seen on the climb out to the missed approach altitude when the pilot did not retract the flaps upon acceleration.

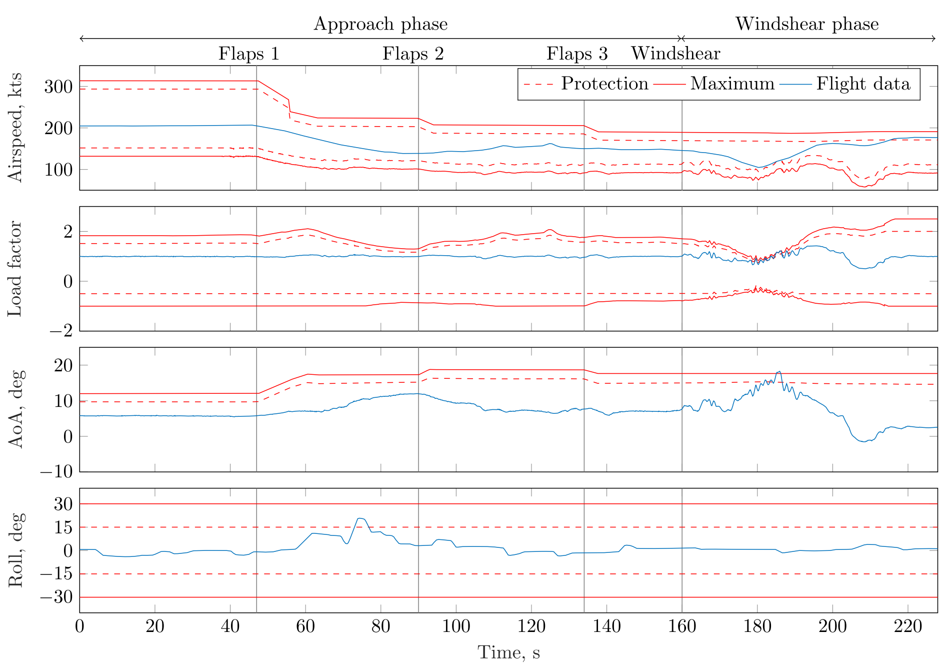


Fig. 20 Typical flight data: a windshear scenario at EGLL with the PFD+.

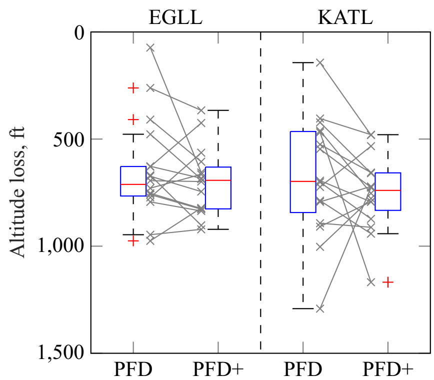


Fig. 21 Mean windshear altitude loss per pilot.

2. Performance

Overall there seems to be little effect of the display on the performance scores (see Sec. IV.G for their definition), but there are some differences between the two airports. Especially in the windshear scenario at KATL, the PFD shows a much larger spread than the PFD+ (Fig. 21), which is not observed at EGLL. The other scenarios only showed marginal effects and thus their data are not visualized here for brevity. At EGLL the PFD+ leads to a slightly lower average sidestep score (−0.05 NM), while at KATL it is higher for the PFD+ (0.1 NM). Finally the go-around also shows a small difference, with five scores below 50% for EGLL and only two of such scores for KATL. On average, go-around performance with the PFD+ at KATL was 95% versus 98–99% for the other three airport–display combinations. Wilcoxon signed-ranks tests show no significant differences for any of the performance scores: windshear at EGLL ($Z = -0.724$, $p = 0.469$) and KATL ($Z = -1.293$, $p = 0.196$), sidestep at EGLL ($Z = -1.028$, $p = 0.304$) and KATL ($Z = -0.159$, $p = 0.874$), and finally the go-around at EGLL ($Z = -0.035$, $p = 0.972$) and KATL ($Z = -0.175$, $p = 0.861$).

3. Secondary Task

Combining the flights of all pilots, there were 734 ATC calls that required a reply. Just 22 of those were not or incorrectly answered. Further analysis shows that the vast majority of ATC requests that were missed occurred while the aural windshear or stall warnings were active, or when the pilot was already transmitting a message, and not the result of workload differences. The ratio of correct replies is therefore not a useful workload measure in this experiment.

4. Time Outside the Safe Flight Envelope

The mean time spent outside the various limits of the SFE is shown in Fig. 22, where only flight phases are shown for which there was more than one excursion in the entire experiment. Roll protection limit excursions ($\phi > 15°$) mostly occurred during the localizer intercept and in the runway sidestep. Only during one windshear the roll protection was very briefly activated, while it was never activated in the go-around phase. Figure 22a shows that in both approach and sidestep the excursions were slightly shorter with the new display. A Wilcoxon's signed-rank test shows that the change in approach is significant ($Z = -3.206$, $p < 0.001$) while in the sidestep it is not ($Z = -1.034$, $p = 0.301$). For the maximum roll limit ($\phi > 30°$), there were too few violations to run a similar analysis.

Speed excursions were primarily seen during windshear and approach (Fig. 22b). In approach, these excursions were generally caused by a decreasing maximum speed upon flap extension. When climbing out of the windshear, flaps were often retracted too late, while the airspeed increased rapidly. In windshear, pilots seem to spend less time in the high-speed protection with the PFD+, but this decrease is not significant ($Z = -1.619$, $p = 0.105$). A similar, but significant, effect is seen during approach ($Z = -2.521$, $p = 0.012$). In the sidestep and go-around, there were too few overspeed occurrences for any statistical analysis. The maximum speed was only exceeded once, during a windshear with the original PFD.

As expected, the AoA limits are almost exclusively exceeded during the windshear. Figure 22c shows the time spent above the protection limit. Only one pilot never exceeded the AoA protection limit. There was a small but not significant decrease in time with the PFD+ ($Z = -0.795$, $p = 0.427$).

5. Control Activity

The root-mean-square (RMS) control deflections of the side stick are given in Figs. 23 and 24 for pitch and roll, respectively. Control activity is highest in the pitch axis during the windshear scenario. In the roll axis, most control activity is seen during the sidestep and to a lesser extent during the approach phase. There are no significant differences between the two displays, although pitch control activity appears slightly higher in windshear with the PFD+ ($Z = -0.879$, $p = 0.379$), while roll control activity seems slightly lower in the sidestep ($Z = -0.465$, $p = 0.642$).

C. Subjective Results

Subjective results were collected through a series of questionnaires. The results are discussed per questionnaire, starting with the one that was presented after each single flight, followed by the questionnaire that wrapped up a series of six flights with a single display configuration, and finally the questionnaire that was posed at the end of the experiment.

1. Postrun Questionnaire

A short questionnaire after each single run allows seeing how the display and haptics are experienced in the three scenarios. Figure 25 shows the answers to the question, “Did you have the feeling you missed any essential information?” Wilcoxon signed-ranks tests indicate that the display had a significant effect on both the lack of information in the windshear scenario ($Z = -2.691$, $p = 0.007$) and sidestep scenario ($Z = -2.121$, $p = 0.034$). The go-around scenario showed no significant results ($Z = -0.756$, $p = 0.450$). In the windshear eleven of the 16 pilots indicated less lack of essential information in the presence of the PFD+ and for three pilots the display

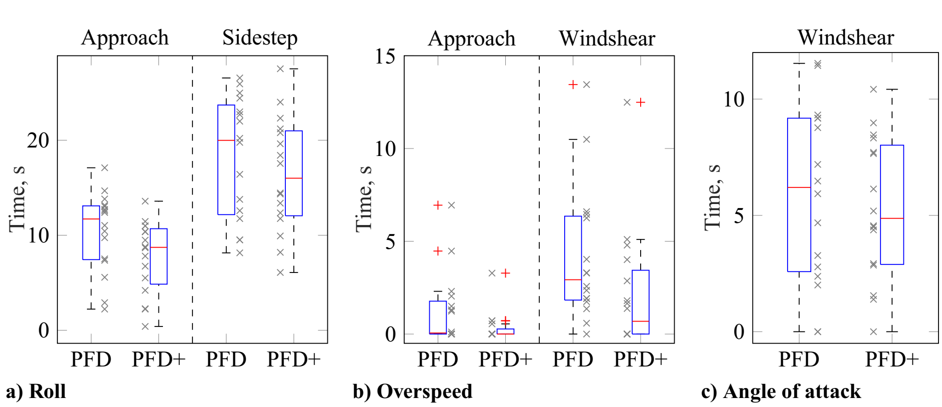


Fig. 22 Mean times in protection per pilot. Flight phases in which one or zero excursions into the protection limits were registered are not shown.

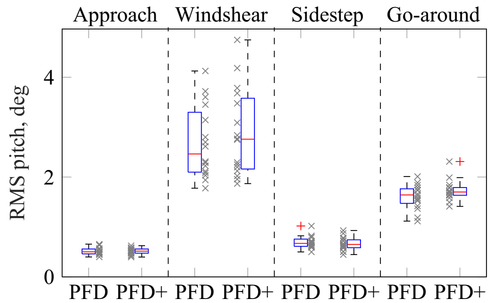


Fig. 23 Mean RMS pitch input per pilot.

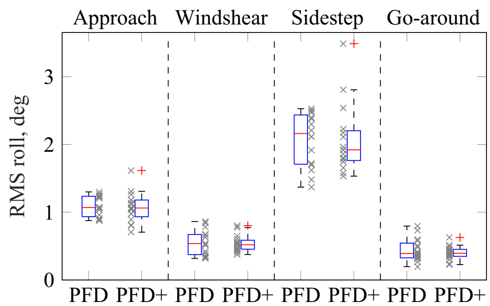


Fig. 24 Mean RMS roll input per pilot.

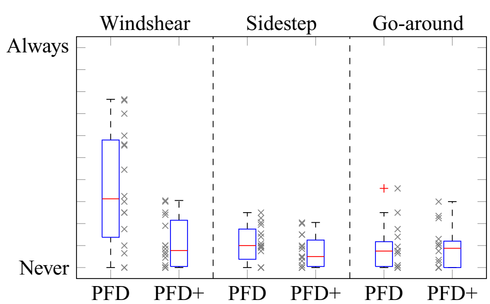


Fig. 25 Postrun question: Did you have the feeling you missed any essential information?

version did not make any difference. Particularly the AoA indication was said to be missed on the original PFD.

No significant difference between displays is observed for any of the scenarios in the control metric regarding the question, "Did you have the feeling you were in control?" (Fig. 26). During the windshear, pilots feel slightly more in control with the new display, in correspondence with the indicated lack of information. Ten pilots indicated an improvement with the PFD+ in windshear, five pilots a decrease, and one pilot was indifferent to the display variant. Overall, most pilots had the feeling that they were less in control in the windshear scenario than in the other scenarios.

In terms of subjective workload, the RSME scores, averaged over the two flights per scenario, show that the pilots perceived the highest

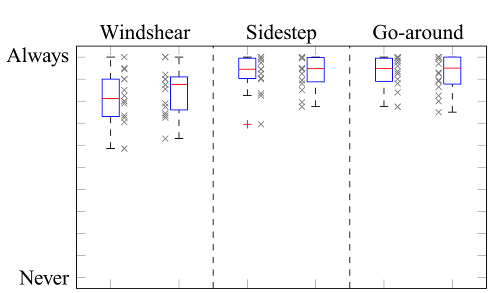


Fig. 26 Postrun question: Did you have the feeling you were in control?

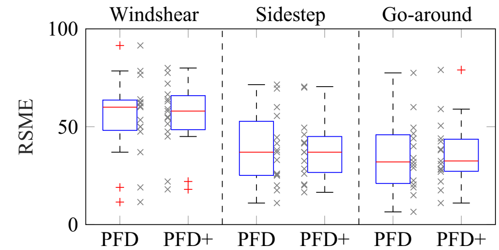


Fig. 27 Mean RSME scores per pilot.

effort in the windshear scenario (Fig. 27). The effort in the sidestep scenario is less and comparable to that in the go-around scenario. There are no statistically significant differences observed between the two displays.

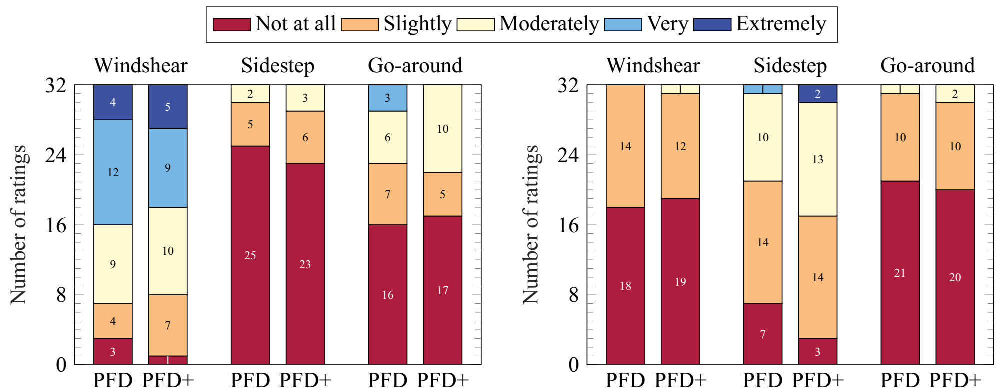
When asked about the usefulness of the haptic feedback on the pitch and roll axis of the stick, it can be seen that the pilots considered the haptic pitch cues most helpful in the windshear scenario (Fig. 28). Pitch did not help in the sidestep scenario, but provided some help in the go-around scenario. Roll was somewhat helpful during the sidestep, but much less than the pitch cues in windshear. In the other scenarios, roll cues were not so helpful. None of the axes shows a significant change in subjective haptic usefulness between the two display variants.

Results of a similar usefulness questionnaire regarding the various display indications are shown in Fig. 29. It reveals that pilots consider the airspeed indication useful in all scenarios, but especially in windshear. The AoA indication is even more useful in the windshear scenario, and for some pilots also in the go-around. The indication of bank is somewhat helpful during the sidestep scenario, but not in the other scenarios. And finally, the load factor indication is almost never useful according to the pilots, who often mentioned that they did not look at it at all.

2. *Postblock Questionnaire*

The Van der Laan ratings, which were collected after six consecutive flights with one of the display options, are shown in Fig. 30 averaged per category [25]. The ratings show a small, insignificant positive effect of the PFD+ on usefulness (Fig. 30a). No such difference was observed in the system acceptance (Fig. 30b). Nevertheless, the spread did reduce in both categories when the PFD+ was used. When splitting the two groups of pilots, the mean of the usefulness rating of the first batch of six flights appeared to be higher than that of the second batch, irrespective of the display order. Apparently, the pilots considered the system less useful once they had practiced more with it. The mean of the acceptance rating did not change much between the first and second batches, but group B showed a greatly reduced spread with the PFD+, whereas group A did not. One pilot from group A gave the lowest rating of all pilots on both usefulness and acceptance when flying the PFD. His ratings were significantly higher with PFD+. As shown in Fig. 30c, only two pilots gave the system a negative usefulness rating, both when flying with the PFD. The PFD ratings showed a strong correlation between usefulness and acceptance with a Pearson correlation coefficient $\rho = 0.877$, while the correlation was weaker with the PFD+ ($\rho = 0.757$).

To get a better understanding of what might have lowered their ratings, the pilots were asked what they would remove from the haptic system, if anything at all. No differences were observed between the two displays, with the exception of the neutral point shift at high speed. Two pilots would like to remove this cue with the PFD, but not with the PFD+. The neutral point shifts in general were not noticed by the pilots unless they explicitly paid attention. One pilot attributed this to his "flying with my fingertips." Furthermore, four pilots who would like to see the pulse removed were annoyed by the strict limits in bank. They also considered the pulses in pitch a nuisance when extending the flaps brought them above the 20 kts margin toward the maximum speed limit while still flying below the maximum flap extension speed. The pulse itself was said to have the potential of a good attention grabber, as long as the limits are set to realistic values.



a) Pitch **b) Roll**

Fig. 28 Subjective usefulness ratings of the two haptic axes in helping to stay inside the FE limits.

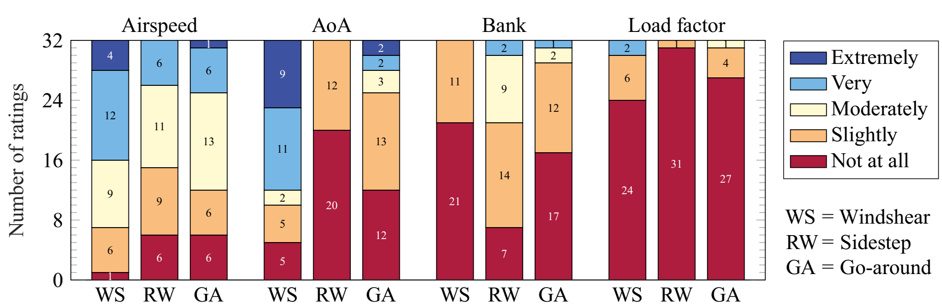
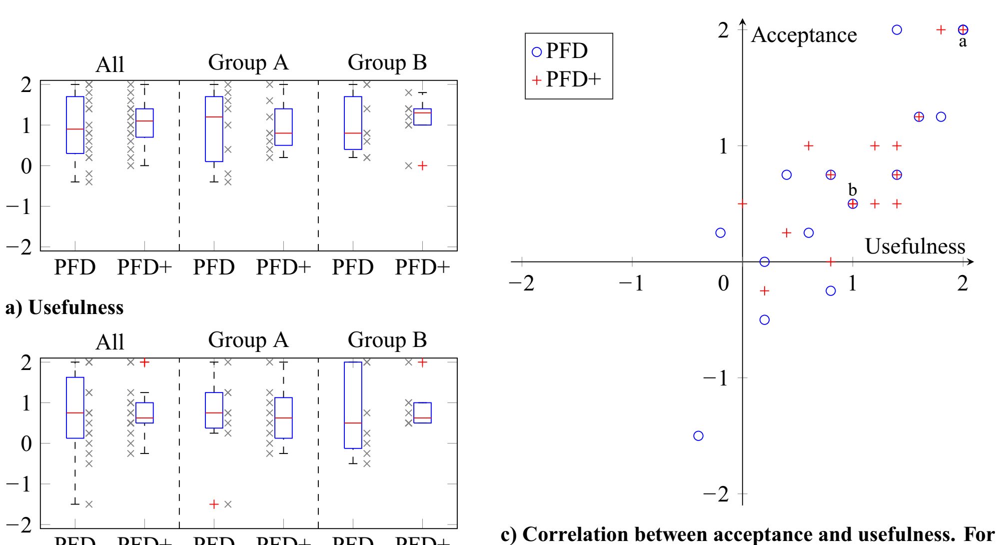


Fig. 29 Subjective usefulness ratings of PFD+ display elements in helping to stay inside the FE limits.



a) Usefulness

b) Acceptance

c) Correlation between acceptance and usefulness. For clarity: a) three overlapping PFD markers, b) two overlapping PFD+ markers

Fig. 30 Van der Laan ratings per pilot and display.

The same question was asked regarding the display indications, assuming that the haptics would not change. The load factor is the only indication that should be removed according to a majority of eleven pilots (four from group A and seven from group B), with the other indications receiving at most three nominations in total for removal.

Asking about the usefulness of the various haptic cues in preventing envelope excursions, all cues except for the stick shaker were considered more useful with the PFD+ (Fig. 31). The increasing stiffness and shifts of neutral point stand out with considerably higher ratings. The pulse was slightly more useful with the PFD+, while the stick shaker was considered slightly less useful. The number of "*not at all*" ratings for the pulse and neutral point shifts correspond to the similar number of pilots that indicated that these should be removed from the system.

The same question was asked about the various elements of the display indications (Fig. 32). The indication of the protected limit (beyond which the SFE is exited) was considered just slightly more useful than the indication of the maximum FE limit. Despite a slight inclination toward useful there was no clear consensus between the pilots on whether the thickening of the indication is a useful aid in preventing envelope excursions.

The Modified Cooper–Harper (MCH) ratings in Fig. 33 show little differences between the old and new PFDs, except that the spread was less with the new display and there were fewer ratings of 4 and worse. When looking at the PFD+ ratings for each group separately, it can be

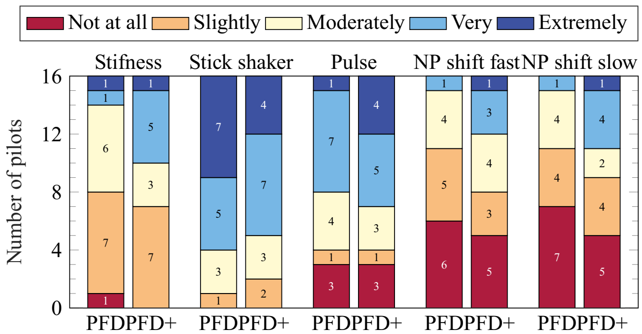


Fig. 31 Usefulness of haptic feedback cues.

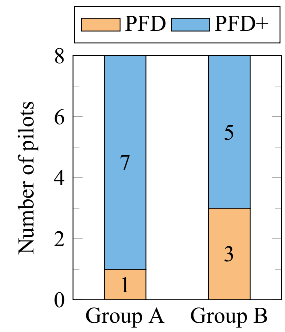


Fig. 34 Preferred display.

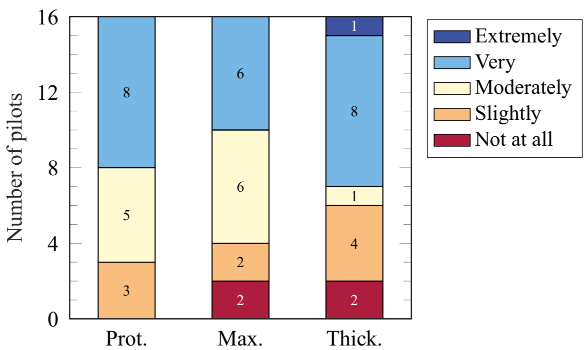


Fig. 32 Usefulness of display cues.

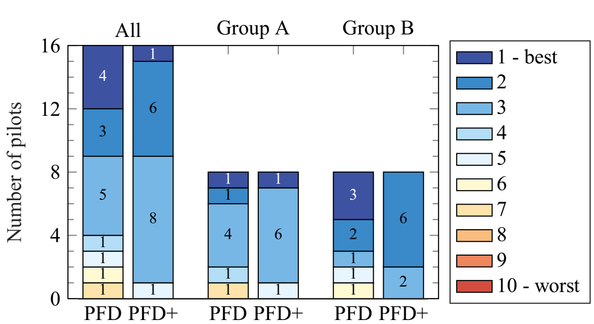


Fig. 33 Modified Cooper–Harper ratings.

observed that the rating was 3 on average for group A, while it was 2 on average for group B. To get from a 2 to a 3 or vice versa, one must answer differently on the question, "Does the system support efficient decision making?" An MCH rating of 1 or 2 indicates that this question was answered with "yes," while a rating of 3 or more can only be chosen when the question is answered with "no."

3. *Postexperiment Questionnaire*

At the end of the experiment, the pilots completed one final questionnaire. Most pilots in both groups indicated that they would like to see the haptics combined with the PFD+ (Fig. 34). In group B the preference was less pronounced than in group A, but there was still a small majority in favor of the PFD+ over the original PFD.

Apart from this binary question, several statements were posed to get a better understanding of how the pilots experienced the system and the experiment itself. The results are shown in Fig. 35. From the figure, a small positive effect of the PFD+ on understanding the haptic cues can be observed. With the PFD, which lacked an indication for the overspeed protection at 20 kts below V_{MO}, some pilots experienced pulses in the stick that they could not explain. Pilots also indicated to be able to return faster to the SFE upon exceeding the envelope when using the PFD+.

Almost all pilots were of the opinion that their understanding of the haptics and display increased throughout the experiment—the so-called "learning effect." Nevertheless, a small majority of pilots thought that the system does not require lots of training. The vast majority of pilots was of the opinion that the system would help prevent critical situations and if such situations do occur that the system would help solve them. In fact, almost all pilots thought that implementation of such a system would have a positive effect on safety; only one pilot thought that safety would be unchanged. Finally, there was no consensus on whether the display is too distracting. Pilots who said it was, often attributed this to the strict bank angle limits leading to—when being accustomed to normal bank limits—premature warnings on the bank scale.

In terms of simulation fidelity, all aspects of the experiment were considered acceptable or better by the vast majority of pilots (Fig. 36). Two "*unrealistic*" ratings on flight dynamics and weather were given by a pilot from group A, who also gave the lowest rating of all pilots on the side stick and ND. The other "*unrealistic*" rating for weather was by another group A pilot. There were considerable comments on the flight dynamics model, primarily about the thrust setting not matching that of a real Airbus and a too high sensitivity in pitch, which were also primary complaints in our earlier research. In terms of weather, some pilots thought that the windshears were too strong compared to their usual training scenarios and some attributed the effect of wind on the aircraft to the weather system instead of the flight dynamics model. The projected environment (terrain, airport, and sky) was rated acceptable or better by all pilots.

When taking a closer look at the two—for this experiment—most important simulation elements, the PFD and the side stick, it is clear that both were sufficiently realistic. Pilots were in general very positive about the realism of the PFD, saying that it resembled the real instrument very well. Most criticism was about the nervousness of the speed trend vector and occasional disappearance of the flight path vector (FPV). The FPV only disappeared during the training sessions at KSEA when the aircraft was flying a heading of exactly 180°. This was only discovered on the third experiment day and was therefore left unfixed for the remainder of the experiment. The ND scored mostly acceptable or better, although some pilots missed the track indication from the real aircraft to help them line up with the runway.

The nominal feeling of the side stick was considered at least acceptable by all but one pilot. Several pilots commented that the pitch and roll axes are more separated in the real stick, allowing for separate inputs in either axis. With the simulated stick, they found it difficult to only apply pitch inputs without inadvertent roll input.

VI. Discussion

Previous research has indicated that adding visualizations to haptic feedback improves user acceptance of this feedback and possibly also yields performance and safety benefits [7]. Our experiment results indeed show a slight improvement in acceptance and safety with the newly designed display, but not any change (neither increase nor decrease) in performance. The following discussion is split into parts that follow the hypotheses. It concludes with the experiment setup and an overall system evaluation.

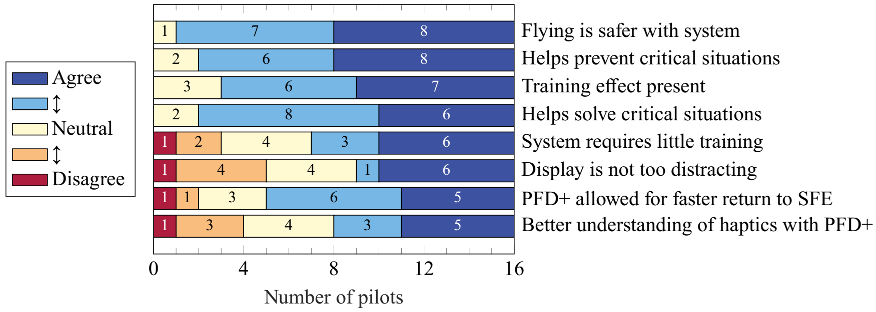


Fig. 35 Subjective postexperiment ratings.

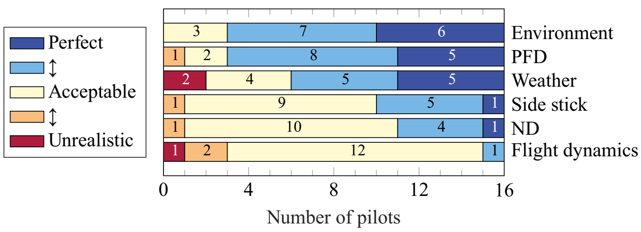


Fig. 36 Subjective simulation ratings.

A. Workload

When it comes to workload, there were only small changes observed in control activity, increasing or decreasing depending on the scenario. All changes lacked statistical significance. The secondary task, replying to ATC requests, turned out to be unusable for workload analysis due to the small number of missed ATC requests. Handling these requests was too easy for the pilots and did not draw their focus away from the primary task. A comparable result was seen in an experiment with a similar setup [6]; therefore future research should make use of a different secondary task to aid measuring workload. The subjective RSME rating showed no change in workload either, nor did any of the pilots hint on a change in workload in the debriefing. Thus, it is reasonably safe to conclude that the PFD+ does not lead to a change in workload, rejecting Hypothesis H1. The fact that there is no *increase* in workload either makes the PFD+ an acceptable display from the perspective of workload.

B. Performance

Concerning the go-around scenario, several pilots indicated that it was "unusual" to maintain the instructed 2000 ft/ min on go-around, so they sometimes forgot to pay attention to the vertical speed. Another possible cause of the low scores for this scenario is the standard procedure to start reducing the rate of climb some 10% below the target altitude, while the score was based on the climb rate up to 100 ft (ca. 5%) below the missed approach altitude.

While the performance measures in the sidestep and go-around scenarios were not expected to see significant improvements with the PFD+, there were strong expectations that the AoA indication would lead to better windshear performance. In theory, it allows pilots to fly at the maximum performance of the aircraft, reducing the altitude lost during recovery. This is, however, not reflected in the results. A possible explanation is that the indication persuaded pilots to pitch up further than the standard 17.5° dictated by procedures. A larger pitch angle makes it harder to recover the aircraft once stalled, as demonstrated by the two windshear crashes that both occurred with the PFD+. Limiting the indication to a fixed maximum—similar to Boeing's PLI [19]—to prevent excessive pitch (pilot following symbol) may diminish this problem. Another potential cause of poor performance was the ambiguity of the AoA indicator's reference. Aligning the "whiskers" such that they touch the upper side of the fixed aircraft symbol when the AoA margin is zero would solve this ambiguity, while also making it easier for pilots to "ride" on the limit. Concluding, the new display seems to neither significantly improve nor deteriorate performance. Hypothesis H2 is thus also rejected.

C. Safety

During windshear, pilots flying the PFD+ spent slightly less time outside the SFE at high AoA and airspeeds. The decrease in time in overspeed protection during the approach phase clearly shows that the stringent 20 kts high-speed margin, only visible to pilots with the PFD+, changed pilot behavior when not only communicated through haptic feedback. Similar behavior was seen in roll. While the time spent in roll protection also significantly reduced, the artificially strict bank limits may have had a big impact on pilot behavior in roll. To ensure that the pilots would enter the roll protection, the bank angle limits in the experiment were artificially reduced compared to the real aircraft. Many pilots indicated that this was unrealistic and perceived the roll cues in the haptic system and display as a nuisance because they activated while flying at a bank angle perfectly acceptable in normal operation. A different scenario setup may allow for the standard bank angle limits to be used, while still bringing the pilots near the edges of the SFE. Nevertheless, pilots did respect the bank limits more when shown on the PFD+. Hypothesis H3 (i.e., the margin to the FE boundaries would become larger with the PFD+) cannot be accepted, however, due to a lack of statistically significant differences. There does seem to be a small effect of the display that warrants further research.

While the objective effect on safety was limited, a large share of the pilots does expect that the system would improve safety when implemented. The data do not provide a definitive answer on whether that can be attributed to the haptics, the display, or the combination of both. Previous research suggests that the haptic system by itself is already seen as a safety improvement, so the effect of the display may be limited here [6].

D. Pilot Appreciation

Overall, most pilots preferred the PFD+ over the conventional PFD, suggesting an improved acceptance of the haptic system in combination with the new display. This is confirmed by the increased usefulness ratings of the haptic cues with the PFD+. Still, Van der Laan and MCH ratings did not indicate a significant change in appreciation of the system as a whole. A possible explanation for this is that the haptic feedback, which was always enabled, was a more prominently present novelty for the pilots and thus had a larger impact on their systemwide ratings than the display. Testing a baseline condition, with no haptic feedback, would show the effect of just the haptic feedback. Previous research did include such a condition, but did not use the Van der Laan and MCH rating scales [6,28]. Based on these findings Hypothesis H4 cannot be unequivocally accepted.

E. Display Indications

As hypothesized, the load factor indication was considered the least useful indication by the pilots. They often indicated that they did not look at it at all for mainly two reasons. Firstly, it is simply not needed, because whenever the load factor limits are reached there is always another limit crossed (in the conditions from the experiment that is indeed true). The other reason is that the indication is added to the left of the speed tape, where in the actual Airbus there is nothing. The new indication was therefore most likely not included in the scanning pattern, but eye-tracking measures would be needed to confirm this. It is worth noting that several pilots considered the addition of a load factor indication "extremely useful" during the briefing at the *start* of the experiment, but then changed their opinion after flying with it. More training may improve this, but combining all results, it is expected that a load factor indication brings no extra benefit over the other indications. In future research, the load factor indication can be removed to reduce visual load and to make the display better fit in the standard Airbus display size.

The AoA indication, on the other hand, was much more appreciated by the pilots. The only pilot who stated to remove it proposed to show the AoA and load factor indications only in critical situations, like windshear or terrain avoidance maneuvers. Although it did not bring the expected performance benefit, it gave most pilots the feeling that they were better informed about the state of the aircraft. It is probably also the reason why the stick shaker is considered less useful with the PFD+, as stall information is now also communicated through the AoA indication. Hypothesis H5 is thus accepted.

A limitation of the decision to adhere to only visualize the cues and limits as communicated by the haptic feedback system was the inability to add indications outside the PFD, such as on the engine displays. Especially for speed limits, engine controls may be a desirable control input worth investigating in future research. The presented design is furthermore limited to modern commercial airliners with digital flight instruments, where FE and state information is readily available. In analog steam-gauge style cockpits, as often found in general aviation or older aircraft, this information first needs to be made available [29].

F. Experiment

Looking back at the experiment itself, the use of two pilot groups with different display orders was a valid choice, as some dependent measures showed a stark contrast between the two groups. This can probably be primarily attributed to the learning effect. Haptic feedback was new for almost all pilots, and those who did fly with it before did so over a year earlier. Even though the pilots received considerable training, they were clearly still getting more accustomed to both the simulator and researched systems as the experiment progressed. Subjective results may have also been affected by the fact that the pilots did not fly a baseline condition with haptic feedback disabled and the original PFD. Although this could have helped determine whether any changes are caused by the haptic feedback itself, or the display, the experiment would also have lasted much longer, possibly even 2 days per pilot, which is impractical.

In the aim for realistic scenarios and a high ecological validity, pilots were given much freedom, which lead to challenges in the data analysis. It could help to limit this freedom in future experiments. Extending the route on the ND from the FAF to the starting position can encourage pilots to intercept the localizer at more consistent angles than flown in this research. Using the autopilot to bring the aircraft to a predetermined state and then hand-over control to the pilot on the onset of an event may also help and is an accepted method in flight training [30]. As with any simulator experiment, the simulator itself may also have influenced the results. To minimize the impact of differences between the real aircraft and the simulation, pilots were given considerable time to familiarize themselves. Together with this research's focus on the PFD and side stick, both rated as sufficiently realistic by the pilots, the differences with respect to the real aircraft are considered acceptable with insignificant influence on the outcome.

Finally, the lack of simulator motion may have influenced pilot behavior. Especially in stall conditions, pilots are known to sometimes over-react when they do not feel the load factor [31]. Displaying the load factor was expected to make up for this lack of information. However, as discussed before, pilots did not pay much attention to the load factor indicator, so this cannot be assumed to be an adequate replacement. An experiment involving haptic feedback, PFD+, and motion cueing should be conducted to see whether motion indeed has any effect.

G. Overall System Evaluation

Wrapping up, supplementing haptic feedback through the PFD+ produces no big improvements, nor does it introduce drawbacks. Because most pilots appreciated the display, although with a number of simplifications, the integration of the input and output spaces seems to be a feasible solution. The display appears to fulfill its main design goal: increasing the understanding and appreciation of haptic feedback as a way to communicate FE boundaries. The system was only evaluated in approach scenarios, though, while a substantial number of LOC-I accidents in recent years occurred during initial climb-out and cruise [32]. Testing the haptics and display in a cruise situation where pilots suddenly find themselves in alternate law could show the potential of the system in a wider range of flight phases. A future design iteration may benefit from changes to the haptic feedback, for example, inhibiting the overspeed warning if the flaps are already transitioning to a setting allowing for a higher V_{MO}. Care must then be taken that the visualizations remain conformal to the cues from the haptic feedback.

VII. Conclusions

This study explored visual ways to supplement haptic feedback in communicating FE boundaries to pilots. The resulting display design is unique in integrating *both* the limits of the FE *and* the limits of the control inputs in one display. The display also shows the direction and magnitude of the haptic feedback applied on the side stick. To accomplish this, the standard A320 PFD has been enhanced with new indications for AoA, airspeed, bank angle, and load factor.

The design presented in this paper did not yield significant differences in performance compared to the original PFD. Small but significant changes were observed in the time spent outside the SFE regarding roll angle and airspeed, hinting on a potential safety improvement. Subjective ratings showed that the display led to a small increase in pilot appreciation of the haptic feedback. The display increased pilots' understanding of what the haptics were aiming to communicate and helped them to keep their aircraft within the SFE limits.

In conclusion, the proposed display can help increase pilot appreciation of haptic feedback. The combined system can lead to an improvement in aviation safety by reducing LOC-I accidents. Future research should focus on improving the display and experiment design. The unused load factor indication should be removed to reduce clutter. Especially the AoA indication appears to be useful, but also led to a number of crashes when pilots followed it too closely. Further research is suggested to improve this particular indication and reduce its ambiguity. It is also recommended to test the display with the actual operational bank limits, instead of the artificially reduced ones used in this research.

Acknowledgments

The authors express their gratitude to the 16 pilots who participated, as well as those who did not participate but helped recruiting colleagues. A final word of thanks is dedicated to all personnel that helps keep the simulator in optimal shape such that experiments like this can be conducted with high reliability of the available facilities.

References

[1] "Statistical Summary of Commercial Jet Airplane Accidents," Worldwide Operations, 1959–2020, Boeing Commercial Airplanes, Sept. 2021, https://www.skybrary.aero/sites/default/files/bookshelf/32664.pdf.

[2] "Aviation Occurrence Categories–Definitions and Usage Notes," CAST/ICAO Common Taxonomy Team, May 2021, http://www.intlaviationstandards.org/Documents/OccurrenceCategoryDefinitions.pdf.

[3] "Aircraft Accident Investigation Report PK-AXC," Komite Nasional Keselamatan Transportas, Rept. KNKT.14.12.29.04, 2015, https://bea.aero/fileadmin/uploads/tx_elydbrapports/Final_Report_PK-AXC-reduite.pdf.

[4] "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 Bureau d'Enquêtes et d'Analyses (BEA), July 2012, https://bea.aero/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf.

[5] Van Baelen, D., Ellerbroek, J., van Paassen, M. M., and Mulder, M., "Design of a Haptic Feedback System for Flight Envelope Protection," *Journal of Guidance, Control, and Dynamics*, Vol. 43, No. 4, 2020,

pp. 700–714.
https://doi.org/10.2514/1.G004596

[6] Van Baelen, D., Ellerbroek, J., van Paassen, M. M., and Mulder, M., "Evaluation of a Haptic Feedback System for Flight Envelope Protection," *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2019-0367, 2019.
https://doi.org/10.2514/6.2019-0367

[7] Ho, V., Borst, C., van Paassen, M. M., and Mulder, M., "Increasing Acceptance of Haptic Feedback in UAV Teleoperation by Visualizing Force Fields," *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Inst. of Electrical and Electronics Engineers, New York, 2018, pp. 3027–3032.
https://doi.org/10.1109/SMC.2018.00514

[8] Meuleau, N., Neukom, C., Plaunt, C., Smith, D. E., and Smith, T., "The Emergency Landing Planner Experiment," *ICAPS-11: Proceedings of the Scheduling and Planning Applications Workshop*, Assoc. for the Advancement of Artificial Intelligence, Palo Alto, CA, 2011, pp. 60–67.

[9] Rijndorp, A. D. T., Borst, C., de Visser, C. C., Stroosma, O., van Paassen, M. M., and Mulder, M., "Aviate, Navigate: Functional Visualizations of Asymmetric Flight Envelope Limits," *Journal of Aerospace Information Systems*, Vol. 18, No. 12, 2021, pp. 933–948.
https://doi.org/10.2514/1.I010895

[10] Lombaerts, T. J. J., Schuet, S., Acosta, D., Kaneshige, J., Shish, K., and Martin, L., "Piloted Simulator Evaluation of Safe Flight Envelope Display Indicators for Loss of Control Avoidance," *Journal of Guidance, Control, and Dynamics*, Vol. 40, No. 4, 2017, pp. 948–963.
https://doi.org/10.2514/1.G001740

[11] Lombaerts, T. J. J., Looye, G., Ellerbroek, J., and Martin, M. R. Y., "Design and Piloted Simulator Evaluation of Adaptive Safe Flight Envelope Protection Algorithm," *Journal of Guidance, Control, and Dynamics*, Vol. 40, No. 8, 2017, pp. 1902–1924.
https://doi.org/10.2514/1.G002525

[12] Ackerman, K. A., Talleur, D. A., Carbonari, R. S., Xargay, E., Seefeldt, B. D., Kirlik, A., Hovakimyan, N., and Trujillo, A. C., "Automation Situation Awareness Display for a Flight Envelope Protection System," *Journal of Guidance, Control, and Dynamics*, Vol. 40, No. 4, 2017, pp. 964–980.
https://doi.org/10.2514/1.G000338

[13] *A318/A319/A320/A321 Flight Crew Operating Manual*, Airbus, Toulouse, France, Aug. 2015.

[14] Rustenburg, J. W., Skinn, D. A., and Tipps, D. O., "Statistical Loads Data for the Airbus A-320 Aircraft in Commercial Operations," Univ. of Dayton Research Inst. Rept. DOT/FAA/AR-02/35, Dayton, OH, April 2002.

[15] Wickens, C. D., Lee, J., Liu, Y., and Becker, S. G., *An Introduction to Human Factors Engineering*, 2nd ed., Pearson–Prentice-Hall, Upper Saddle River, NJ, 2004, pp. 185–193.

[16] "Electronic Flight Displays," Advisory Circular 25-11B, Federal Aviation Administration, July 2014, https://www.faa.gov/documentlibrary/media/advisory_circular/ac_25-11b.pdf.

[17] Roscoe, S. N., "Airborne Displays for Flight and Navigation," *Human Factors*, Vol. 10, No. 4, 1968, pp. 321–332.
https://doi.org/10.1177/001872086801000402

[18] Roscoe, S. N., Corl, L., and Jensen, R. S., "Flight Display Dynamics Revisited," *Human Factors*, Vol. 23, No. 3, 1981, pp. 341–353.
https://doi.org/10.1177/001872088102300309

[19] Cashman, J. E., Kelly, B. D., and Nield, B. N., "Operational Use of Angle of Attack on Modern Commercial Jet Airplanes," AERO, 12, Boeing Commercial Airplanes, 2000, pp. 10–22, https://www.boeing.com/commercial/aeromagazine/aero_12/aoa.pdf.

[20] Lombaerts, T. J. J., Looye, G., Seefried, A., Neves, M., and Bellmann, T., "Development and Concept Demonstration of a Physics Based Adaptive Flight Envelope Protection Algorithm," *IFAC-PapersOnLine*, Vol. 49, No. 5, 2016, pp. 248–253.
https://doi.org/10.1016/j.ifacol.2016.07.121

[21] Van Paassen, M. M., Stroosma, O., and Delatour, J., "DUECA–Data-Driven Activation in Distributed Real-Time Computation," *Modeling and Simulation Technologies Conference*, AIAA Paper 2000-4503, 2000.
https://doi.org/10.2514/6.2000-4503

[22] "OpenStreetMap Maps," OpenStreetMap Contributors, 2019, https://www.openstreetmap.org.

[23] "Windshear Training Aid," The Boeing Company, 1990, https://flightcrewguide.com/wp-content/uploads/2013/11/Windshear-Training-Aid-Vol-1.pdf.

[24] Zijlstra, F. R. H., "Efficiency in Work Behaviour: A Design Approach for Modern Tools," Ph.D. Thesis, Delft Univ. of Technology, Delft, The Netherlands, Nov. 1993.

[25] Van der Laan, J. D., Heino, A., and de Waard, D., "A Simple Procedure for the Assessment of Acceptance of Advanced Transport Telematics," *Transportation Research Part C: Emerging Technologies*, Vol. 5, No. 1, 1997, pp. 1–10.
https://doi.org/10.1016/S0968-090X(96)00025-3

[26] Cummings, M. L., Myers, K., and Scott, S. D., "Modified Cooper Harper Evaluation Tool for Unmanned Vehicle Displays," *Conference on Unmanned Vehicle Systems Canada*, 2006.

[27] Wilde, G. J. S., "Risk Homeostasis Theory and Traffic Accidents: Propositions, Deductions and Discussion of Dissension in Recent Reactions," *Ergonomics*, Vol. 31, No. 4, 1988, pp. 441–468.
https://doi.org/10.1080/00140138808966691

[28] Ellerbroek, J., Martin, M. R., Lombaerts, T. J. J., van Paassen, M., and Mulder, M., "Design and Evaluation of a Flight Envelope Protection Haptic Feedback System," *IFAC-PapersOnLine*, Vol. 49, No. 19, 2016, pp. 171–176.
https://doi.org/10.1016/j.ifacol.2016.10.481

[29] Khan, W., Ansell, D., Kuru, K., and Bilal, M., "Flight Guardian: Autonomous Flight Safety Improvement by Monitoring Aircraft Cockpit Instruments," *Journal of Aerospace Information Systems*, Vol. 15, No. 4, 2018, pp. 203–214.
https://doi.org/10.2514/1.I010570

[30] "Airplane Upset Prevention & Recovery Training Aid for Transport Category Airplanes," Rev. 3, International Civil Aviation Organization, Feb. 2017, https://www.icao.int/safety/loci/auprta/index.html.

[31] Le Vie, L. R., "Survey of Quantitative Research Metrics to Assess Pilot Performance in Upset Recovery," NASA TM-2016-219191, April 2016.

[32] "Loss of Control In-Flight Accident Analysis Report," International Air Transport Association, 2019, https://www.iata.org/contentassets/b6eb2adc248c484192101edd1ed36015/loc-i_2019.pdf.

D. Casbeer
Associate Editor