Supplementing Haptic Feedback Through the Visual Display of Flight Envelope Boundaries

de Rooij, Gijs; Van Baelen, Dirk; Borst, Clark; van Paassen, Rene; Mulder, Max

DOI
10.2514/6.2020-0373

Publication date
2020

Document Version
Final published version

Published in
AIAA Scitech 2020 Forum

Citation (APA)
de Rooij, G., Van Baelen, D., Borst, C., van Paassen, R., & Mulder, M. (2020). Supplementing Haptic Feedback Through the Visual Display of Flight Envelope Boundaries. In *AIAA Scitech 2020 Forum: 6-10 January 2020, Orlando, FL* (pp. 1-25). Article AIAA 2020-0373 (AIAA Scitech 2020 Forum; Vol. 1 PartF). American Institute of Aeronautics and Astronautics Inc. (AIAA). https://doi.org/10.2514/6.2020-0373

Supplementing Haptic Feedback Through the Visual Display of Flight Envelope Boundaries

Gijs de Rooij *, Dirk Van Baelen †, Clark Borst ‡, Marinus M. van Paassen § and Max Mulder ¶
Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands

This paper describes the design and evaluation of a visual display in supplementing haptic feedback on the side stick as a way to communicate flight envelope boundaries to pilots. The design adds indications for the limits in airspeed, load factor, angle of attack and angle of bank to a standard Airbus primary flight display (PFD). The indications not only show the limits of the flight envelope, but also indicate magnitude and direction of the haptic cues. Fifteen professional Airbus pilots and one Airbus sim instructor participated in an experiment in the SIMONA Research Simulator at Delft University of Technology. Several approaches in three different scenarios were flown in alternate law with the old and new PFD, while haptic feedback was always enabled. Objective results do not show clear improvements with the new display, 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 focus on improving the design by removing unused indications and setting up an experiment with a bank scenario that allows the use of operational bank limits rather than artificially reduced limits.

I. Introduction

Aviation is one of the safest forms of transport, yet there are still improvements to be made. In the last decade, loss of control in-flight (LOC-I) has been the primary cause of fatal commercial jet airplane accidents [1]. There are multiple slightly different 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 the pilot to control the aircraft [2]. Modern fly-by-wire (FBW) aircraft are protected from such FE excursions, but when automation fails, the pilots find themselves in a stressful situation and do not always know what to do 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 unrecoverable prolonged stall. All 162 people on board perished when the aircraft crashed into the Java Sea. This and other incidents like Air France 447 [4] show that once protections are lost, pilots lack clear cues on their position with respect to the flight envelope and how to return to the envelope if they exceed it.

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 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 looked at the design of displays that can show (more) flight envelope information. The primary flight display (PFD) seems to be the preferred location to integrate such information, altough some projects designed stand-alone displays. A common factor in most existing solutions is the separation of output and input space, showing either the limits of the envelope [8–11] or the limits in control inputs that would otherwise bring the aircraft

*Ph.D. student, Delft University of Technology - Control & Simulation, g.derooij@tudelft.nl. Member AIAA
†Ph.D. student, Delft University of Technology - Control & Simulation, d.vanbaelen@tudelft.nl. Student Member AIAA
‡Assistant professor, Delft University of Technology - Control & Simulation, c.borst@tudelft.nl
§Associate professor, Delft University of Technology - Control & Simulation, m.m.vanpaassen@tudelft.nl
¶Professor, Delft University of Technology - Control & Simulation, m.mulder@tudelft.nl. Associate Fellow AIAA

outside that envelope [12]. No previous research is known on an aircraft display specifically integrating the limits on the input and output space together with information on associated haptic feedback.

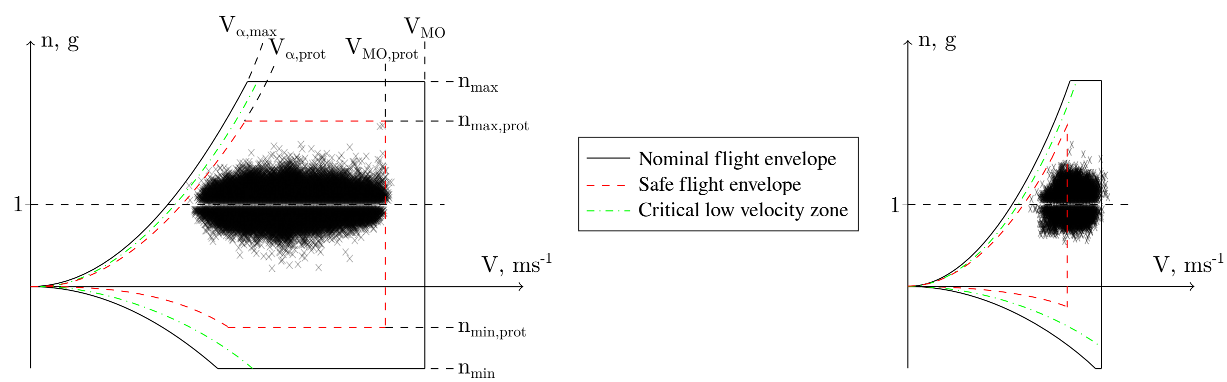
This study builds on the foundations of the aforementioned research by investigating the design of a display that integrates the input and output space, while also showing the force and direction cues of the haptic feedback. The paper starts with background information on flight envelopes and haptic feedback in Section II. Section III presents the display design and explains the rationale behind it. The display was tested in a human-in-the-loop simulator experiment involving 16 professional pilots as explained in Section IV, to asses the added value of said display. Section V lists the results of the experiment, which are then discussed in Section VI together with some recommendations for further research. Finally, Section VII concludes the article.

II. Background

A basic understanding of the flight envelope is required to grasp the working of the haptic feedback and consequential 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 flight envelope (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 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 in addition moves the lower and upper load factor limits to 0 and 2 g respectively when the flaps are not up [13], but the model from this paper keeps the load factor limits at -1 and 2.5 g in order to match the haptic feedback.



(a) Flaps clean **(b) Flaps 3**

Fig. 1 Typical flight envelopes with velocity (V) versus load factor (n) [5]. Augmented with load factor data for 10,066 A320 flights [14]. The actual envelopes depend on the aircraft's configuration and loading.

Safety margins are added to the FE to create a so-called safe flight envelope (SFE), indicated by the red dashed line in Fig. 1. The associated *prot*ection margins are 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 0.5 g, lower speed margins vary along the envelope, and high-speed margin is fixed at 20 kts below V_{MO}. Another margin can be distinguished near the lower velocity 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, nevertheless some flights do get close to the limits of the SFE. On the contrary, with the significantly smaller FE corresponding to a flaps 3 configuration, the majority seems to operate outside the SFE. This can be explained by the fact that the fixed 20 kts overspeed margin of the SFE is not used in real-life operations.

B. Haptic feedback

The working principles of the haptic feedback system are best explained using the haptic profile, shown in Fig. 6c, which gives the stick deflection δ and the 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. The full description can be found in our previous paper [5]. While only longitudinal haptic feedback was considered 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, depicted by the in-set graph in Fig. 6f. Second, continuing to steer the aircraft out of the SFE, results in a progressively increased stiffness as shown by the asymmetric profile on Fig. 6f. Third, when zero stick input is insufficient to return to the SFE for low velocities, the neutral point moves as shown on Fig. 6i. Fourth, Fig. 6i shows the stick shaker when crossing the critical lower velocity indicated on the FE in Fig. 6g by the dashed green line. Fifth, Fig. 6l shows the shifted neutral point of the stick during an overspeed situation to indicate the automatic pitch up command.

III. Display design

The haptic feedback system from Section II.B was tested with professional pilots in previous research, resulting in a recommendation to investigate the addition of a display to visualize the haptic cues [6]. Combining haptic feedback with a visual display could fulfill the important principle of multiple resources when presenting 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

In order to support the haptic system, the indications on the display have to match with the forces felt through the side stick in both magnitude and direction. From the cues discussed in Section II.B, the discrete cue and changing stiffness can be visualized by an ordinary spring (upper part of Fig. 2) that is positioned next to the side stick. When the aircraft approaches the edge of the SFE as discussed in Section II.A the spring moves towards 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 tick 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. The maximum compression is reached when the two vertical lines touch each other. Beyond that maximum the spring coefficient does not change any further.

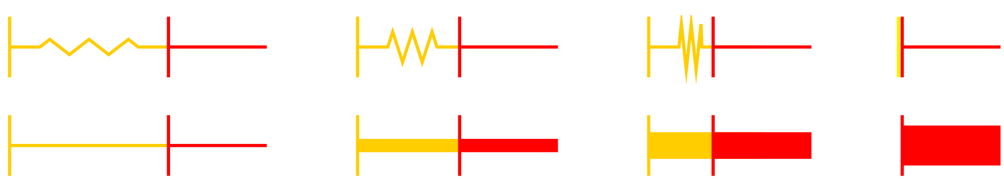


Fig. 2 Spring (top) and piston (bottom) symbols with increasing levels of compression.

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 part of Fig. 2). Apart from visualizing the 'feel' from the haptics in both magnitude and direction, these indications also show the pilot in which direction he 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].

The other cues from the haptic system are not explicitly visualized. The stick shaker is a trigger to bring the pilot's attention to the low speed rather than an actual limit, so no extra indication is added. 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 Wicken'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 flight envelope. The actual flight envelope limits are indicated in red.

In order 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 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 3 shows the PFD with all of the flight envelope indications in place. The various elements are discussed in greater detail below.

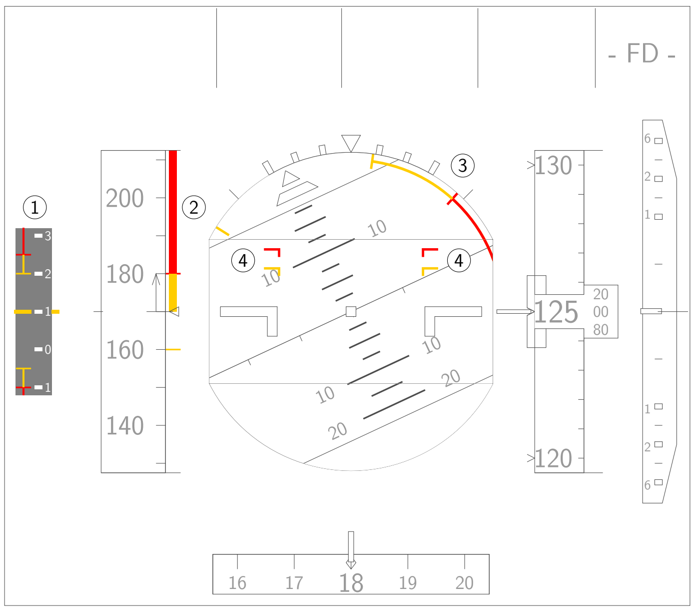


Fig. 3 Wireframe view of the Airbus A320 PFD with additional load factor indicator ① and flight envelope limits for airspeed ②, bank angle ③ and angle of attack ④.

B. Novel display indications

1. Load factor

The first addition is a load factor indicator to the left of the airspeed tape (Fig. 4). The new indicator consists of a tape showing the load factor currently acting on the aircraft. Similar to the 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 major tick marks every 1 g and minor tick marks every 0.5 g. The flight envelope limits are indicated by horizontal lines that attach to vertical lines running away from the fixed 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. 4. 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). In order 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. 3). 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.

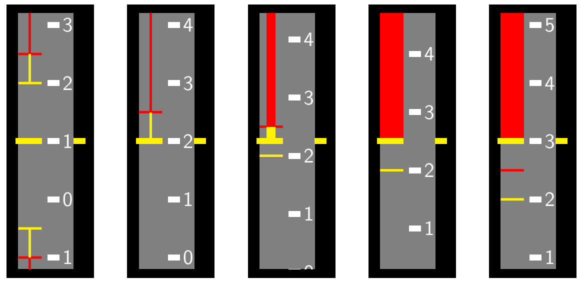


Fig. 4 Load factor indicator progressively reaching and eventually exceeding the upper limit.

A similar indication on the lower side of the speed tape corresponds to the low speed part of the flight envelope, 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 impeding stall.

One potential issue with the above described representation is that the nose has to 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 [15]. 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 angle of attack) 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. 3). 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 towards the center of the display from the side that the aircraft is banking to. According to Wickens' 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. 3, the pilot should roll left to lower the bank angle.

4. Angle of attack

An indication for margin to stall angle of attack (AoA) is added to the PFD as shown in Fig. 5. 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) [17]. 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. The whiskers are placed besides the pitch ladder to not obstruct the ladder.

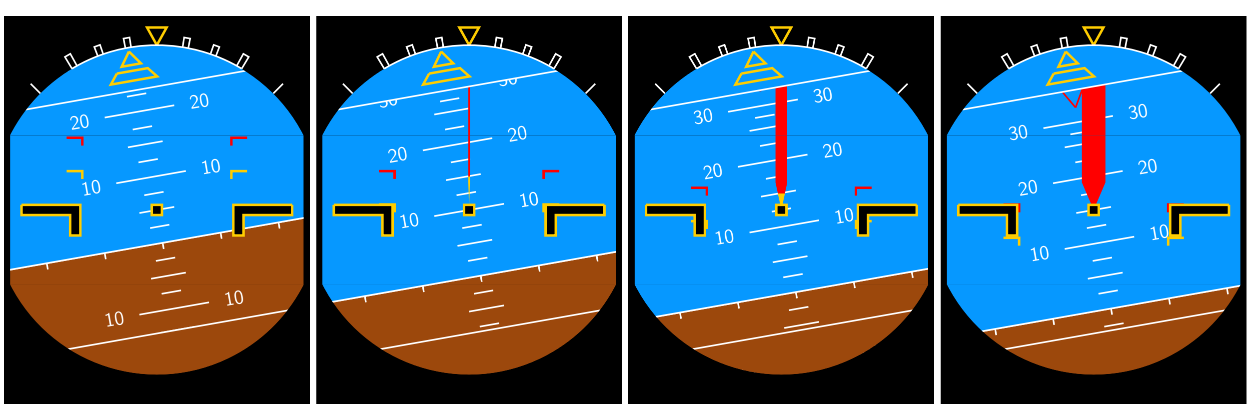


Fig. 5 AoA indicators relative to the fixed aircraft symbol progressively reaching the AoA limit.

C. Typical windshear recovery

To illustrate the synergy between the flight envelope, display and haptic feedback, Fig. 6 shows the display indications during a typical windshear escape procedure side-by-side with the flight envelope 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 tick on the stick's pitch axis, as well as an increased stick-back stiffness, to alert him 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 angle of attack indication on the pitch scale starts growing in width, as the angle of attack 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 neutral point of the stick 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 tick warns the pilot of an imminent excursion and the stick moves backwards to help the pilot bleed of airspeed. The spring stiffness of the stick is increased to inform the pilot of the distance to the ultimate flight envelope limit, as visualized by the widening of the piston on the speed tape.

IV. Method

Since pilots are expected to interact with the display, its design was tested in a human-in-the-loop simulator experiment. The goal of the experiment was to evaluate the interaction of pilots with the display, see what it does to their control strategy and whether it improves their subjective perception of the haptic feedback.

A. Participants

Fifteen professional Airbus pilots, all male, from four airlines and one male Airbus A320 synthetic flight instructor (SFI) participated in the experiment. The experience of all participants is shown in Table 1. They were divided over two groups (A and B) that experienced a different display order. Four pilots previously participated in our haptic feedback evaluation, namely A5, A6, B1 and B6. It is worth noting that the second officers – while not certified to operate the aircraft below 20,000 ft – did receive a 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 nine in an Airbus. All pilots had received upset recovery and prevention training (UPRT) and had experienced alternate law in simulator training.

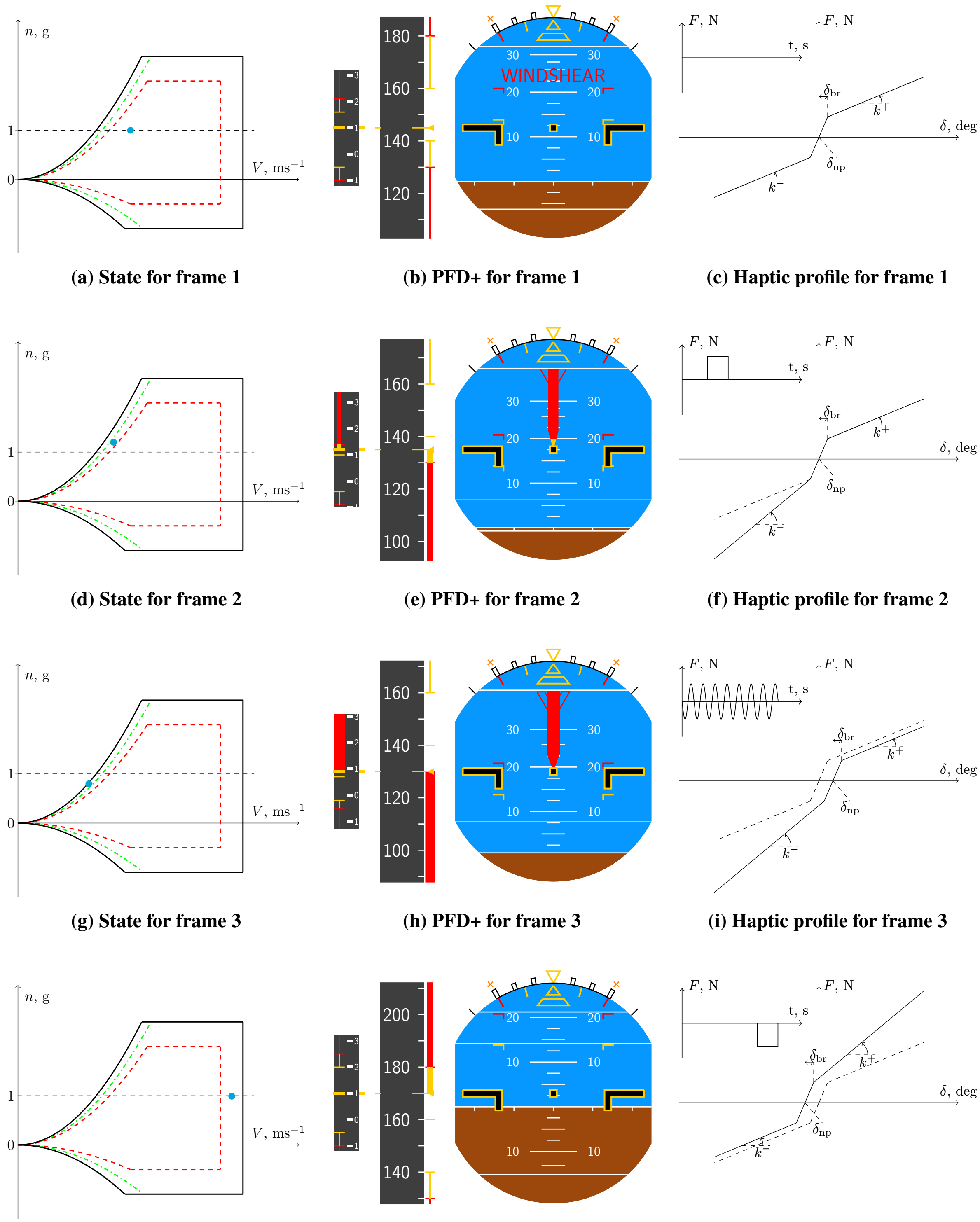
B. Apparatus

The experiment took place in the Simulation, Motion, and Navigation (SIMONA) Research Simulator (SRS) at Delft University of Technology. The simulator's exterior and interior are shown in Fig. 7 and 8 respectively. SIMONA is a six degrees 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. Since the model did not include a landing gear, all flights were automatically stopped upon reaching 50 ft above ground level (AGL).

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 [18]. The PFD and ND were drawn using OpenGL (see Fig. 14) and very closely resembled the real Airbus displays.

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



(a) State for frame 1 (b) PFD+ for frame 1 (c) Haptic profile for frame 1

(d) State for frame 2 (e) PFD+ for frame 2 (f) Haptic profile for frame 2

(g) State for frame 3 (h) PFD+ for frame 3 (i) Haptic profile for frame 3

(j) State for frame 4 (k) PFD+ for frame 4 (l) Haptic profile for frame 4

Fig. 6 Typical windshear recovery procedure. The left column shows the flight envelope, the center column shows an excerpt of the PFD and the right column shows the associated haptic pitch profile.

Table 1 Participants in the experiment.

Pilot	Age	Flight hours	Airbus flight hours	Position	Main aircraft type
A1	52	13,500	2000[2]	SFI	A320
A2	48	13,500	700	First officer	A330
A3	27	2800	2300	First officer	A320
A4	56	10,000	6000	Captain	A330
A5[1]	57	9500	9000	Captain	A320
A6[1]	47	15,000	1500	Captain	A330
A7	28	1200	600	Second officer	A330
A8	25	2300	200	Second officer	A330
B1[1]	48	16,000	5000	Captain	A330
B2	50	16,000	5000	Captain	A330
B3	43	12,500	7500	Captain	A320
B4	30	3000	400	Second officer	A330
B5	49	13,000	2000	Captain	A330
B6[1]	31	5500	5300	Captain	A320
B7	47	13,950	3300	Captain	A330
B8	39	8787	6178	Captain	A320
Mean	42	9784	3561		
Std. dev.	11	5235	2758		

[1]Pilot participated in previous haptic feedback research.

[2]These are simulator hours. Participant is a former Boeing pilot and current SFI for the Airbus A320.

C. Procedure

All participants engaged in the procedure outlined in Table 2. They were divided over two groups, with an equal distribution of aircraft types. 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 circa five hours per pilot.

Table 2 Experiment procedure.

<table>
<tr><td></td><td>45 min</td><td>40 min</td><td>30 min</td><td>30 min</td><td>60 min</td><td>60 min</td><td>20 min</td></tr>
<tr><td>Group A</td><td rowspan="2">Briefing</td><td rowspan="2">Familiarization</td><td rowspan="2">Training flights</td><td rowspan="2">Lunch</td><td>PFD flights</td><td>PFD+ flights</td><td rowspan="2">Debriefing</td></tr>
<tr><td>Group B</td><td>PFD+ flights</td><td>PFD flights</td></tr>
</table>

1) **Briefing** – At the start of the day, the pilots received a short introduction, signed a consent form 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 them were not completely resembling 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 flight envelope.
   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 simple right-hand circuit to runway 36L at Amsterdam Airport Schiphol (EHAM) was flown twice with the baseline PFD. Note that this is a non-standard approach and therefore no instrument landing system or precision approach path indicator was provided.

Fig. 7 Exterior of SIMONA at TU Delft.

Fig. 8 Interior of SIMONA at TU Delft.

Next, the pilots performed the following exercises over the North Sea to experience the haptic cues:

1) Pilot induced stall by maintaining altitude with idle throttle.
2) Overspeed by full throttle and pitching down.
3) Nose-dive followed by a strong back stick input to reach the high-g region.
4) Rolling to the left and right.
5) 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 manoeuvres were flown over the North Sea as before, apart from the closed-eyes exercise.

3) **Training** – The training phase was setup to more closely resemble operational flights and prepared the pilot for the actual measurement flights. Four approaches were flown towards runway 16R of Seattle Airport (KSEA), for which the layout is shown in Fig. 9. 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.

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+

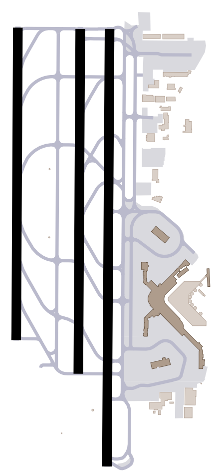


Fig. 9 Airport diagram of KSEA [19].

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

re-familiarize themselves with the flight model, haptic feedback and when applicable PFD+ after a (lunch) break. Next 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 to them and any open 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 ATC calls that they had to reply to. Each pilot's callsign reflected the company that the pilot was employed at: '{Company} 107'. To ensure 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 callsign 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 $\sigma = 2.5$ seconds). The texts were read out loud by a female American-English accent from the Festival[†] text-to-speech generation library, developed by the University of Edinburgh.

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. However, 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 in order 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. The layouts can be found in Fig. 10 and 11. 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.

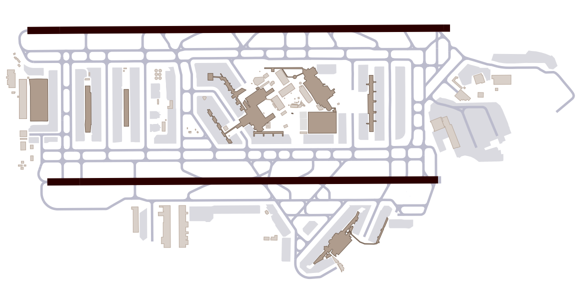


Fig. 10 Airport diagram of EGLL [19].

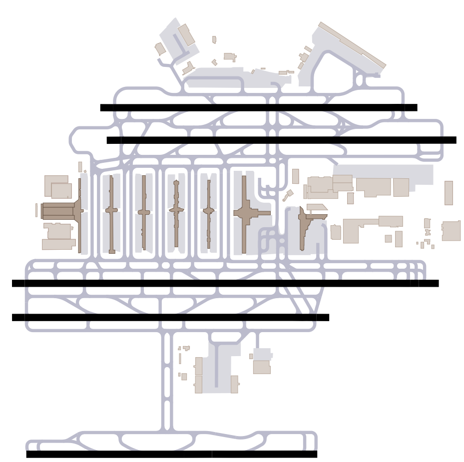


Fig. 11 Airport diagram of KATL [19].

[†]Available at `http://www.cstr.ed.ac.uk/projects/festival/`

Each flight started circa 12 NM from the airport in trimmed flaps up condition at 215 kts and an intercept heading of circa 45°, towards 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 12 shows a typical trajectory towards EGLL.

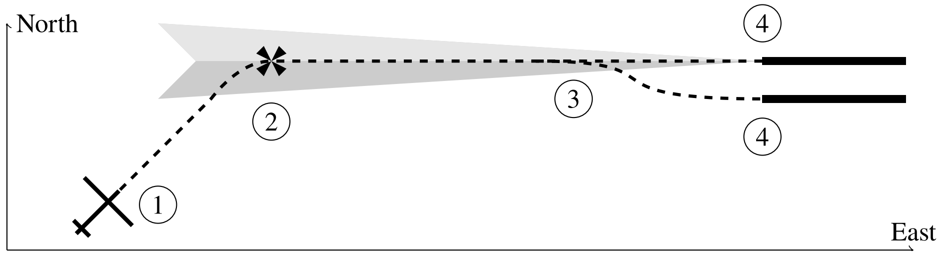
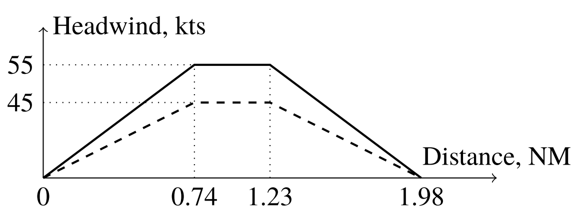


Fig. 12 Flight plan for EGLL (not to scale). Start ①, localizer interception at FAF ②, scenario triggering point ③ and end of flight ④.

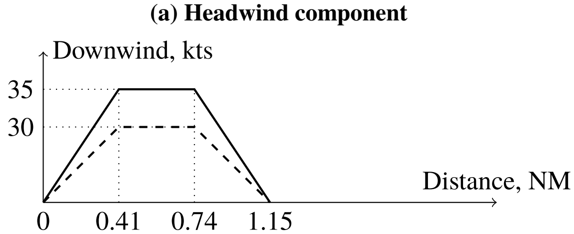
2. Scenario

The pilots were subjected to three scenarios. These were automatically triggered upon descending through a pre-determined 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 take-off wind model from the FAA [20] with wind components as shown in Fig. 13. An approach windshear model was not used because it was found not to ensure that the aircraft would fly near the limits of the flight envelope. In the training runs, the strength of the windshear was reduced while keeping the same distances, as indicated by the dashed lines in Fig. 13.
   In accordance with the FCOM procedure, 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.



(a) Headwind component



(b) Downwind component

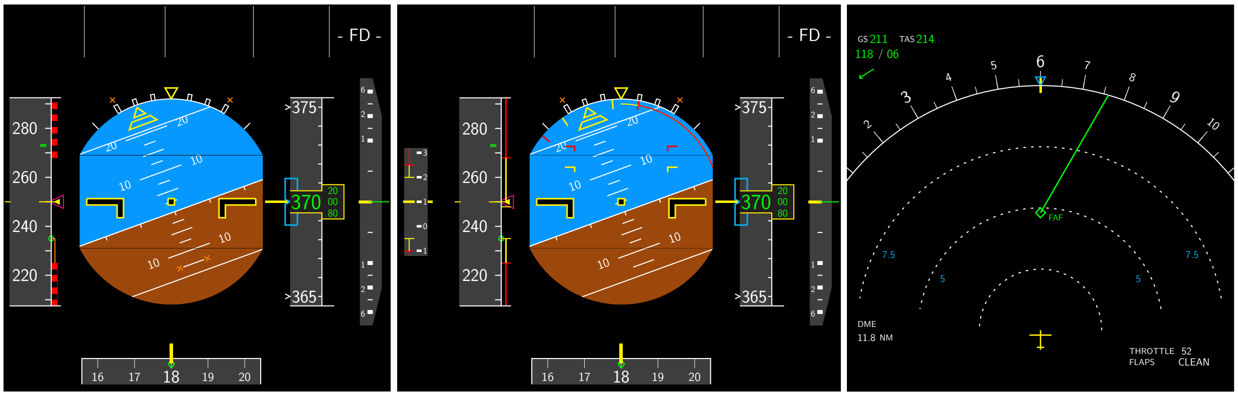
Fig. 13 Windshear model, based on [20]. The dashed profile was used in the training runs.

Table 4 Scenario triggering altitudes.

Scenario	Airport	Triggering altitude			
		ASL, ft		AGL, ft	
		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

3. Display

Two variants of the PFD were used in the experiment (Fig. 14): the original PFD was a replica of the PFD on the real A320, while the new PFD+ had several new indications as discussed in Section III. The A320-like ND was the same throughout the experiment and always showed the final approach fix and runway threshold as waypoints (Fig. 14c). This display also showed the current throttle and flap settings to compensate for the absence of their normal indicators in the simulator.



(a) PFD **(b) PFD+** **(c) ND**

Fig. 14 Displays used in the experiment.

F. Control variables

The aircraft model and haptic feedback settings were the same in all flights. The aircraft had a total mass of 64,841.7 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 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 afterwards computed from this data or handwritten notes on the secondary task.

- Control activity – Root mean square of the stick deflection angle in degrees.
- Margins to flight envelope limits – Both the flight envelope limits and aircraft states were measured in terms of airspeed, angle of attack, load factor and roll angle. The flight envelope margin was computed off-line.
- 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:
  - Windshear – Total altitude loss in feet from start of windshear till lowest point during recovery.
  - Sidestep – 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.
  - 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.
- Workload through 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.

- After each flight:
  - Workload through a RSME questionnaire [21].
  - Situation awareness through two questions on a linear scale ranging from 'Never' (0) to 'Always' (100):
    1) *Did you have the feeling you were in control?*
    2) *Did you have the feeling you missed essential information?*
  - 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 flight envelope through a five-point Likert scale question per item labeled as *not at all, slightly, moderately, very* and *extremely*.
- After both consecutive sets of six flights:
  - System acceptance through Van der Laan rating [22] and Modified Cooper-Harper rating [23].
  - Five-point Likert scale questions on three statements, with labels at the minimum (*disagree*), middle (*neutral*) and maximum (*agree*).
  - Questions on usefulness of individual haptic and display properties in helping the pilot to stay within the flight envelope limits. Five-point Likert scale labeled as *not at all, slightly, moderately, very* and *extremely*.
  - Open question on what haptic cue(s) and/or display element(s) to add to the system, if any.
  - Open question on what haptic cue(s) and/or display element(s) to remove from the system, if any.
- At the end of the experiment:
  - Question on the pilot's display preference (PFD or PFD+) in combination with the haptic system.
  - Five-point Likert scale statements on the haptics, display and experiment with a minimum (*disagree*), middle (*neutral*) and maximum (*agree*) label.
  - Five-point Likert scale question on the safety effect of the system, with a minimum (*unsafer*), middle (*unchanged*) and maximum (*safer*) label.
  - 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 expected to follow risk homeostasis theory (a trade-off between performance and perceived level of risk) [24]. However, 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 flight envelope 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 angle of attack 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 in objective results as shown in Section V.B and subjective results in Section V.C that stem from the questionnaires. Whenever statistical tests are performed, these are Wilcoxon signed-rank tests with a 95% confidence interval, unless explicitly stated otherwise.

A. Data selection

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 crashed their aircraft during the measurement flights by not recovering from a stall upon windshear occurrence. Pilot B2 crashed on the first measured windshear, while pilot B5 crashed on his second windshear. Both where flying with the PFD+ when they crashed and had already experienced two successful windshears in the training flights. B2 indicated after the flight that he did not follow the procedure from the FCOM, but relied on the AoA indication on the PFD+. B5 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 re-started when the pilot (A5) 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.

B. Objective results

All flown tracks for both airports are shown in Fig. 15. The freedom of the pilots to choose their flight path is clearly visible. Some pilots steered away from the localizer to give themselves a smaller intercept angle, while other pilots steered towards the localizer to overfly the FAF while lined up with the runway. While the intercept angle varied per pilot, this was found to be constant irrespective of the display. 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 vastly different approaches in terms of airspeed and corresponding flight envelope limits.

This freedom comes with several challenges for the analysis of the data. For a fair comparison, each flight is therefore cut into two sections 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.

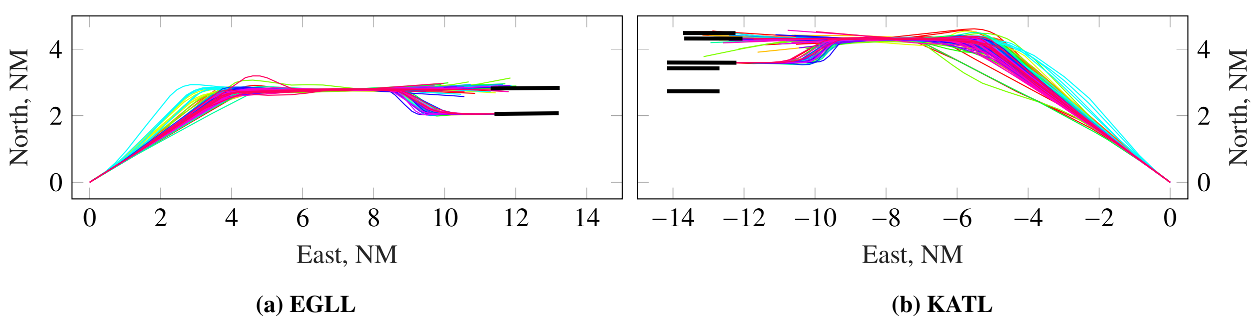


Fig. 15 Flight tracks of all flights combined, colored per pilot.

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. Notably pilots A1 and A5 had much higher velocities, generally this corresponds to a smaller loss of altitude. These pilots are consistently flying fast though, irregardless of the display variant (Fig. 16). 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 provided full back stick upon encountering the shear. One of the pilots explained that he inadvertently thought he was flying in normal law. Third, flap extension time is different per pilot, leading to different performance during the initial approach.

1. Typical data

Figure 18 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 five seconds, 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.

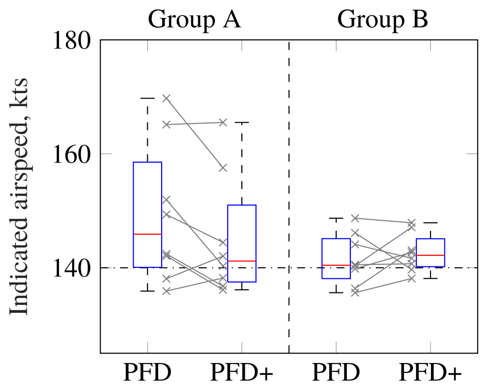


Fig. 16 Mean indicated airspeed per pilot at start of windshear (V_{APP} = 140 kt).

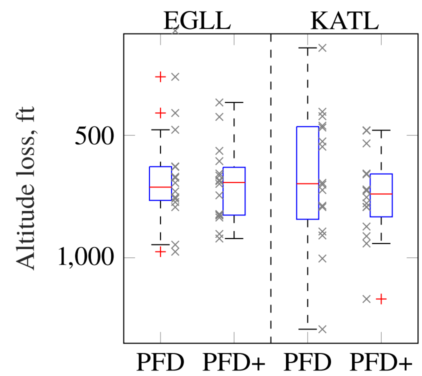


Fig. 17 Mean windshear altitude loss per pilot.

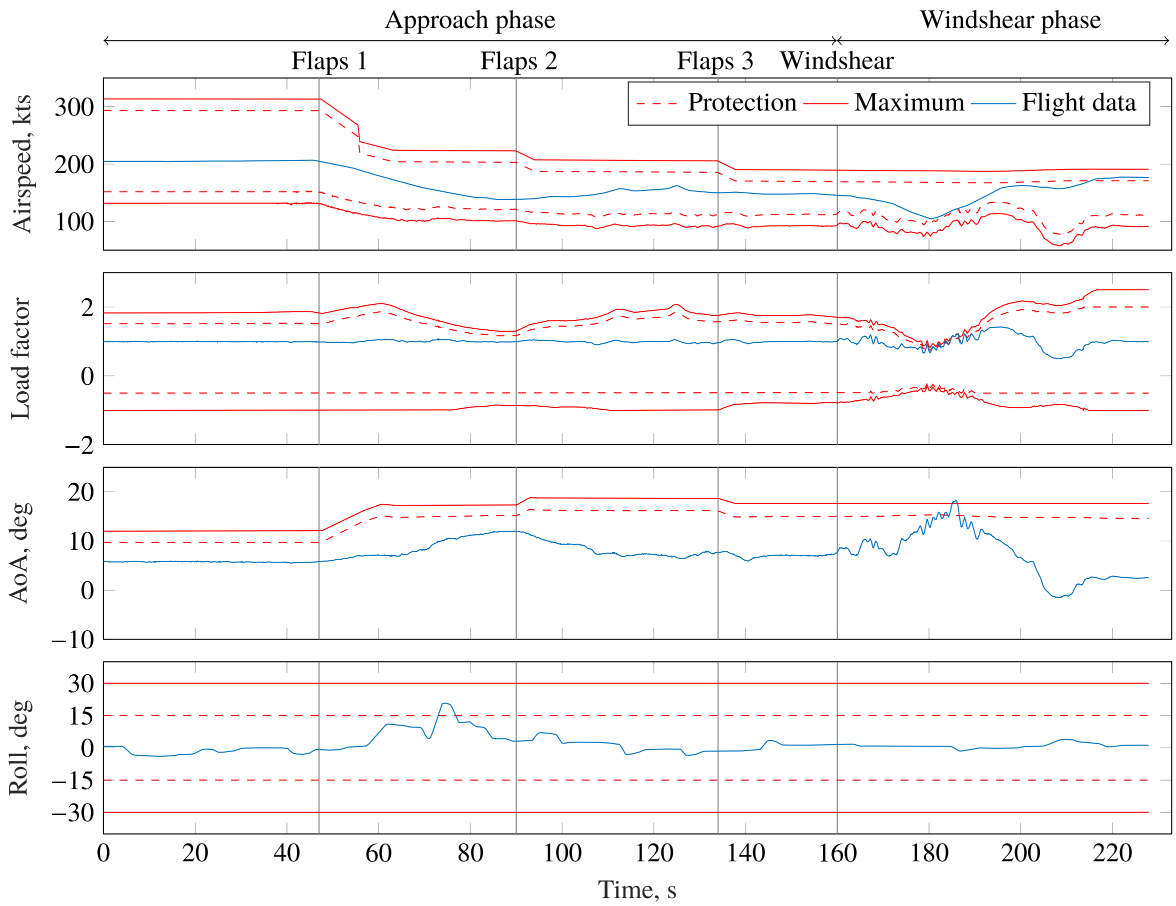


Fig. 18 Typical flight data: pilot A2 flying a windshear scenario at EGLL with the PFD+.

2. Performance

Overall there seems to be little effect of the display on the performance scores, 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. 17), which is not observed at EGLL. The other scenarios only showed marginal effects and thus their data is not visualised here for brevity. At EGLL the PFD+ leads to a slightly lower sidestep score, while at KATL the PFD+ has a higher score. Finally the go-around also shows a small difference, with more low scores at KATL than at

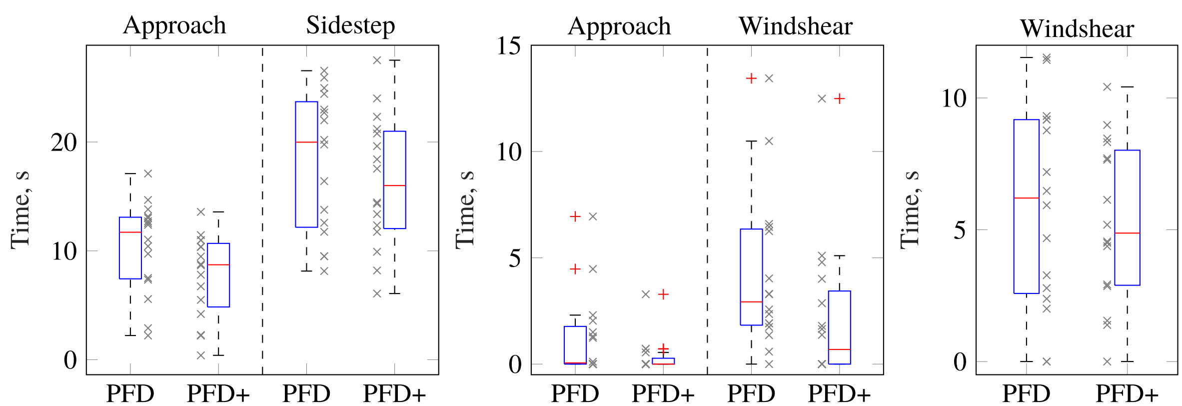
EGLL. Wilcoxon signed-ranks tests show no significant differences for any of the performance scores. Windshear at EGLL (Z = -0.724, p = 0.469) and KATL (Z = -1.293, p = 0.196), sidestep at EGLL (Z = -1.028, p = 0.304) and KATL (Z = -0.159, p = 0.874) and finally the go-around at EGLL (Z = -0.035, p = 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. 19, 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 interception 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 19a 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, $p < 0.01$) while in the sidestep it is not (Z = -1.034, p = 0.301). For the maximum roll limit ($\phi > 30°$), there were too little violations to run a similar analysis.



(a) Roll **(b) Overspeed** **(c) Angle of attack**

Fig. 19 Mean times in protection per pilot. Flight phases in which one or zero excursions into the protection limits were registered are not shown.

Speed excursions were primarily seen during windshear and approach (Fig. 19b). 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 little overspeeds for any statistical analysis. The maximum speed was only exceeded once, during a windshear with the original PFD.

As expected, the angle of attack limits are almost only exceeded during the windshear. Figure 19c shows the time spent above the protection limit. Only one pilot (A1) never exceeded the AoA protection limit. There was a small decrease in time with the PFD+ that is not significant (Z = -0.795, p = 0.427).

5. Control activity

The root mean square (RMS) control deflections of the side stick are given in Fig. 20 and 21 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, even though 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).

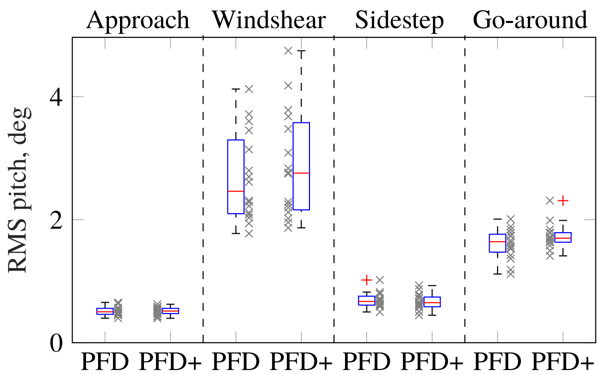


Fig. 20 Mean RMS pitch input per pilot.

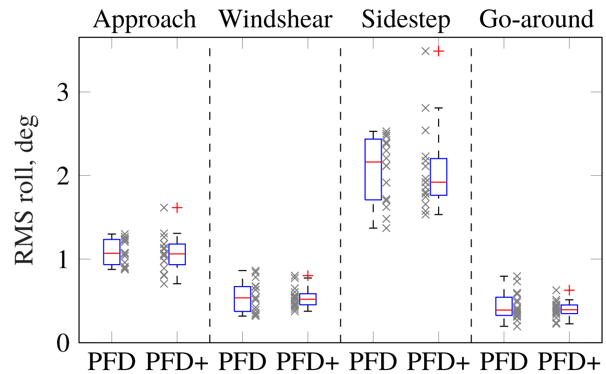


Fig. 21 Mean RMS roll input per pilot.

C. Subjective results

Apart from objective data, subjective results were collected through a series of questionnaires. The results are discussed per questionnaire, starting with the questionnaire 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. Post-run questionnaire

A short questionnaire after each single run allows to see how the display and haptics are experienced in the three scenarios. Figure 22 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 11 of the 16 pilots indicated less lack of essential information in the presence of the PFD+ and for three pilots the display version did not make any difference. Especially the angle of attack 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. 23). 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 they were less in control in the windshear scenario than in the other scenarios.

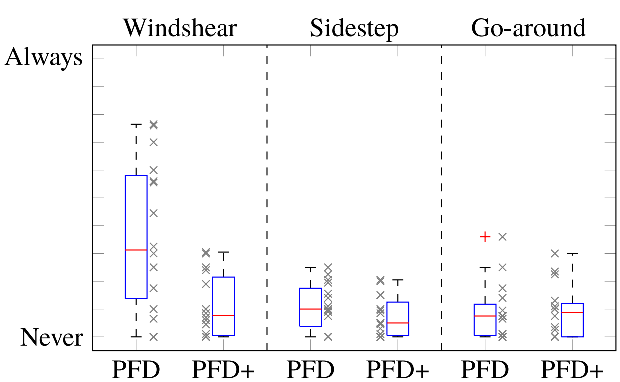


Fig. 22 Post-run question: *Did you have the feeling you missed any essential information?*

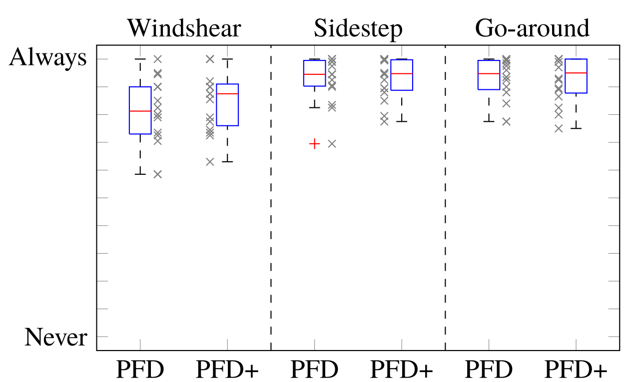


Fig. 23 Post-run question: *Did you have the feeling you were in control?*

In terms of subjective workload, the RSME scores, averaged over the two flights per scenario, show that the pilots perceived the highest effort in the windshear scenario (Fig. 24). 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.

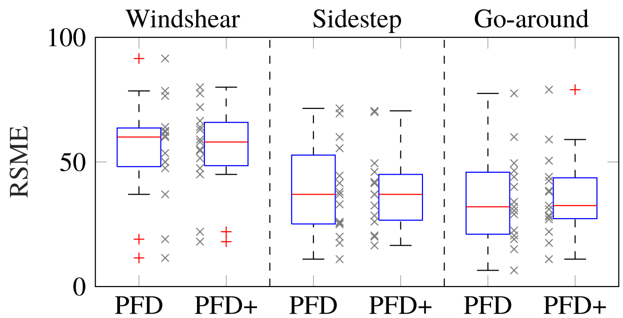
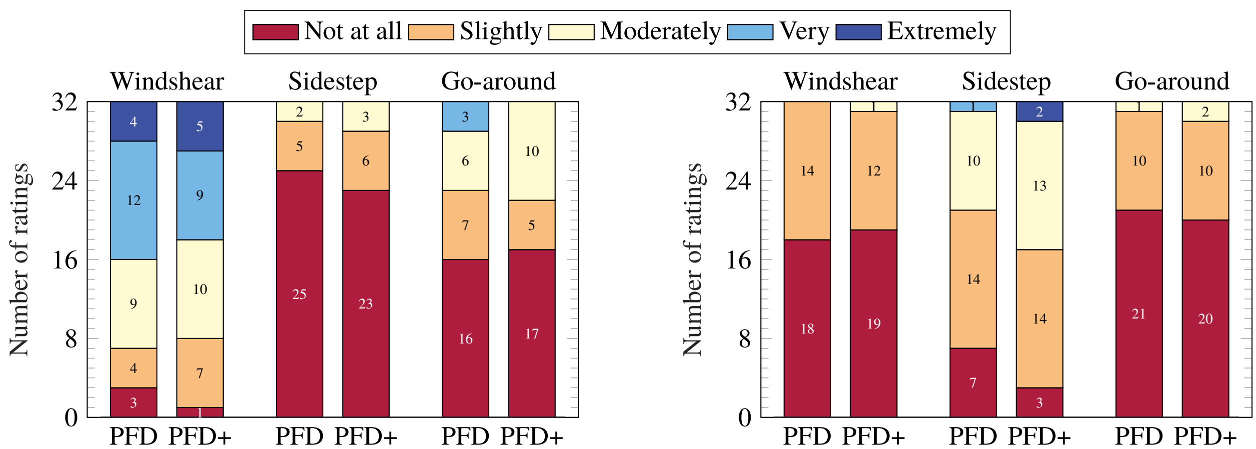


Fig. 24 Mean RSME scores per pilot.

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. 25). 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. For both axes there is no significant change in subjective haptic usefulness between the two display variants.



(a) Pitch

(b) Roll

Fig. 25 Subjective usefulness ratings of the two haptic axes in helping to stay inside the flight envelope limits.

Results of a similar usefulness questionnaire regarding the various display indications are shown in Fig. 26. 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. Post block questionnaire

The Van der Laan ratings, that were collected after six consecutive flights with one of the display options, are shown in Fig. 27 after being averaged per category [22]. The ratings show a small insignificant positive effect of the PFD+ on usefulness (Fig. 27a). No such difference is observed in the acceptance of the system (Fig. 27b). 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 batch, but group B shows a greatly reduced spread with the PFD+, whereas group A does 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. 27c,

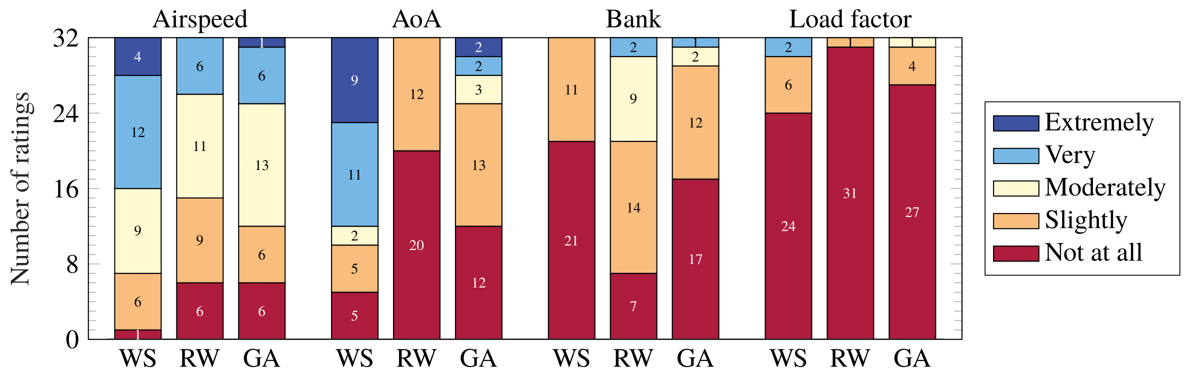
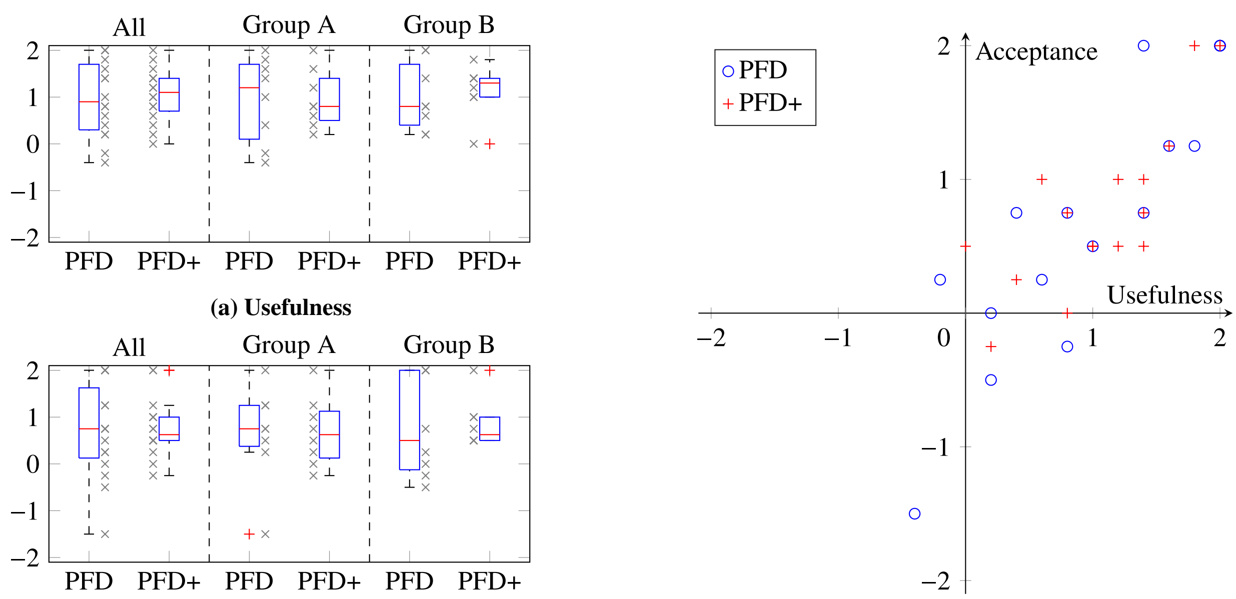


Fig. 26 Subjective usefulness ratings of display elements in helping to stay inside the flight envelope limits.

only two pilots gave the system a negative usefulness rating, both when flying with the PFD. The PFD ratings show a strong correlation between usefulness and acceptance with a Pearson correlation coefficient $\rho = 0.877$, while the correlation is weaker with the PFD+ ($\rho = 0.757$).



(a) Usefulness

(b) Acceptance

(c) Correlation

Fig. 27 Van der Laan Ratings.

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 finger tips'. Furthermore, four pilots that would like to see the tick removed were annoyed by the strict limits in bank. They also considered the tick in pitch a nuisance when extending the flaps brought them above the 20 kts margin towards the maximum speed limit while still below the maximum flap extension speed. The tick itself was said to have the potential of a good attention grabber, as long as the limits are set to realistic values.

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 11 pilots (four from group A, 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 are considered more useful with the PFD+ (Fig. 28). The increasing stiffness and shifts of neutral point stand out with considerably higher ratings. The tick is slightly more useful with the PFD+, while the stick shaker is considered slightly less useful. The number of '*not at all*' ratings for the tick 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. 29). The indication of the protected limit (beyond which the SFE is exited) is considered just slightly more useful than the indication of the maximum flight envelope limit. Despite a slight inclination towards useful there is no clear consensus between the pilots on whether the thickening of the indication is a useful aid in preventing envelope excursions.

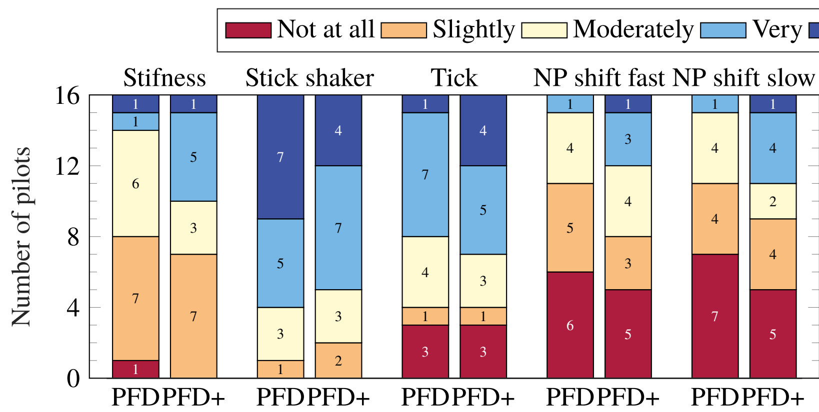


Fig. 28 Usefulness of haptic feedback cues.

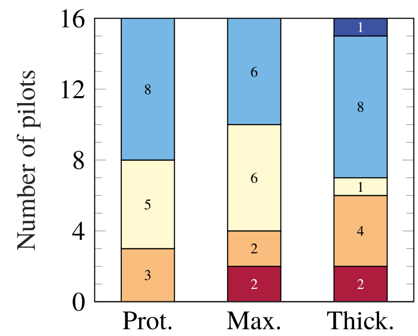


Fig. 29 Usefulness of display cues.

The Modified Cooper-Harper (MCH) ratings in Fig. 30 show little differences between the old and new PFD, except that the spread is less with the new display and there are less ratings of 4 and worse. When looking at the PFD+ ratings for each group separately, it can be observed that the rating is 3 on average for group A, while it is 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?*' A 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*'.

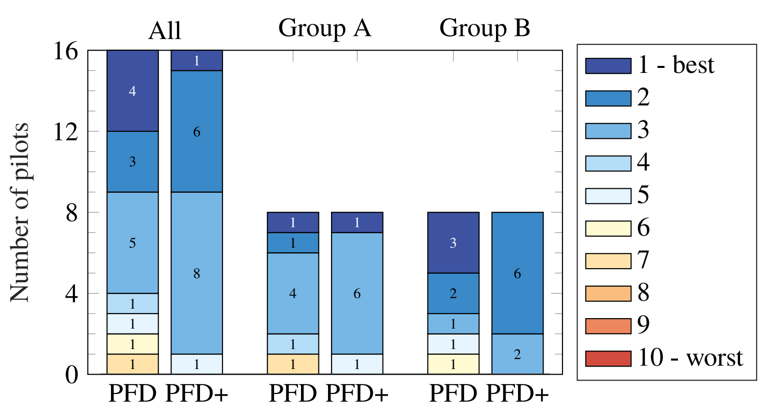


Fig. 30 Modified Cooper-Harper ratings.

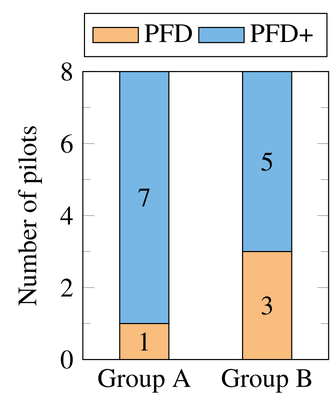


Fig. 31 Preferred display.

3. Post experiment questionnaire

At the end of the experiments, the pilots had to fill in one final questionnaire. When asked which system had their preference, most pilots in both groups indicated that they would like to see the haptics combined with the PFD+ (Fig. 31). In group B the preference is less pronounced than in group A, but there is 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. 32. From the figure a slight 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}, numerous pilots experienced ticks 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 the system does not require lots of training. The vast majority of pilots is 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 implementation of such a system would have a positive effect on safety; only one pilot thought safety would be unchanged. Finally, there is no consensus on whether the display is too distracting. Pilots that 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.

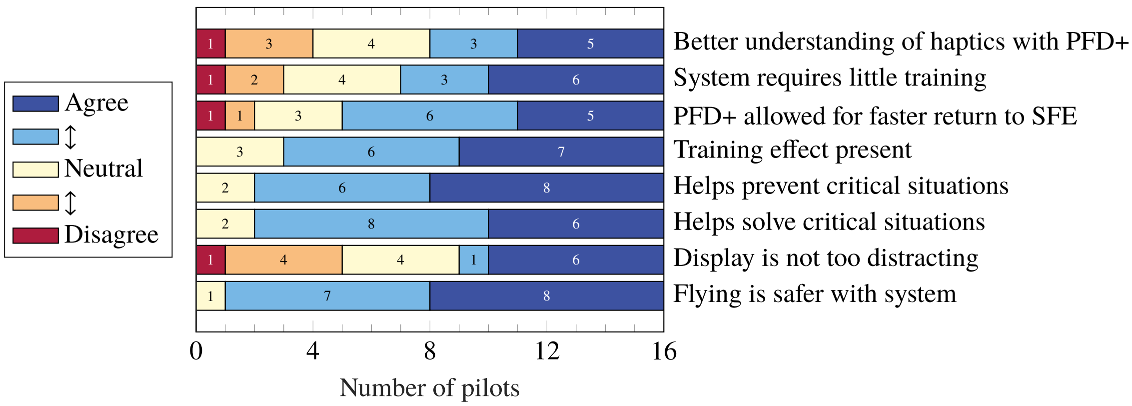


Fig. 32 Subjective post experiment ratings.

In terms of simulation fidelity, all aspects of the experiment are considered acceptable or better by the vast majority of pilots (Fig. 33). Two '*unrealistic*' ratings on flight dynamics and weather were given by pilot A7, who also gave the lowest rating of all pilots on the side stick and ND. The other '*unrealistic*' rating for weather was by pilot A8. 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 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. The projected environment (terrain, airport and sky) was rated acceptable or better by all pilots.

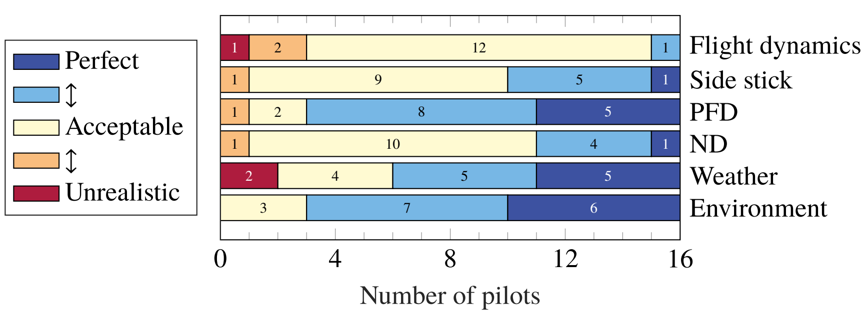


Fig. 33 Subjective simulation ratings.

When taking a closer look at the two – for this experiment – most important simulation elements, the side stick and the PFD, it is clear that both are sufficiently realistic. 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 it was said to be difficult to only apply pitch inputs without inadvertent roll input.

Pilots were in general also very positive about the realism of the PFD, saying 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.

VI. Discussion

Previous research has indicated that adding visualizations to haptic feedback improves user acceptance and possibly also performance and safety metrics [7]. The current experiment indeed shows a slight improvement in acceptance and safety with the newly designed display. It does not, however, show an increase in performance. The following discussion is split in parts that follow the hypotheses. It concludes with the experiment setup and an overall system evaluation.

A. Workload

When it comes to workload, there were only some small changes observed in control activity both in positive and negative direction, depending on the scenario. All changes lacked statistical significance. The secondary task, replying to ATC requests, actually turned out to be unusable for workload analysis due to the small number of ATC requests. 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, however, 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 makes the PFD+ an acceptable addition in terms of workload.

B. Performance

Concerning the go-around scenario, several pilots indicated that it was ‘unusual’ to maintain 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. Limiting the indication to a fixed maximum – similar to Boeing’s pitch limit indicator (PLI) [17] – to prevent excessive pitch (pilot following symbol) may diminish this problem. Another potential source 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 angle of attack margin is zero would solve this ambiguity, while also making it easier 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 clearly communicated. Similar behaviour 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. In order to ensure 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 this was unrealistic and perceived the roll cues in the haptic system and display as a nuisance since 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. Nevertheless pilots did respect the bank limits more when shown on the PFD+. The hypothesis H3 that the margin to the flight envelope boundaries would become larger with the PFD+ can, however, not be accepted 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 rather limited, a large share of the pilots does expect that the system would improve safety when implemented. The data do not provide an answer on whether that can be attributed to the haptics, the display or the combination of both. Previous research does suggest 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 old PFD, suggesting an improved acceptance of the haptic system in combination with the new display. This is confirmed by the increased usefulness of the various 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 bigger impact on their system-wide 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, 25]. Based on the preceding, 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 not included in the scanning pattern. 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 fit in the standard Airbus display size.

The AoA indication on the other hand was much more appreciated by the pilots. The only pilot that said to remove it, does like to see the AoA and load factor indications in certain critical situations, like windshear or terrain avoidance maneuvers. Although it did not bring the expected performance benefit, it gave 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.

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 that did fly with it before did so over a year earlier. Even though the pilots received considerate 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. This could have helped determine whether any changes are caused by the haptic feedback itself, or the display.

In the aim for realistic scenarios, pilots were given a lot of freedom which lead to challenges in the data analysis. It could help to limit this freedom in future experiments. For example by showing the route on the ND all the way from the start, instead of from a distant waypoint onward. Using the autopilot to bring the aircraft to a pre-determined state and hand-over control to the pilot on the onset of an event may also help and is an accepted method in flight training [26]. 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 motion may have influenced pilot behavior. Especially in stall conditions, pilots are known to over-react when they do not feel the load factor [27]. 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 can not be assumed to be an adequate replacement. An experiment involving haptic feedback, PFD+ and motion cueing should be conducted to see whether motion has any effect.

G. Overall system evaluation

Wrapping up, the PFD+ brings no big improvements nor any large deteriorations. Since pilots seem to like the display, albeit with a couple of modifications, the integration of input and output space 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 flight envelope boundaries. At least in approach scenarios, while a substantial number of LOC-I accidents in recent years occurred during cruise. Testing the haptics and display in a cruise situation where the pilots suddenly find themselves in alternate law could show the potential of the system in a wider range of flight phases.

VII. Conclusion

This study looked into the effect of providing a visual display as a complement to haptic feedback in communicating flight envelope boundaries. The resulting display design is unique in displaying not only the limits of the flight envelope or the limits of the control inputs, but both in one display. In addition to that, the display shows the direction and force of the haptic feedback that is applied to the side stick. To accomplish this, the standard A320 PFD has been enhanced with new indications for angle of attack, airspeed, bank angle and load factor. The display was evaluated by inviting 16 professional Airbus pilots to TU Delft's SIMONA research simulator, where they flew two approaches in each of three different scenarios with both the original and modified PFD.

Unlike hypothesized, the design presented in this article 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. On the subjective front, the display proved to result in a small increase in pilot appreciation of haptic feedback. The display increased the pilots' understanding of what the haptics were trying to communicate and helped pilots stay within the limits of the SFE.

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. On the opposite, especially the AoA indication appears to be useful, but also lead to a number of crashes when pilots followed it too closely. Further research is therefore suggested to improve this particular indication and reduce its ambiguity. It is also recommended to test the display with the actual bank limits, instead of the reduced ones used in this research.

Acknowledgments

We would like to express our gratitude to the 16 pilots that participated, as well as those that 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 – 2016," Tech. rep., Boeing Commercial Airplanes, Jul. 2017.

[2] CAST/ICAO Common Taxonomy Team, "Aviation occurence categories - Definitions and usage notes," , Dec. 2017.

[3] "Aircraft Accident Investigation Report PK-AXC," Tech. rep., Komite Nasional Keselamatan Transportas, 2015.

[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," Tech. rep., BEA, Jul. 2012.

[5] Van Baelen, D., Ellerbroek, J., van Paassen, M. M., and Mulder, M., "Design of a Haptic Feedback System for Flight Envelope Protection," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Kissimmee (FL)*, 2018, p. 16. doi:10.2514/6.2018-0117, AIAA-2018-0117.

[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, San Diego (CA)*, 2019, p. 25. doi:10.2514/6.2019-0367, AIAA-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," *IEEE International Conference on Systems, Man, and Cybernetics*, 2018.

[8] Meuleau, N., Neukom, C., Plaunt, C., Smith, D. E., and Smith, T., *The Emergency Landing Planner Experiment*, NASA, 2011.

[9] Rijndorp, A. D. T., Borst, C., de Visser, C. C., Stroosma, O., Mulder, M., and van Paassen, M. M., "Aviate, navigate: Functional visualizations of asymmetric flight envelope limits," *AIAA Information Systems-AIAA Infotech at Aerospace*, 2017.

[10] Lombaerts, T., 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. doi:10.2514/1.G001740.

[11] Lombaerts, T. J. J., Looye, G., Ellerbroek, J., and y Martin, M. R., "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. doi:10.2514/1.G002525.

[12] Ackerman, K. A., Seefeldt, B. D., Xargay, E., Talleur, D. A., Carbonari, R. S., Kirlik, A., Hovakimyan, N., Trujillo, A. C., and Gregory, I. M., "Flight envelope information-augmented display for enhanced pilot situation awareness," *AIAA Infotech at Aerospace*, 2015.

[13] *A318/A319/A320/A321 Flight Crew Operating Manual*, Airbus, 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," Tech. rep., FAA, Apr. 2002.

[15] Wickens, C. D., Lee, J., Liu, Y., and Becker, S. G., *An Introduction to Human Factors Engineering*, 2^{nd} ed., Pearson Prentice Hall, 2004.

[16] "Electronic Flight Displays, Advisory Circular 25-11B," Tech. rep., FAA, Jul. 2014.

[17] "Operational use of Angle of Attack on Modern Commercial Jet Airplanes," *Aero*, edited by J. Langer, 12, Boeing Commercial Airplanes Group, 2000.

[18] van Paassen, M. M., Stroosma, O., and Delatour, J., "DUECA - Data-Driven Activation in Distributed Real-Time Computation," Modeling and Simulation Technologies Conference, 2000, pp. 1–7. doi:10.2514/6.2000-4503.

[19] OpenStreetMap contributors, "OpenStreetMap maps," `https://www.openstreetmap.org`, 2019.

[20] "Windshear Training Aid," Tech. rep., Federal Aviation Administration, 1990.

[21] Zijlstra, F. R. H., "Efficiency in work behaviour: A design approach for modern tools," Ph.D. thesis, Delft University of Technology, Nov. 1993.

[22] 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, 1997, pp. 1–10. doi:10.1016/S0968-090X(96)00025-3.

[23] Cummings, M. L., Myers, K., and Scott, S. D., "Modified Cooper Harper Evaluation Tool for Unmanned Vehicle Displays," *Proceedings from UVS Canada: Conference on Unmanned Vehicle Systems Canada '06*, 2006.

[24] 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. doi:10.1080/00140138808966691.

[25] Ellerbroek, J., Rodriguez y Martin, M., Lombaerts, T., 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. doi: 10.1016/j.ifacol.2016.10.481, 13th IFAC Symposium on Analysis, Design, and Evaluation of Human-Machine Systems (HMS 2016).

[26] *Airplane Upset Prevention & Recovery Training Aid for Transport Category Airplanes*, ICAO, Feb. 2017. Rev. 3.

[27] Le Vie, L., "Survey of Quantitative Research Metrics to Assess Pilot Performance in Upset Recovery," Tech. rep., NASA, Apr. 2016.