JOURNAL OF AEROSPACE INFORMATION SYSTEMS
Vol. 18, No. 12, December 2021

Aviate, Navigate: Functional Visualizations of Asymmetric Flight Envelope Limits

Tom A. D. T. Rijndorp,* Clark Borst,† Coen C. de Visser,‡ Olaf Stroosma,§ Marinus M. van Paassen,¶ and Max Mulder**
Delft University of Technology, 2629 HS Delft, The Netherlands

https://doi.org/10.2514/1.I010895

Current aircraft flight deck interfaces do not provide information on how a performance-altering failure constrains an aircraft's flight envelope. As a result, it is difficult for flight crews to plan maneuvers in order to reach navigation targets. This study presents the results of the conceptual development of constraint-based interface symbology that aims to address this issue. The proposed symbology is designed to integrate with both the primary flight display and navigation display. A small-scale, pilot-in-the-loop experiment ($N = 9$) was conducted to assess the effectiveness of the used symbology in terms of flight performance and pilot usability. A simplified dynamic model with an asymmetric flight envelope was used to purposefully manipulate various levels of damage severity and corresponding flight envelopes. Results show that although the modifications to the primary flight display generally did not show statistically significant improvements, presenting flight envelope constraints as a reachable navigation envelope on the navigation display generally did do so for severe failures. The visualized envelope occasionally resulted in improved tactical control decisions at reduced workload levels. A future study involving a lager sample size and increased simulation realism should substantiate the discovered results.

I. Introduction

TODAY, the task of flying an aircraft is predominantly done by means of interacting with the on-board automated systems or autopilot. In certain events such as engine failures, control surface failures, or structural damage, pilots will need to re-evaluate the conditions under which they are operating. They may need to determine a new mission objective and adjust their control strategy accordingly.

The pilot's mantra here is "aviate, navigate, communicate" [1], summarizing the tasks that need to be performed in their order of prioritization. That is, first, make sure that the aircraft regains and maintains stable flight. Then, derive a strategy to reach the (newly formulated) destination state, and only then inform other stakeholders of the issue at hand. In commercial aviation, "aviate" and "navigate" tasks in failure conditions are usually allocated to the captain and first officer, respectively.

Performance-altering events cause changes in the safe flight envelope, making maneuvers that were safe to execute with a nominal flight envelope, dangerous or even impossible [2–4]. This complicates a flight crew's ability to both aviate and navigate. A direct result of changes in the safe flight envelope is that the flight crew may become unaware of the limits of the impaired flight envelope. This poses challenges in terms of maintaining control of the aircraft, but also in understanding implications on higher-level goals such as trajectory planning and landing.

When control is regained after an initial upset, the issue that remains is finishing the (new) mission objective in the safest way possible. However, being able to *fly* an aircraft does not necessarily imply that one is able to successfully *navigate* an aircraft to a target state as well [5]. In the presence of performance-altering failures, like those that result in asymmetric left and right turn maneuverability, control actions need to be carefully planned and executed to prevent one from leaving recoverable regions of the flight envelope [2]. Note that in this paper, the scope of navigation is limited to the planning and execution of (a sequence of) control inputs that bring the aircraft into a desired state toward a navigational target (e.g., make a left or right turn, at a certain roll angle, to align the aircraft with the runway localizer plane).

Existing research contains proposals for flight deck interfaces such as fully automated emergency replanning and landing systems [6], planner modules that involve the pilot to select the best candidate from a set of trajectories defined by the automation [7–11], as well as interactive planners that require the pilot to interact with the planner module by iteratively adjusting and reviewing a landing plan [12]. The benefit of such automated solutions is that they alleviate the pilot's cognitive effort required to formulate a solution, thereby freeing up valuable cognitive resources for completing other tasks. As articulated in the Ironies of Automation [13,14], the downside of such an approach is that pilots may experience difficulty in 1) *understanding* why the system proposed a specific solution, 2) *judging* the validity of the solution, and 3) *hypothesizing* what alternative solutions exist. Especially in abnormal situations unanticipated in the design of the automation, pilots are generally regarded as creative and adaptive problem solvers who can save the day under such difficult situations, as witnessed in the Hudson River landing [15]. To facilitate adaptivity and creativity within safe bounds, pilots require proper information.

In recognition of these human–automation interaction issues, research is also considering providing visual (and even haptic [16]) information to pilots about flight envelope limitations using augmentations on existing flight displays [2,17–19]. These studies go beyond the visual cues that portray flight envelope constraints typically found on current operational flight displays (e.g., stall speeds and maximum speeds marked on the speed tape). Instead of additional interface enhancements that present single optimized solutions that pilots are expected to follow (e.g., flight director guidance), visualizations that present boundaries and margins for safe control actions aim to give pilots more flexibility in deciding the best course of action. The majority of such visualizations focus on supporting either control

Presented as Paper 2017-1297 at the AIAA Information Systems-AIAA Infotech @ Aerospace, Grapevine, TX, January 9–13, 2017; received 3 September 2020; revision received 28 May 2021; accepted for publication 20 June 2021; published online 10 August 2021.

*M.Sc., Section Control & Simulation; tomrijndorp@gmail.com.
†Assistant Professor, Section Control & Simulation; c.borst@tudelft.nl.
‡Assistant Professor, Section Control & Simulation; c.c.devisser@tudelft.nl.
§Researcher, Section Control & Simulation; o.stroosma@tudelft.nl. Senior Member AIAA.
¶Associate Professor, Section Control & Simulation; m.m.vanpaassen@tudelft.nl.
**Professor, Section Control & Simulation; m.mulder@tudelft.nl. Associate Fellow AIAA.

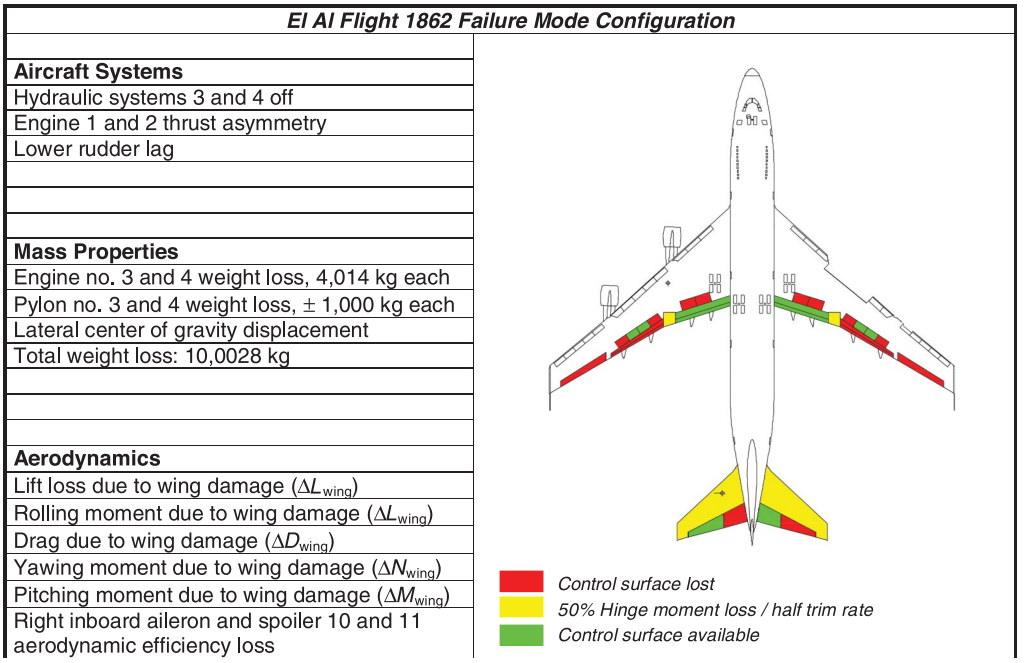

Fig. 1 El Al Flight 1862 failure mode [24].

(i.e., maintain stability) or navigation (i.e., long-term route planning) tasks in both nominal and off-nominal flight conditions. However, as "aviate" and "navigate" tasks take place at different time scales, it remains unclear how to best support pilots in situations involving asymmetric maneuvering limits in roll and yaw axes. In such situations, it may be important to communicate information that allows pilots to relate the impact of critical short-term maneuvering information (e.g., attainable roll rates and angles) on medium-term reachable navigation states (e.g., ground-referenced flight directions and positions). In this work the definition of the maneuvering envelope from [3,20,21] is used, which can be analytically derived using nonlinear reachability analysis given a model of the system dynamics.

The goal of this paper is to introduce new visualizations geared toward providing critical aviate and navigate information resulting from asymmetric flight envelopes. Here, the emphasis lies on detailing the visualizations and gaining empirical insights into their usefulness in terms of both flight and human performance measures by means of an exploratory human-in-the-loop experiment. For this purpose, a simplified asymmetric maneuvering envelope in the roll and yaw axes, resulting from asymmetric dynamic aircraft behavior, was developed that allowed us to purposefully manipulate various levels of damage severity. This work therefore may complement previous and similar research efforts (e.g., [22,23]) by providing another design example, along with empirical insights, that can guide the aviation industry in developing more informative pilot-support displays. This paper is structured as follows: Section II describes the theoretical motivation underpinning this study. Section III presents the dynamic model that was used along with its design considerations, and Sec. IV describes the interface development process. Section V details the design of human-in-the-loop experiment that was conducted, and Sec. VI describes the experiment's results. Discussion, recommendations, and conclusion sections complete this paper.

II. Theoretical Motivation

A. Flight 1862

This research is motivated in part by El Al Flight 1862 (Flight 1862), a Boeing 747-200 freighter flight in which the two starboard engines detached from the wing shortly after takeoff [5]. The separated engines also inflicted damage upon the starboard wing's leading edge and hydraulics systems, rendering multiple control surfaces (among which the outboard ailerons) inoperable. A schematic overview of the failure mode is shown in Fig. 1.

The original investigation mentioned that at higher angles of attack, the difference in lift between the left and right wings would increase as a result of the right wing leading edge damage. Therefore, increasing the angle of attack would generate a right-wing-down rolling moment [25].

At the final stage of the flight, the angle of attack was increased, most likely in order to reduce the descent rate. This led to a further increase in drag on top of the already higher-than-nominal drag caused by the sideslip angle, as well as an additional decrease in speed. This probably caused the pilots to increase the engine thrust. This series of events led to a large right-wing-down rolling moment due to the asymmetric lift distribution, the asymmetric thrust, and reduced right-hand inboard aileron effectiveness. At this point, control of the aircraft was lost [25].

Inspection of Flight 1862's Digital Flight Data Recorder (DFDR) data indeed revealed that yoke deflections of up to 60 deg out of 80 deg were required in order to maintain straight flight [26]. As a result of the aileron deflection limits, roll-rate performance suffered. An independent investigation by the Netherlands Aerospace Laboratory (NLR) and Delft University of Technology (TU Delft) revealed that, in hindsight, control could have been maintained by maintaining a higher airspeed, resulting in more control authority and higher maneuverability [26].

While Flight 1862's failure mode posed a significant challenge for the flight crew in terms of maintaining control, the failure also heavily affected the aircraft's maneuvering and navigation capabilities. The flight crew had no other means for developing an understanding of these capabilities besides observing the aircraft's response to their control inputs. In fact, one of the decisions that stood out was that the crew decided to land on Runway 27[††] with a tail/crosswind approach instead of the runway in use, Runway 06. Runway 06 offered a headwind approach and was likely easier to navigate to.

We may never be able to fully understand all factors in the decision-making process of the flight crew, and therefore we do not know whether the best possible decision was made in terms of runway selection and navigation planning. With the available avionics (displays), the information available to the flight crew was limited, creating limitations for diagnostics and for understanding how the aircraft's *functionality* would be impacted. That is, no explicit information was available on how the aircraft's aviate and navigate capabilities were affected by the failure. When the outcome of an

[††]At Amsterdam Airport Schiphol, The Netherlands.

action is uncertain, it is difficult to plan ahead. It is our expectation that, in the event of a failure, providing pilots with a means to understand an aircraft's limited capabilities from a functional (i.e., action-relevant) perspective may improve the tactical planning process.

B. Developments in Enabling Technology

To develop a system that presents information on attainable (navigation) states, knowledge of the flight envelope is required. Unfortunately, estimating the flight envelope of a damaged aircraft even in an offline setting is far from trivial [3]. Although much research is currently being done on fault detection and isolation (FDI) and postfailure flight envelope estimation [4,27–36], fully functioning online systems do not yet exist. One reason for this lack of existing implementations is the high dimensionality of the state space and the corresponding computational complexity. A second, more fundamental, reason is that all reachability-based envelope prediction methods require an accurate *global* aerodynamic model, which in general is not available after a failure event. A promising solution to this issue in the form of a database-driven envelope prediction method was recently introduced by Zhang et al. [4].

The primary driver for research in flight envelope prediction and estimation is the field of fault-tolerant flight control (FTFC). The main goal of FTFC is to improve the controllability of aircraft after a system fault has occurred or damage has been sustained. Fault-tolerant control systems are, to some extent, a double-edged sword. As these systems simplify the control task, potentially by seamlessly blending different control effector inputs in such a way that aircraft control is natural and predictable [37], it will be harder for a pilot to observe the amount of remaining control authority and thereby how safe a certain state or maneuver actually is. For this reason, flight envelope prediction is an important field of study as it enables flight envelope protection (FEP) under off-nominal flight conditions, preventing pilots from operating outside the more constrained safe flight envelope.

However, even with a fault-tolerant, flight-envelope-protected control system, but without an understanding what the off-nominal flight envelope implies in terms of maneuvering and navigation capabilities, pilots may still not be able to formulate a safe navigation strategy with a damaged aircraft. In an effort to communicate these impaired aircraft capabilities to pilots, this research offers a way to convey such information in a visual manner. As such, it aims to bridge the gap that exists between being able to control an aircraft with off-nominal performance and being able to understand how this performance reduction affects current and subsequent maneuvering capabilities.

III. Model Development

Testing these new visualizations in a human-in-the loop experiments generally requires a dynamic aircraft model with an off-nominal flight envelope. Additionally, the model's flight envelope should be known. It was mentioned that fully operational flight envelope prediction systems do not yet exist. And when the resulting aircraft dynamics are difficult to control, a fault-tolerant control system may have to be added: in a study on emergency landing planning by Meuleau et al. [11], it was found that adding a fault-tolerant controller was a prerequisite for effective use of the planner interface, as not having a fault-tolerant controller can render the control task prohibitively difficult (as the pilot needs to allocate cognitive resources to remain airborne).

Figure 2 shows the control loop of a hypothesized future aircraft system including a flight envelope estimation module, fault-tolerant control system, the pilot, and the (graphical) interface. It can be observed that the pilot is interacting with the flight control system (FCS), which in this case is assumed to be fault tolerant. The fault or failure that alters the flight envelope of the aircraft also impairs the pilot's trained ability to estimate reachable aircraft states. To this end, a fault detection and diagnosis (FDD) module coupled with a flight envelope estimation module may assist the pilot in interpreting the new, off-nominal flight envelope when this information is presented on the electronic flight instrument system (EFIS) screens, as, for example, demonstrated in [2]. However, it should be noted that online flight envelope prediction is an active research field. Such systems are not yet operational, and those proposed are complex; see, e.g., [4] for a proposed online envelope prediction and protection system architecture.

As it is generally not possible for a pilot to discriminate between performance limitations imposed by the FCS and those resulting from the aircraft's dynamics, the approach chosen in this study is to limit model complexity by lumping together the flight envelope estimation, FDD, FCS, and dynamics blocks, resulting in a simplified plant with equivalent dynamics that can output data to the interface and accept pilot control inputs. This simplified, combined system integrates the aircraft's dynamics and returns flight envelope data, as is depicted as "Combined system" in Fig. 2.

Although aircraft maneuvers in the vertical plane are relevant to this field of research, as a first step, the focus of this study is on the lateral maneuvering capabilities and the control strategies pilots develop as a result of the remaining performance. As such, the simplified model did not need to consider vertical motion in this first exploratory study.

To constrain the flight envelope, a consequence of Flight 1862's wing damage was considered. An asymmetric lift distribution requires compensating aileron deflections in order to maintain equilibrium. This results in a nonzero aileron deflection to maintain straight flight. A model with these properties was obtained, starting from the GARTEUR RECOVER Boeing 747 model developed by Smaili et al. [39]. First, the model was trimmed for straight and level flight for a given flight condition and aircraft configuration. Next, the model was linearized, resulting in a set of continuous-time state space matrices. Then, new system and input matrices were formed using only the lateral states p (roll rate), ψ (heading angle), r (yaw rate), β (side slip), and ϕ (roll angle) and the aileron input δ_a [40]. Assuming the presence of an effective yaw damper and turn coordination system in the FCS, the system can be further simplified to the extent that only the roll subsidence eigenmode is retained, resulting in a simple transfer function that describes the roll behavior of the aircraft [41].

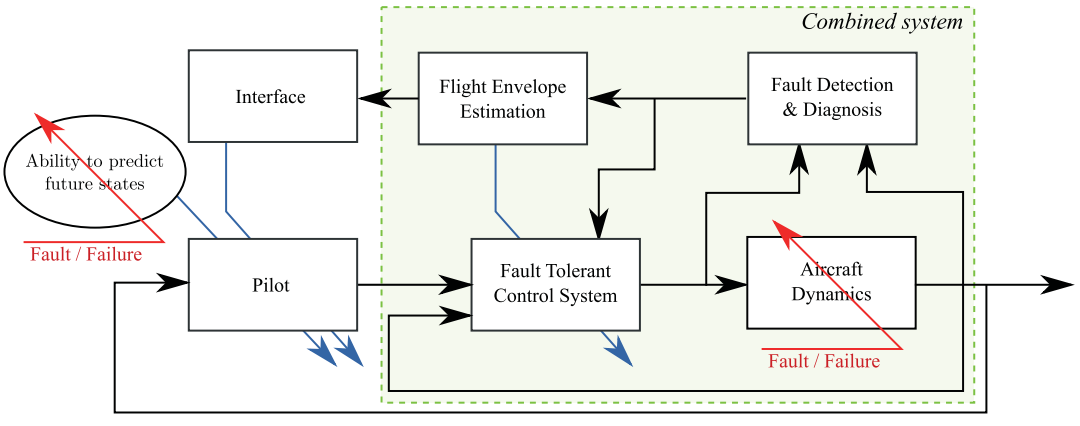

Fig. 2 The control loop including pilot and interface (adapted from [38]).

The model structure can be seen in Eq. (1). The gain K_p and time constant τ_p depend on the airspeed V. Speed-dependent values for these parameters can be derived by performing the procedure outlined above for multiple trim points.

$$\begin{aligned}\dot{p}(t) &= -\frac{1}{\tau_p(V)}p(t) + \frac{K_p(V)}{\tau_p(V)}\delta_a(t)\\ \dot{\phi}(t) &= p(t)\\ \dot{\psi}(t) &= \frac{g_0}{V(t)}\phi(t)\\ \dot{x}^E(t) &= V(t)\cos\gamma\cos\psi(t)\\ \dot{y}^E(t) &= V(t)\cos\gamma\sin\psi(t)\end{aligned} \tag{1}$$

With this result, a simulated failure can be introduced with a controllable severity, characterized by the aileron or yoke deflection required to maintain a zero roll rate. To obtain an aircraft model that is still controllable by a human in a simulation study, this value should be chosen in between the left and right control effector's saturation limits. The larger the magnitude of this value, the more asymmetric the roll behavior of the aircraft will be.

Now, two approaches can be taken to modify the navigation performance (i.e., ability of the aircraft to intercept and follow a path) resulting from this simplified model:

1) One can increase the airspeed, thereby reducing the aileron deflection required for steady flight, resulting from the lower angle of attack. This would represent the *same* damage case at a *different* airspeed.

2) One can keep the airspeed constant, but alter the required aileron deflection. This would represent a *different* damage case at the *same* airspeed.

Irrespective of the approach taken, the result is that the larger the steady-state aileron offset gets, the more asymmetric the aileron authority becomes. This results in lower attainable roll rates in one direction, and higher rates in the other (note that for the nominal case, the figure would have rotational symmetry about the origin). The resulting lateral maneuvering envelope is shown conceptually in Fig. 3. For this study, the shaded area of safe combinations of roll rate and roll angle can be considered the safe flight envelope.

Figure 3 conceptually illustrates the set of safe roll rates given the current roll angle. As clockwise (positive) roll rates move away from the left roll angle limit, these rate/angle combinations are safe for both negative roll angles as well as slightly positive roll angles (marked ① and ②, respectively). Using the same reasoning, the same holds true for negative rates at positive angles and slightly negative angles. In the top right corner (see ③), it can be seen that there are combinations of roll rates and roll angles that cause the roll angle limit to be exceeded. These combinations should be avoided. Equivalently, the same holds for the combinations in the lower left corner (marked ④). However, because the attainable roll rates at ④ are lower and the opposite aileron authority is higher, these rates can be controlled back to zero faster. Lastly, we may trim the controls by offsetting the stick input with the stick input that yields a zero roll rate, effectively simulating the actions of a stabilizing fault-tolerant controller. In many failure cases a significant offset input is required on the stick roll command to reach zero roll rate; this is visible in the asymmetry of positive and negative roll rate limits.

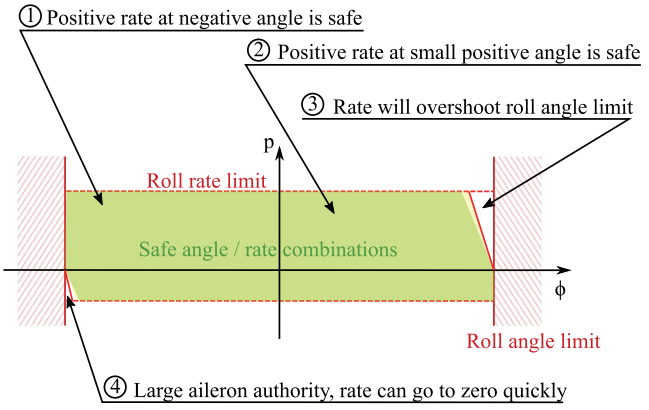

Fig. 3 Relation between safe roll rates and roll angles at a given airspeed for a simplified model with asymmetric aileron authority.

The result from Fig. 3 mainly explains how wing damage may impair the lateral maneuvering capabilities. However, it does not yet clarify how reduced maneuvering capabilities become relevant to the higher level goal of *navigating* an aircraft. To this end, Fig. 4 is introduced. This figure conceptually illustrates how the lateral aircraft states relate to one another. This representation is assumed to be similar to a pilot's mental model of how lateral aircraft dynamics function. The representation is also similar to the simplified aircraft model that was introduced in this section. The exact relationship between control surface authority and reachable navigational states over time is complex, and even more so in the light of a performance-altering failure. This is why system faults and failures like those experienced by Flight 1862 and other flights with wing damage and/or aileron failures are hard to cope with: suddenly, the aircraft's dynamics are different from the dynamics the pilot became familiar with during numerous hours of flight training and accumulated experience on a particular aircraft class. As a result, the set of reachable navigational states becomes uncertain.

For the simplified model described in this section, the exact relationship between aileron authority and reachable states can be found by integrating the set of differential equations in Eq. (1). We provide more detailed graphical representations in the following section.

IV. Interface Development

A. Design Requirements

As the goal of the interface symbology presented in this work is to provide pilots with an understanding of an impaired flight envelope, and how that impacts various safe aviate and navigate states, it was chosen to consider a constraint-based approach that visualizes boundaries and margins. Figure 5 shows an abstract representation underlying such an approach, where the idea is that the physical system boundary determines the safe action–state space in which *any* action that reaches the target state is acceptable, except those that will cross the boundary (and thus lead to accidents). Note that this view is inspired by the underpinnings of ecological interface design [42], which has been successfully applied in various application domains involving dynamic control tasks [43,44].

Translating this concept to the problem at hand, the interface must provide perceivable information on the initial and goal states, the dynamic boundary for safe control actions (i.e., the flight envelope limits), and a preview (incorporating a certain look-ahead time) of where the current action will lead. Note that the boundaries illustrated in Fig. 5 are dynamic, meaning that actions can deform, shrink, and/or expand the space of possibilities.

The more complete this information is, the better pilots will be supported in planning safe actions. The challenge is, however, to create a design that balances interface complexity (e.g., clutter) against usability (e.g., workload). This is especially important when we want to support pilots in hazardous situations that require swift pilot responses.

Given the above-mentioned requirements and the expected operational conditions, it was decided to augment an existing

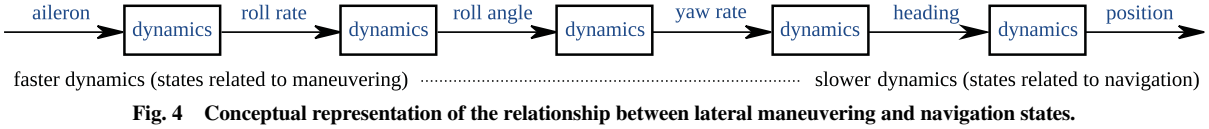

Fig. 4 Conceptual representation of the relationship between lateral maneuvering and navigation states.

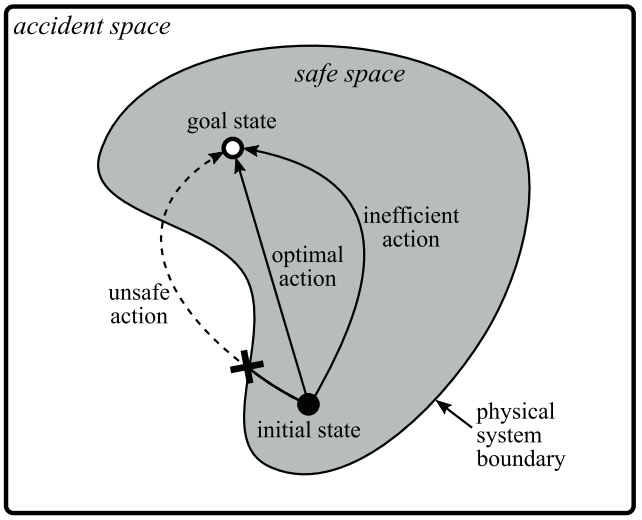

Fig. 5 Abstract representation of the safe and unsafe control spaces.

primary flight display (PFD) and navigation display (ND) instead of adding a new interface to the limited display real-estate in cockpits. Additionally, a separate interface would also demand a pilot's divided attention, which would be undesirable in time-critical failure conditions. As pilots are already familiar with the PFD and ND, getting accustomed to the augmentations may also require less training time [44].

B. Proposed Symbology

Interface elements were added to both the (Boeing-style) PFD and ND. These elements can be seen in Fig. 6, labeled ① through ⑤.

1. *Attainable Roll Rates*

The attainable roll rates (shown as asymmetric in the figure due to the presence of the failure) are shown by ①. These were visualized as a green-shaded arc on the roll angle indicator, with the extents of the arc indicating the maximum steady-state roll rate presented as a roll angle preview over a look-ahead time of 5 s. This preview time was chosen to provide an appropriate scaling factor given the dynamics of the model introduced in Sec. III. In this example, the arc covers an area of 30 deg to the right, implying that the maximum steady-state roll rate is 6 deg per second. The attainable roll rate limits follow directly from Eq. (1) by substituting the aileron deflection limits, as is done in Eq. (2).

$$p_{\min} = K_p \delta_{a_{\max}} \quad p_{\max} = K_p \delta_{a_{\min}} \tag{2}$$

The main motivation for including this feature was that it provides a means to directly perceive the impaired flight envelope in terms of roll dynamics. In the asymmetric case presented here, it is directly visible that performance will indeed be asymmetric and how the two attainable roll rate limits relate to one another, e.g., a ratio of approximately 6 to 1 in the example provided here. It is clear which turn can be initiated faster, which is useful for evasive maneuvers, and it gives a first impression on whether leveling out of a turn will be relatively fast or relatively slow.

2. *Current Roll Rate*

The arrow inside the shaded bar presented the current roll rate, again presented as an angle preview (see ②). In the figure, the arrow covers almost the entire shaded bar, indicating that the aircraft is currently rolling at almost its maximum rate. The difference between the current roll rate and the maximum rate indicates how much more rolling performance is available, which will be attained over time if the maximum control input is delivered.

3. *Roll Angle Preview*

The shaded triangles (see ③) indicate the reachable set of roll angles within the next 10 s. Again, this time interval was chosen to provide a good resolution given the dynamics of the model. This is done by calculating the largest roll angle that can be obtained by exerting a continuous maximum control input, initiating a rolling motion either to the left (negative) or to the right (positive). The shaded blue triangles remain stationary if the calculated angle would be larger than the roll angle limit represented by the white triangles. This is the case in Fig. 6 for the right roll limit (shown on the left of the PFD). Alternatively, after 10 s a roll angle of 10 deg left can be attained (shown on the right). This symbol was included to provide an alternative representation to the attainable roll rate limits to see whether pilots would have a clear preference for one symbol or the other: the attainable rate limits are stationary on the PFD screen (attached to the attitude scale), whereas the triangles rotate along with the attitude indicator (and the horizon).

4. *Roll-Out Indicator*

The roll-out indicator (ROI) (see ④) presented information on the angle covered in order to bring the current roll rate to zero. In the figure, the current roll angle is zero degrees, but the roll rate is about 3 deg per second. The ROI deviates approximately 12 deg from the current roll angle, implying that (even given the maximum *opposite* stick input) the aircraft would roll out to 12 deg of roll. Making the connection to Fig. 3, we may realize that when the current state of the aircraft is at ③ in that figure, the ROI would be located beyond

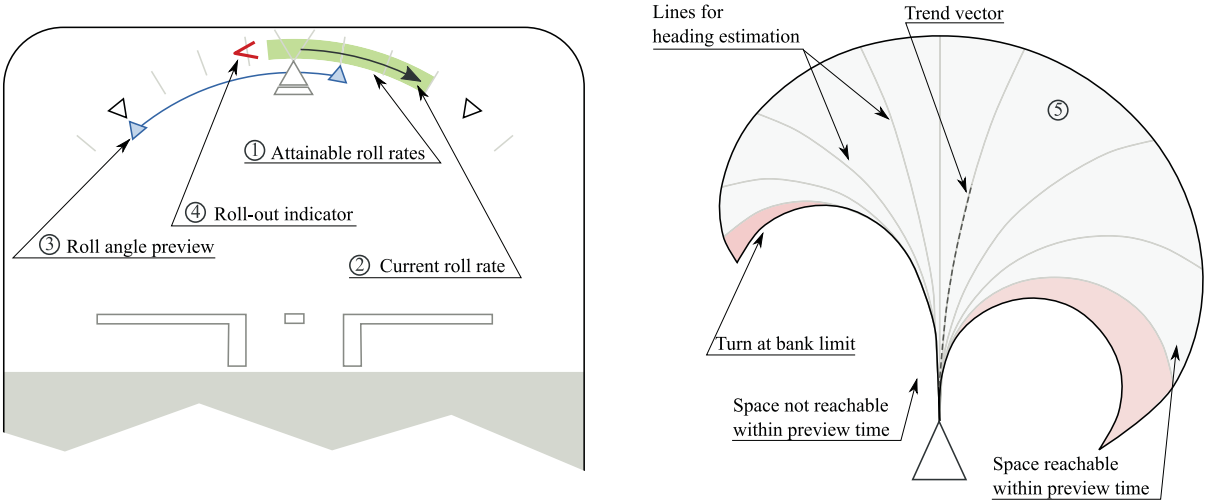

Fig. 6 Interface modifications on the PFD (left) and ND (right).

the right roll angle limit (which is located on the left in a Boeing PFD) as the combination of roll angle and rate is no longer safe. Therefore, the pilot is required to keep the ROI within the indicated roll angle limits. This symbol was included because it was hypothesized that the asymmetric rate performance would make it harder to not exceed the roll limits as was clarified in Fig. 3. In a (potentially autopilot-controlled) flight control system without FEP, this symbol aids the pilot in staying within the maneuvering envelope. In a system with FEP, this symbol would explain why the control system is reducing the roll rate even though the roll limit is not yet attained.

5. Navigation Envelope

The ND was augmented with one additional symbol: the navigation envelope (NE) (see ⑤ in Fig. 6). This symbol provided a 1-minute preview of the reachable navigational space by using the model's dynamics [Eq. (1)] and limitations [Eq. (2)]. The preview time was selected based on a tradeoff between introducing too much clutter (the NE will wrap back onto itself with a large preview time) and expected ND zoom ranges. The ownship symbol and standard position trend vector are included in the figure to clarify where the NE would appear. Inside the shaded area, curves were included in order to facilitate heading estimation: the pilot can estimate the heading corresponding to any position by mentally interpolating in between the curves.

This is illustrated in Fig. 7: if a pilot would want to intercept the vector shown in the figure, the NE explicates the control actions that can be employed to achieve this goal. To define a successful procedure, a pilot would mentally slide the vector down toward the ownship symbol. In this figure, the vector would then first become tangential to curve 1 at the point indicated. This means that a pilot could initiate a turn, i.e., position the trend vector on curve 1, when the vector would touch curve 1 and intercept the vector at point A. The same is true for curve 2 and point B, or curve 3 and point C, respectively. It should be noted that these trajectories would require an increasingly larger maximum roll angle. In fact, to intercept the vector at point C it can be seen that a turn at the maximum roll angle would be required. If no action would be taken and the vector would end up in the position indicated with 4, an attempted intercept will cause the aircraft to fly past the vector. This could also signify the need to pursue a different approach strategy altogether.

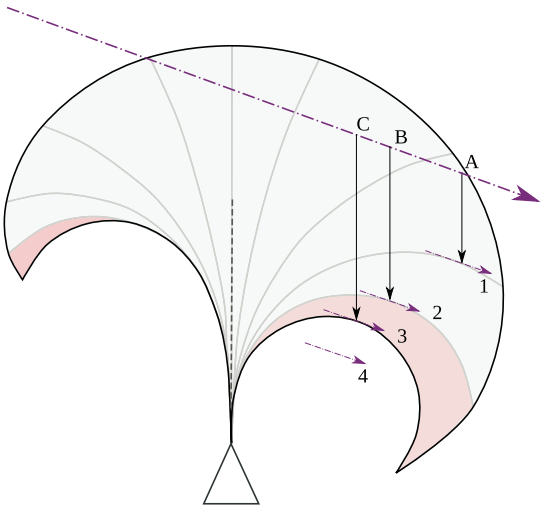

Fig. 7 Estimating different turns that will intercept the given vector.

Note that the curves in the figure do not represent turns at a given roll angle; instead, they represent turns at a given *target* roll angle. That is, using the dynamic model, a control sequence to go from the current roll angle to the target roll angle is calculated using a doublet input with the maximum aileron authority. If the target roll angle is reached within the preview time, the remaining part of the calculated turn is performed at the target roll angle. Figure 8 shows what the NE may look like given different initial roll angles and an asymmetric aileron failure. The shape of the NE hints what turns could be more favorable in terms of recovery opportunities; given that higher roll rates to the left can be reached, turns toward the right are slower, but can also be recovered from easier.

C. Expected Interface Usage

The goal of the novel interface symbology is to allow pilots to understand and act upon cues that indicate off-nominal performance in an effective manner. Rasmussen's taxonomy of the levels of cognitive control [45] may shed a light on how the interface should be used and what level of cognitive control the different cues support.

Rasmussen stated that a distinction can be made between three levels of cognitive control: skill-based behavior (SBB), rule-based behavior (RBB), and knowledge-based behavior (KBB). Operators employ SBB for simple, familiar tasks that require little to no mental effort, e.g., the tracking task of keeping an aircraft's wings level while flying in mild turbulence. RBB is employed when familiar solutions exist to familiar problems and is triggered by "if–then" rules, e.g., completing check lists, flying a standard rate turn, etc. KBB is triggered in unfamiliar situations where neither SBB nor RBB can be used to solve a problem. It generally involves laborious and error-prone problem-solving activities, such as fault detection and diagnosis and estimating the operational impact of the failure. For instance, when a standard rate turn cannot be properly executed because the desired rate cannot be attained, a pilot will need to start reasoning on what may have happened to the aircraft that is causing this observation: the turn indicator could be broken, there may be something wrong with the flight controls, and/or the aircraft itself is damaged. Additionally, KBB will be involved in assessing the operational impact of the failure, such as estimating what navigational states can and cannot be reached.

Which levels of behavior are used to perform a certain task ultimately depends on a number of factors, such as the operator's training and familiarity with the given situation, as well as the available support for the task, e.g., the interface. In current operations when experiencing failures, flying the aircraft, performing system diagnosis, and navigating are divided between the captain and first officer. In this study, however, it is expected that a single pilot performs both the flying and near-term navigation tasks, freeing up cognitive resources of the pilot not-flying to, for example, diagnose the cause of the failure.

It is expected that both the additional PFD and ND symbology can lower the required cognitive demand by allowing for "shortcuts" in the decision-making process, which may enable pilots to employ SBB and RBB instead of exercising KBB that would have been needed without the additional cues. Here, the NE is hypothesized to play an important role in lowering KBB to RBB or even

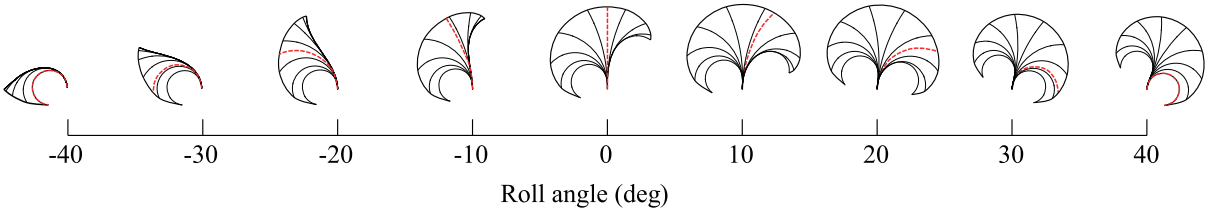

Fig. 8 Different NE shapes indicating how the reachable states are affected by the roll angle (asymmetric roll performance). The red dashed lines indicate the trajectory that will be flown given the current state.

SBB when flying to safe navigation targets in asymmetric flight envelopes.

First, the NE provides an intuitive glance on the asymmetry of the lateral flight envelope. Observing the asymmetry can be qualified as SBB and deducing what turns are more favorable than others in certain situations can be regarded as RBB (e.g., if a sharp right turn would be initiated, then rolling back to wings level will take more time). The additional cues on the PFD mainly support the pilot in controlling the immediate aircraft states within the safe boundaries of the envelope. Here, the attainable roll rate bar and the roll angle preview support RBB as they visualize boundaries for aiming (i.e., SBB) the current roll rate and roll angle.

Note that some KBB will still be required when using the NE for planning the initial trajectory, as full trajectories leading to a far-away navigation target (e.g., the runway localizer) are not visualized. However, with progressively smaller distances to the navigation target, the NE will provide increasingly more cues as to which control action to perform, up to the point where there is only the possibility of immediately giving a control input to intercept the navigation target. At these last stages, no KBB is required for factoring the failure information into the tactical planning and mentally integrating future states.

V. Experiment

A pilot-in-the-loop experiment was performed to assess the pilot's ability to understand and work with the proposed PFD and ND enhancements as well as the performance of pilot actions in flight conditions featuring asymmetric flight envelopes. Note that the purpose of the experiment was not to evaluate the display cues under realistic failure conditions featuring high levels of stress. Such situations would be difficult to recreate in a simulator study anyway [46]. As such, the psychological relevance of, for example, workload measurements should only be considered relative to the current experimental settings.

A. Apparatus

The experiment was conducted in TU Delft's SIMONA research flight simulator. For this experiment, the six-degree-of-freedom motion base was not used and a single pilot operated the simulator. Participants controlled the simulation exclusively by means of a right-handed (first officer seat) control-loaded side stick. In front of the participant, two screens showed the ND (left) and PFD (right) displays (see Fig. 9). An outside visual consisting of flat terrain was added for realism and providing optical flow cues, but no additional cues could be obtained from this visual.

B. Participants

Nine pilots with various degrees of professional experience participated. Their qualifications are shown in Table 1. All pilots had experience with glass-cockpit displays.

C. Procedure

A high-level overview of the experiment procedure is shown in Table 2. The elements of the procedure are elaborated upon below.

1. *Briefing*

Initially, pilots were briefed on the tasks they were about to perform and the interfaces that would be used during the experiment. The constraints of the experiment (e.g., fixed airspeed, side stick control only) were also explained at this stage.

2. *Training*

After the briefing, a training phase commenced in which the participants were given the time to familiarize themselves with the different interface combinations and the dynamic model of the aircraft (as outlined in Sec. III). At this point, no failure was present in the model and the model was linearized with the outboard ailerons as active control surface ($K_p = 0.3984\ \text{s}^{-1}, \tau_p = 1.130\ \text{s}$).

During the training phase, participants were encouraged to track a reference trajectory consisting of multiple straight segments with increasingly sharper angles between them. This prompted them to exert a significant amount of stick deflections such that they could clearly observe the behavior of the different interface symbols. Additional clarifications relating to the new symbology were still being provided at this stage.

During all phases of flight, participants were instructed not to exceed the visualized roll angle limits. In this event, a stall warning

Table 1 Pilot qualifications

License type	No. of pilots
Private Pilot License (PPL)	1[a]
Commercial Pilot License (CPL)	1
Airline Transport Pilot License (ATPL)	7[b]

[a]This participant had no experience with the ND position trend vector.
[b]One participant was still in training with four months remaining; three participants had experience as research test pilot.

Table 2 Chronological overview of participant activities

Receive briefing
Fly training runs
Identify failure
Select runway
Intercept localizer
Fill in workload assessment
Fill in questionnaire

Fig. 9 The interior and exterior of the SIMONA research flight simulator. The interior reflects the setup that was used to conduct the experiment.

would be shown, the simulation would be halted, and the event would be treated as a loss-of-control event.

3. Measurement Runs

After the training phase, six trials were performed under different experimental conditions. The trials were split up in distinct parts, drawing inspiration from the GARTEUR FM-AG(16) benchmark [39]. These parts are failure identification, runway selection, localizer intercept, and workload assessment.

First, the pilot was given the opportunity to identify the failure in flight, while not pursuing any other goals at the same time. The motivation for this was that in a real-life performance-altering event, the first thing a pilot would do would be to get an impression of the altered flight envelope. Note that diagnosing the cause of failure (i.e., the reason why the flight envelope was altered) was not part of the task.

Second, the simulation was halted and the aircraft was repositioned to a predetermined position in the vicinity of an airport with two runways. The pilot was asked to indicate which runway would be preferred for performing a landing, given the current aircraft state, the orientation of the runways, and the nature of the failure.

Third, the pilot was asked to perform a "raw data" (i.e., without flight director) localizer intercept. The pilot was instructed to perform the intercept in the best possible way while minimizing the flying time and the localizer error along a defined segment. It was indicated that as the cause of the failure was not known, the failure could potentially get worse. In practice, this meant the pilots were encouraged to only fly an indirect approach when they thought they would not be able to intercept the localizer sufficiently well by performing a direct approach. A direct approach as defined here consists of one turn toward the localizer, bringing the aircraft on an intercepting heading, and one turn to align with the localizer. Participants were instructed to complete the localizer intercept before the glide slope would be captured and to keep minimizing the error until that point. The simulation was halted when the glide slope intercept point was passed. Note that these two waypoints were both portrayed on the ND by star-shaped icons, as shown in Fig. 12.

As the fourth and last step of each measurement run, the participant was asked to complete a workload self-assessment by means of filling in a NASA Task Load Index (TLX) form [47].

D. Independent Variables

Two independent within-participant variables were defined for this repeated-measures experiment: the failure condition ("FAIL"), which had two levels, and the EFIS display configuration ("DISP"), which had three levels. Hence, in total there were six different experimental conditions in this experiment. A Latin square distribution was used to mitigate order effects.

1. Failure Condition

The two levels of FAIL were named "moderate" and "severe." Their differences are in the aileron offsets required to maintain a constant roll angle, as illustrated in Fig. 10. The safe maneuvering constraints as introduced in Fig. 3 were calculated for both failure conditions; these are shown in Fig. 11. In the direction of impaired roll authority—the most challenging aspect of the asymmetric envelope—the maximum rates were 1.41°/s (left) and 0.70°/s (right) for the moderate and severe failure conditions, respectively.

2. Display Configuration

The three levels of DISP were named "Base," "PFD+," and "ND+." Base, or baseline configuration, consisted of a standard PFD and ND with the addition of roll angle limits on the PFD in the form of "barber poles," representing the stall angle that was not to be exceeded. Such

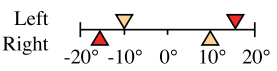

Fig. 10 Inboard aileron deflection for straight and steady flight for moderate (inner triangles) and severe (outer triangles) case.

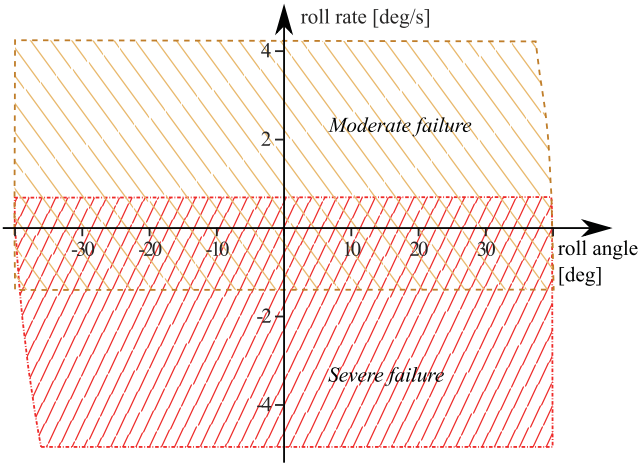

Fig. 11 Lateral maneuvering envelope for the two failure levels.

cues represent the state-of-the-art in communicating control limits to pilots.

The PFD+ level added only the PFD symbology that was introduced in Sec. IV, and the ND+ level included both the PFD cues and the NE on the ND. The different configurations are summarized in Table 3. A screen capture of the actual PFD and ND with the additional symbology as it was used in the experiment is shown in Fig. 12. The rationale for including both the PFD+ and ND+ conditions was to investigate how much the additional cues on the PFD alone would already enable a pilot to make better decisions regarding reaching near-term navigation targets.

E. Control Variables

The parameters that were used to linearize the GARTEUR RECOVER Boeing 747 model are listed in Table 4. Additionally, the resulting control parameters that were used in the experiment to simulate the aircraft behavior as referenced in Eq. (1) are shown in Table 5. Using this model to drive the aircraft's behavior and the visualizations ensured a perfect match between the aircraft response and the visualized information on the PFD and ND. Note that wind and turbulence were not simulated. Furthermore, altitude was constrained; that is, the "combined system" that was introduced in Sec. III only allowed the pilot to perform coordinated turns at constant altitude. For the sake of experimental control, the input and output degrees of freedom were constrained in order to facilitate comparing the decisions and control actions that different pilots made.

F. Scenario

The scenario is shown in Fig. 13, where the aircraft symbol represents the initial position and heading. The diverging semishaded lines represent the two localizer centerlines leading to the two runways indicated at the top of the figure. Each localizer is marked with the two waypoints that were introduced in Sec. V.B; these waypoints are also shown.

For the runway selection task, the participant was asked to indicate a preference for executing the localizer intercept on runway ① or ②. It was indicated that the runway properties were similar in both dimensions and the available support, and that there was an absence of wind. This made sure that environmental conditions (e.g., runway orientation and wind direction) would not bias decisions toward a specific solution.

For the localizer intercept task, runway ① was always "randomly selected" for performing the localizer intercept task in order to gather sufficient data points. In 50% of the trials, both the failure and the runway layout were mirrored along the north axis in order to prevent scenario and failure mode recognition. This is also why the runway layout was designed to be slightly asymmetric with respect to the initial position and heading.

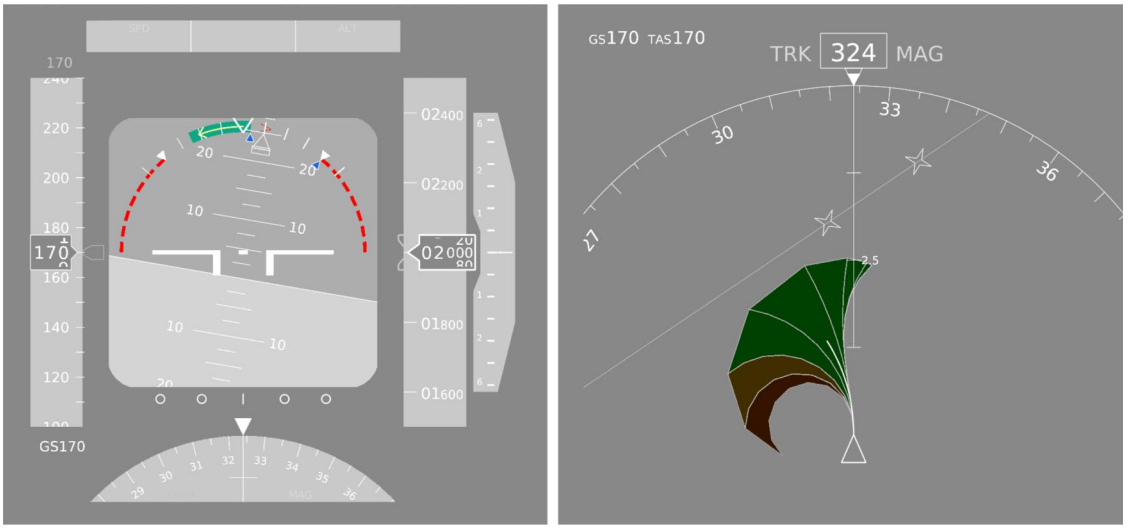

Fig. 12 Additional display elements in the PFD and ND for the ND+ display condition.

Table 3 Interface elements available for the different display configurations

	Display condition		
Interface element	Base	PFD+	ND+
Roll limits	✓	✓	✓
Roll rate limits		✓	✓
Current roll rate		✓	✓
Roll preview		✓	✓
Roll-out indicator		✓	✓
Standard trend vector	✓	✓	✓
Navigation envelope			✓

Table 4 Linearization parameters of GARTEUR RECOVER Boeing 747 model

Parameter	Value
Mass	200,000 kg
Failure mode	0 (none)
Gear	0 (down)
Flaps	20°
Altitude	2,000 ft
Speed	170 kts
Flight path angle	0°

Table 5 Configuration of experiment model control variables

Parameter	Value
K_p	0.141 s^{-1}
τ_p	1.130 s
Altitude	2000 ft
Speed	170 kts
Flight path angle	0°
Roll angle limits	±40°
Inboard aileron deflection limits	±20°

G. Dependent Measures

The interfaces were evaluated using objective data originating from the logged data, as well as subjective data originating from NASA TLX workload assessments [48] and pilot opinions by means of a questionnaire.

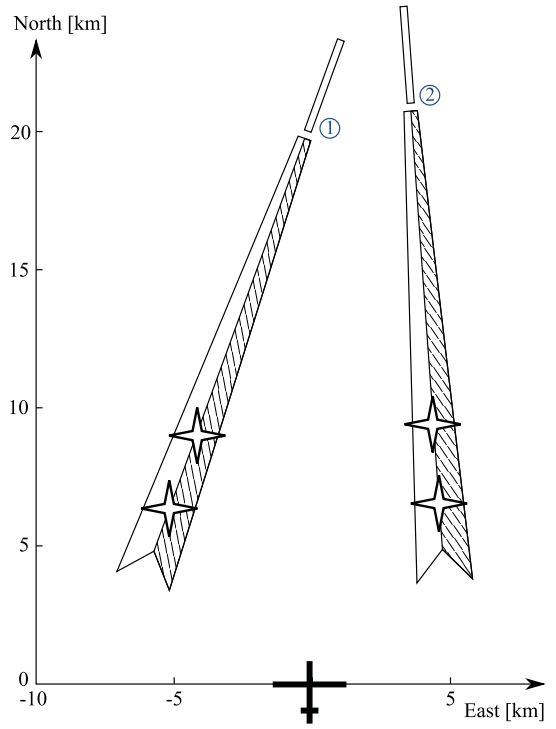

Fig. 13 Scenario used for runway selection task and localizer intercept. The intercept was always performed on runway 1.

1. *Decision Quality*

Two measures were used for assessing the quality of the decision: the preferred runway and the employed navigation strategy. Measuring the decision quality in this manner is inspired by an earlier study [17]. Comparing the runway selection between the different conditions could indicate whether or not the display condition would have an impact on the preferred runway and the consistency in the preferred runway. The employed navigation strategy was characterized as a "direct intercept" or an "indirect intercept." A direct intercept was defined as a trajectory consisting of two consecutive turns with a total flight time of less than 150 s. This strategy was deemed the preferred decision, as it was possible to perform a localizer intercept in this manner within the constraints of the experiment, and the pilots were instructed to employ this approach if they thought it would be feasible. The justification for this is that a direct localizer intercept would, in this experiment, result in a significantly lower total flight time, which would minimize the

probability of damage getting worse over time. For instance, the original investigation of Flight 1862 mentioned that the pilots may have believed that the starboard wing was on fire at the time. An indirect intercept was defined as any strategy resulting in a total flight time of more than 150 s.

2. *Safety*

Safety was evaluated using the following metrics:

1) *Completion time:* It is measured by starting time at the beginning of the intercept run until the aircraft crossed a line perpendicular to the localizer at the second waypoint. Lower completion times are regarded more favorable in terms of safety as failures may evolve over time.

2) *Maximum roll angle during the intercept phase:* Larger roll angles are closer to the boundaries of the safe flight envelope, leading to aerodynamic stall or insufficient thrust to maintain the current altitude.

3) *Maximum roll rate during the intercept phase:* Higher roll rates cause larger structural loads and require more control actions.

3. *Performance*

Performance was evaluated by the following metrics, representing a subset of how instrument landing system (ILS) tracking performance is commonly measured [49]: 1) lateral localizer error at glide slope, 2) heading/track error at glide slope, and 3) yaw rate at glide slope.

For these metrics, lower values generally indicate that the pilot was more successful in tracking the localizer center line.

4. *Workload*

Following every localizer intercept phase, participants were asked to assess their workload by completing a NASA TLX form [48]. A filled in TLX culminates in an overall workload score based on ranking and grading the following categories: mental load, physical load, temporal demand, performance, effort, and frustration.

H. Final Questionnaire

At the end of the experiment, participants were asked to complete a final questionnaire. The questions related to the pilot's self-assessed ability to infer meaning from the different interface elements. Furthermore, participants were asked to comment on their concerns regarding interface usability (e.g., clutter) and training effects that may have been present.

I. Hypotheses

It was hypothesized that pilot decision quality, navigation strategy, and employed control actions in the PFD+ and ND+ conditions would be significantly better than in the baseline conditions. It was further expected that these results would be more pronounced in the severe failure conditions. Significant differences between PFD+ and ND+ conditions were expected to be present in the severe failure case, where the ND+ condition was expected to yield better results. This is because the NE would be most informative in revealing the level of asymmetry of the damaged flight envelope, guiding pilots toward a specific solution and encouraging them to undertake timely control actions. In the moderate failure condition, the NE appears more symmetric, making its added value less pronounced.

VI. Results

In general, all pilots were able to successfully complete the experiment and none of them exceeded the roll angle limits during any of the measurement runs. Given the limited sample size ($N = 9$), nonparametric statistical tests were used to detect significant differences between the display conditions in conjunction with the failure severity. All significant results are marked in boldface.

A. Decision Quality

1. *Runway Selection*

Figure 14 shows whether the pilots preferred to perform a localizer intercept where the first turn would be in the direction of the slower or the faster roll rate. It can be seen that across all conditions, the pilots generally seemed to prefer to first turn with a slower roll rate, the impaired performance turn. However, each pilot occasionally decided on flying the faster turn first; a clear rationale behind this was not given. Cochran's Q test was used to detect significant differences across the experiment condition conditions with respect to runway preference. The results were not found to be significant [$Q(5, 9) = 5.56$, $p = 0.352$].

Pilots were asked to briefly describe how they would fly the approach, which confirmed the statistical results: in most cases, the pilot indicated to "get the hard part out of the way first"; i.e., initiate the first turn into the direction of the impaired roll performance. In 51 out of 54 runs, pilots estimated that they could fly a direct localizer intercept by performing two consecutive turns, while sometimes anticipating the potential need for changing this strategy to a more indirect approach, for instance, by flying a so-called "teardrop" maneuver.

Based on the pilots' initial analyses, it appears that only one pilot seemed to realize that "getting the hard part out of the way first" would raise another issue: upon localizer capture, leveling out after the final turn would again require rolling the aircraft at an impaired rate, potentially causing the pilot to miss the localizer intercept.

2. *Employed Navigation Strategy*

The flown trajectories can be seen in Fig. 15. All trajectories started at (0, 0). The trials that were presented to the pilot in a mirror image have been transformed back in this figure in order to compare the results. The marker closest to the origin represents the waypoint where the intercept should have been completed if possible, and the other marker represents the intersection with the glideslope.

Considering the moderate failure case, in both the Baseline and PFD+ configurations one pilot decided not to attempt a direct approach, though it was not the same pilot each time. With the ND+ configuration, every pilot decided that the intercept could be flown directly. Furthermore, it seems that with the ND+ configuration the pilots opted for a larger intercept angle on average, enabling them to intercept in closer proximity to the first waypoint as requested.

Looking at the severe failure case, it is observed that both with the Baseline and PFD+ conditions, many pilots opted for an indirect approach, employing various strategies to intercept the localizer.

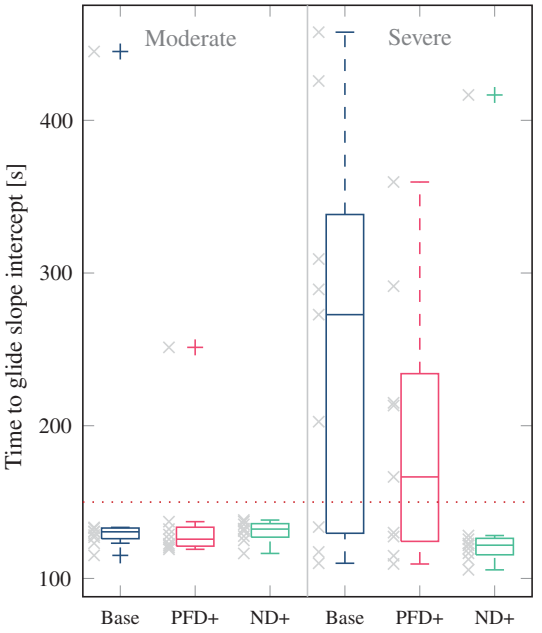

Fig. 14 Preferred runway selected for each condition.

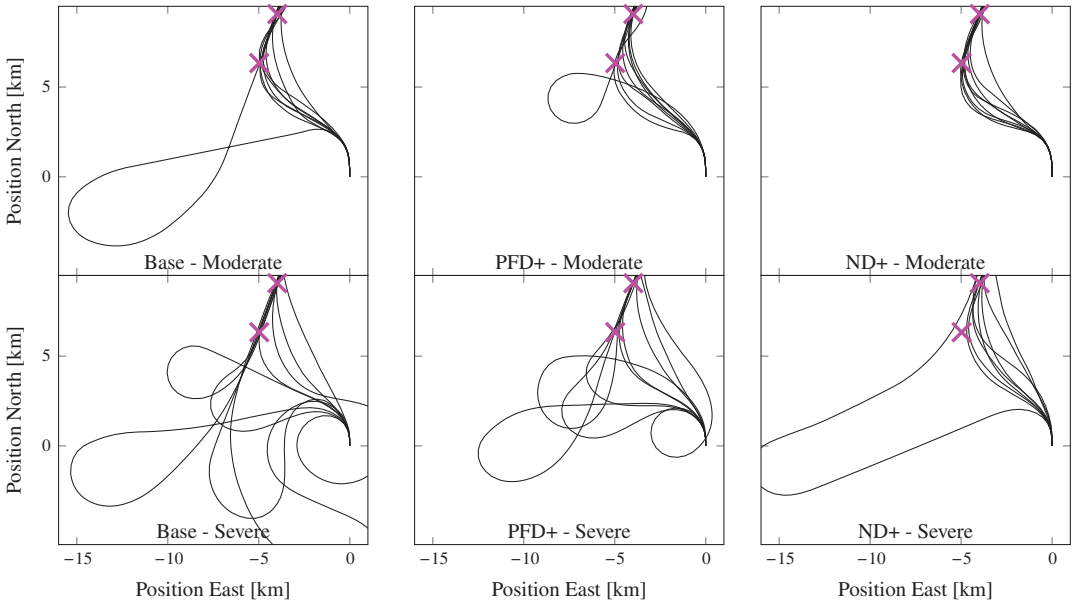

Fig. 15 Flown trajectories for the six experimental conditions.

With the ND+ configuration, pilot behavior was more consistent: only one pilot opted to fly away from the runway first. Cochran's Q test was used for significance with respect to the localizer intercept strategy. The test pointed out that the display level had a significant effect on the chosen approach strategy [$Q(2, 9) = 8.4$, $p = 0.015$]. However, a post-hoc analysis with Bonferroni-corrected pairwise comparisons using McNemar's test did not yield a significant difference.

B. Safety

The employed strategy can also be observed in the time taken to intercept the glide slope, as presented in Fig. 16. The dotted line at 150 s represents the threshold above which an indirect approach was flown. Here, it can also be seen that, as a result of the different indirect approaches in the severe failure case, both the median and variance in the data were higher when the ND+ display condition was not used. Friedman's ANOVA revealed that DISP had no significant effect in the moderate failure case [$\chi^2(2) = 2.667$, $p = 0.264$]. A significant effect was found, however, for the severe failure case [$\chi^2(2) = 8.222$, $p = 0.016$], as was hypothesized. Here, a post-hoc analysis using the Wilcoxon's signed-rank test showed that the ranks for the Baseline and PFD+ levels were significantly different ($Z = -2.547$, $p = 0.011$) at the Bonferroni-corrected significance level of $p = 0.0167$.

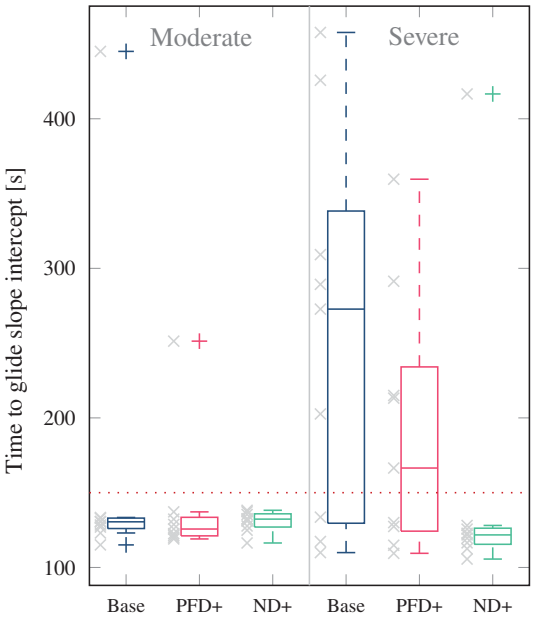

Fig. 16 Completion time.

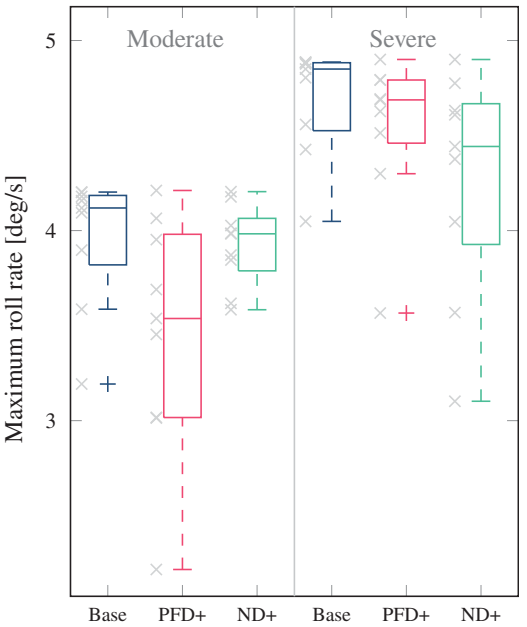

Fig. 17 Maximum roll rate.

Figure 17 shows the maximum roll rates that were registered during the intercept task. It can be seen that with the severe failure case, higher roll rates were commanded. Across the different display levels, the results seem more consistent. In fact, no significant effect was found for either the moderate failure level [$\chi^2(2) = 2.889$, $p = 0.236$] or the severe failure level [$\chi^2(2) = 2.889$, $p = 0.236$].[‡‡]

Finally, looking at Fig. 18, the maximum roll angle that was obtained during the intercept runs was fairly consistent across the different display levels, especially for the moderate failure case. Friedman's test confirmed the lack of a significant difference [$\chi^2(2) = 2.889$, $p = 0.236$]. In the severe failure case, Fig. 18 shows a lower median maximum roll angle for the ND+ level. Although

[‡‡]Friedman's ANOVA indeed yielded the same results for both tests.

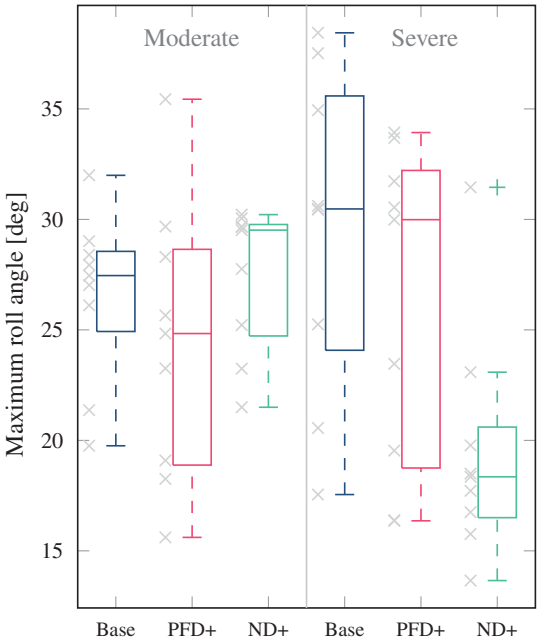

Fig. 18 Maximum roll angle.

there was a significant effect on the maximum roll angle in the severe failure condition [$\chi^2(2) = 6.889$, $p = 0.032$], one should be careful not to interpret this result as the NE causing more conservative steering behavior. Instead, in this scenario the NE made obvious that a left turn at a high roll angle would most likely result in failing to intercept the localizer, suggesting pilots to execute the turn at a lower roll angle. Wilcoxon's signed-rank tests did not show a significant effect for any of the pairwise comparisons.

C. Performance

The box plots in Figs. 19–21 aim to visualize how well the pilots managed to reach the glide slope in a stable and accurate manner. The three figures tell a similar story as, in general, a large lateral error would typically drive a pilot to command a heading change toward the glide slope, which would also temporarily result in a nonzero yaw rate.

It can be observed that, in general, the Baseline display level yielded the best performance in terms of the metrics presented here. This can be explained when one takes another look at the trajectories

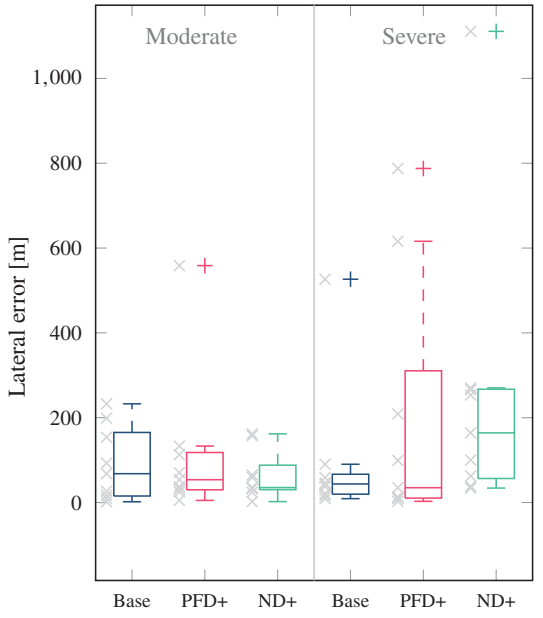

Fig. 19 Lateral error at glide slope.

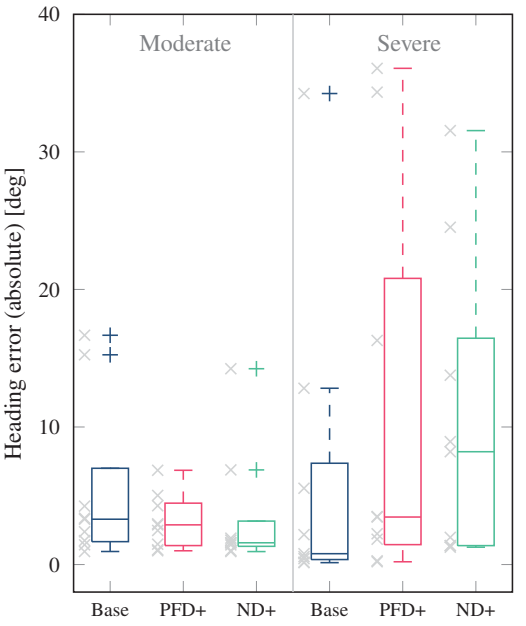

Fig. 20 Heading error at glide slope.

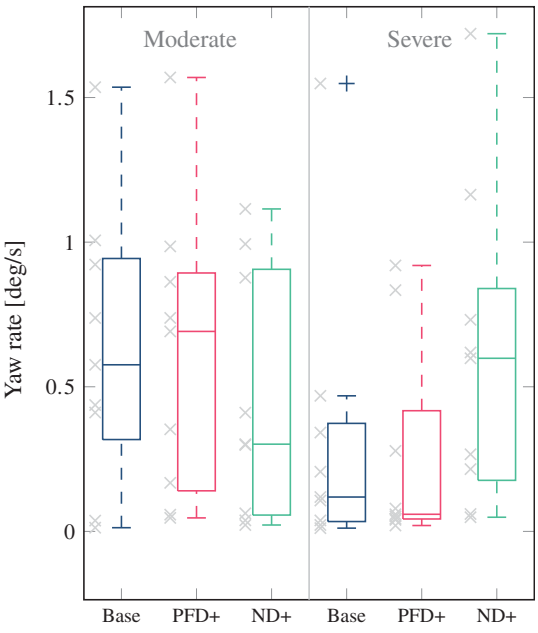

Fig. 21 Yaw rate at glide slope.

in Fig. 15: as the employed navigation strategy was indirect much more often in the Baseline display level, the pilots were able to intercept the localizer further away than when a direct approach was flown. The result of this is that they had more time on their hands to correct for any localizer error. This also explains why the yaw rate at the glide slope was often smaller in the severe failure case: again, in this case an indirect approach was selected more often.

For the lateral error, no significant effects were found for either the moderate [$\chi^2(2) = 1.556$, $p = 0.459$] or the severe [$\chi^2(2) = 0.222$, $p = 0.895$] failure cases. The same holds for the heading error in both conditions [$\chi^2(2) = 1.556$, $p = 0.459$ and $\chi^2(2) = 0.667$, $p = 0.717$], and for the yaw rate [$\chi^2(2) = 2.000$, $p = 0.368$ and $\chi^2(2) = 0.889$, $p = 0.641$].

D. Workload

Figure 22 shows the overall workload that was experienced by the participants according to the TLX results. Additionally, the weights

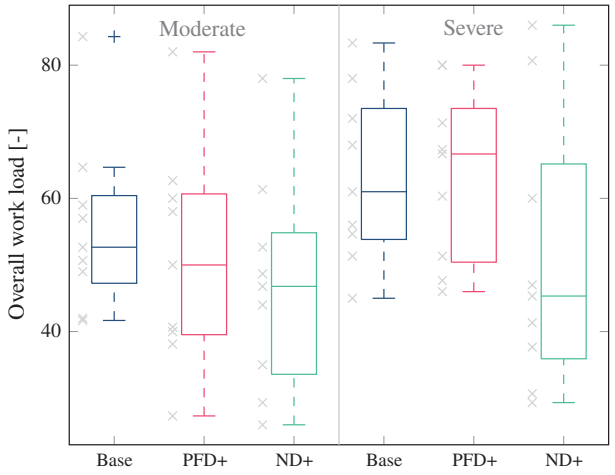

Fig. 22 NASA TLX overall workload as reported by the pilots.

of the different TLX categories are shown for the different conditions in Fig. 23 to illustrate what factor contributed most to the experienced workload. The weights seem consistent across the experimental conditions, which implies that the nature of the task was not perceived fundamentally different as a result of the display and failure conditions. It can be observed that for both conditions with the ND+ level, both mental load and frustration were ranked slightly lower than in the other conditions, which may indicate that pilots felt less discouraged and more secure when using the NE.

When the navigation options were harder to estimate (in the severe failure case), a relative workload reduction of the ND+ condition compared with Base and PFD+ is observed. This may be explained by recognizing that the NE allowed pilots to make decisions using RBB instead of KBB when planning their actions, requiring less cognitive effort.

In Fig. 22, the results show an increase in the median workload with the severe failure condition. This should be expected, as there was a lower margin of error for completing the task in an optimal way. The most obvious outcome is that the median reported workload was lower especially in the ND+ display configuration. This indicates that displaying integrated information may indeed lower the workload for the task.

Furthermore, in the severe failure case, the PFD+ condition seemed to increase the reported workload. Surprisingly, the Baseline and PFD+ conditions resulted in similarly experienced workload patterns. This could imply that the introduced symbology was creating more clutter than providing useful information. An analysis using

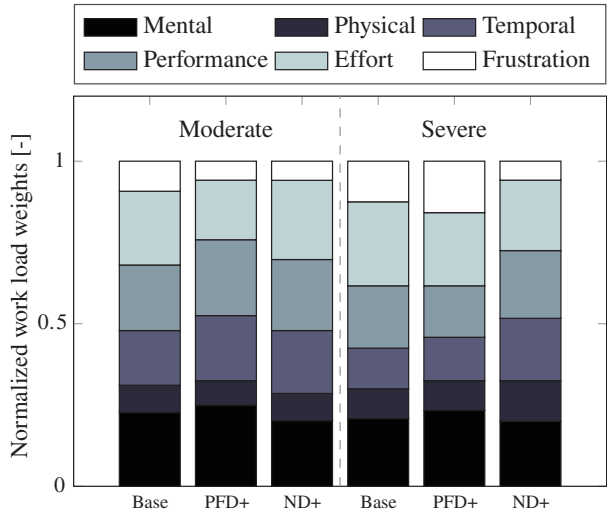

Fig. 23 TLX workload weights.

Friedman's ANOVA with respect to the TLX overall workload did not yield a significant result [$\chi^2(5) = 10.00$, $p = 0.075$].

E. Questionnaire Results

At the end of the experiment, all participants were asked to complete a questionnaire in order to gather additional information on how the new interfaces were received. As one questionnaire response failed to be recorded due to a technical error, only data of eight pilots are presented here.

When asked which feature was the most useful for controlling the aircraft (see Fig. 24), pilots reported that for the PFD, the green-shaded arc indicating the attainable roll rates was found more useful than the other PFD additions. All pilots indicated that they relied on this feature to some extent. The other PFD features were all found to be of very limited use, perhaps introducing more clutter than providing useful information. The NE on the ND, on the other hand, was deemed the most useful feature overall.

When asked whether the different interfaces improved or decreased their performance, results were similar to the responses shown in Fig. 24, where pilots unanimously reported that mainly the modified ND improved their performance. This confirms that the pilots indeed favored a direct localizer intercept with a slightly larger error in terms of deviation and heading over taking additional time to intercept the localizer at a later point in time.

When asked whether the PFD or ND modifications aided the pilots in understanding the consequences of their control actions, again a similar pattern was visible: all pilots found that the PFD modifications helped to some extent, but again the NE was the main driver of this aspect, with most pilots finding this very useful. In terms of the runway selection task, the majority of the pilots thought the PFD modifications were not helpful in the runway selection task, whereas the NE did help in this respect.

Finally, most participants also felt that these or similar interface elements would help them in performing better during a performance-altering emergency, with one pilot remarking that this was only true with respect to the ND. The majority of the participants did not feel that the additions would introduce unacceptable clutter on the displays.

VII. Discussion

The goal of this study was to explore visual ways to communicate the consequences of asymmetric flight envelopes on the immediately controllable aircraft states as well as the reachable navigation states. The results from an exploratory pilot-in-the-loop experiment revealed that portraying additional information on the ND, in the form of an NE, was most useful in helping pilots to assess the severity of the impaired envelope, to reduce their workload and undertake appropriate control actions. This is rather surprising, given that the majority of studies in communicating flight envelope limits mainly focus on portraying additional information on the PFD [2,19]. At this stage, however, it is also important to acknowledge and discuss the limitations of this study so as to provide recommendations for follow-up studies.

Although the initial results look promising, caution is necessary when interpreting the results from this experiment. One of the limitations was the relatively small number of participants, which made it unfeasible to make bold statements about the population mean and variance of the measured signals. The lack of significance in some dependent measures, due to relatively large variability in the data, can be explained in two different ways.

On the one hand, pilots were free to employ any control and navigation strategy they deemed appropriate as the display(s) only show the boundaries of actions rather than presenting a single optimized action. Giving them such freedom generally comes at the cost of more predictable outcomes, as was found in other studies [43]. On the other hand, indications of insufficient training time were observed, as several pilots felt that they became more proficient in using the new interface cues during the experiment. An increased training time would likely have resulted in a better familiarity with the interface elements and perhaps less pilot variability. As reported in

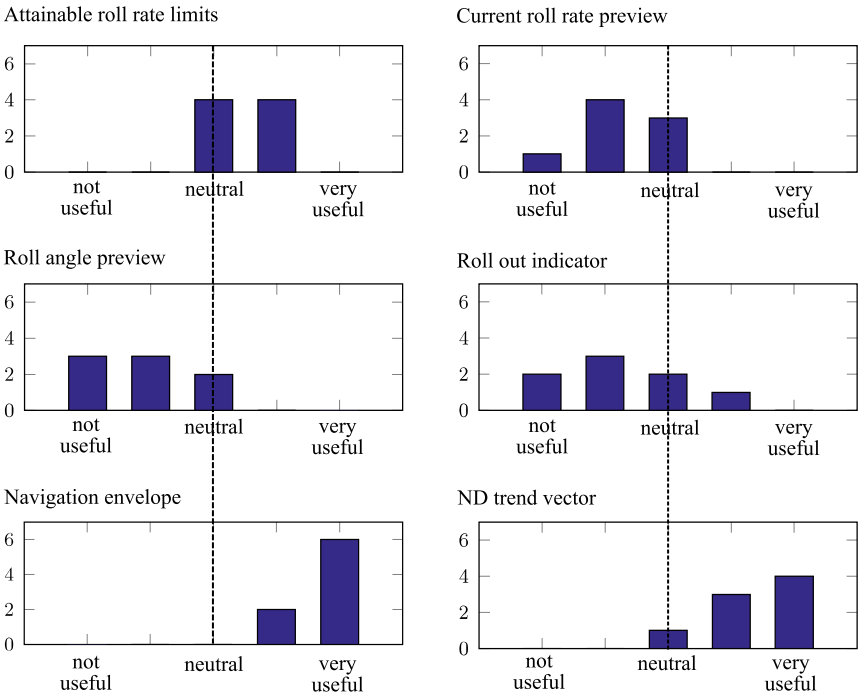

Fig. 24 Pilot usefulness ratings of the designed interface cues.

other studies, providing sufficient training time to get comfortable with constraint-based interfaces seems to be a common problem [43,50]. Note that "pushing the envelope" will also depend on other factors encountered in real-life situations, such as substantial wind gusts and shear, which will likely make pilots more cautious due to uncertainties.

Another limitation is related to the purely two-dimensional navigation task and the simplified aircraft model. This was required for the sake of experimental control, but the question that remains is whether or not the presented interface elements can be modified to 1) provide sufficient and coherent information in the event of a more complex failure mode, such as with a high-fidelity model of Flight 1862, and 2) encode three-dimensional position and flight path information. As flight envelopes (especially off-nominal ones) are high dimensional, nonlinear, and time varying, it is a legitimate concern whether or not constraint-based interface symbology can capture the complete set of all possible control actions in such a way that it is still possible to infer useful information from its visual representation. Perhaps a combination of presenting an optimal 3D emergency trajectory alongside the NE is the best of both worlds; pilots can use the envelope to visually assess the validity of the generated emergency trajectory and either consequently follow its guidance signals to track the reference trajectory or decide to deviate from it.

The simulation realism can also be identified as a limitation in this experiment: the control task was rather simple and the simulator's motion base was not in use. It is obvious that a real-life emergency situation would be much more stressful. Hollnagel states that in unfamiliar situations with significant time pressure and stress levels, operators tend to use a "scrambled" mode of control, where rational decision making does not necessarily occur [51]. If the experiment would have induced higher stress levels through improved realism, increased time pressure, or other means, the results may have shown less effective use of the interface symbology. However, this would be difficult to test as real-life stress levels are difficult to recreate in simulators [46].

Finally, the ND+ condition occasionally resulted in slightly more aggressive maneuvers than with the other display conditions, suggesting that the NE encouraged pilots to "push the envelope." This behavior seems consistent with observations in ecological interface design research, where operators often decide to operate closer to the visualized limits of the system when these are present [17,19,52]. It can be argued that operators presented with constraint-based interfaces have the tendency to push the envelope, implying that such interfaces invite operators to use less safe behavior. However, this would be an oversimplification. Comans et al. have shown that operators are also compelled to honor intentional constraints (as opposed to physical constraints) when these are visualized, indicating that control behavior can still be influenced by introducing additional soft constraints [52].

In case of the NE present in this study, one could imagine applying a different shading to indicate the roll angles that yield a certain load factor. Additionally, one could argue that knowingly operating close to a system's limit is perhaps safer than operating at a safer region in the state space without knowing where the system's limits are: with the NE present, the pilots were more certain that they could execute certain maneuvers, whereas without this information present, they were less certain about the outcomes of their control actions. A good interface should improve operator confidence, without making the operator *overconfident*.

However, operating close to system boundaries becomes an important issue when the system boundaries are less clear: the model that was used in this experiment was time-invariant and deterministic, and its parameters were assumed to be fully known by the system. In reality, this is a luxury we (currently) do not have: safe maneuvering envelopes of aircraft have a high dimensionality, and the space of all possible control inputs affects the energy balance of aircraft, making it hard to visually or mentally chain together series of viable control actions. Furthermore, external factors such as wind direction and wind strength are hard to predict locally in an accurate manner, introducing additional uncertainties. Here, an interface designer will need to determine whether these uncertainties should be visualized in some form or shape (again with the risk of encouraging operators to push the envelope), or to tighten the constraints (i.e., shrinking the available solution space) to account for uncertainties. However, tightening the constraints is also not risk-free as it might eliminate viable (and possibly, for the emergency situation at hand, even optimal) solutions.

VIII. Conclusions

This study set out to explore visual ways to communicate the consequences of asymmetric flight envelopes on the immediately controllable aircraft states as well as the reachable navigation states. Results from a human-in-the-loop evaluation show that portraying

additional information on the ND, in the form of an NE, was most useful in helping pilots to better assess the severity of the impaired envelope, to reduce their workload, and to plan and execute appropriate control actions. Although the results bear limited statistical significance due to the limited sample size combined with relatively large variability in pilot behavior, the trends in the data do indicate that communicating consequences of impaired flight envelopes on the ND is worthwhile to explore further. For future research, it is recommended that emphasis is placed on simulation realism, increasing the sample size, modeling and visually representing parameter uncertainty, and possibly combining the visualizations with an emergency trajectory generator.

References

[1] "ALAR Briefing Note 1.3—Golden Rules," Flight Safety Foundation, 2000, https://flightsafety.org/files/alar_bn1-3-goldrules.pdf.

[2] 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. https://doi.org/10.2514/1.G001740

[3] Nabi, H. N., Lombaerts, T., Zhang, Y., van Kampen, E., Chu, Q. P., and de Visser, C. C., "Effects of Structural Failure on the Safe Flight Envelope of Aircraft," *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 6, 2018, pp. 1257–1275. https://doi.org/10.2514/1.G003184

[4] Zhang, Y., de Visser, C. C., and Chu, Q. P., "Database Building and Interpolation for an Online Safe Flight Envelope Prediction System," *Journal of Guidance, Control, and Dynamics*, Vol. 42, No. 5, 2019, pp. 1166–1174. https://doi.org/10.2514/1.G003834

[5] Edwards, C., Lombaerts, T., and Smaili, H., *Fault Tolerant Flight Control: A Benchmark Challenge*, Springer, Berlin, 2010, pp. 3–45. https://doi.org/10.1007/978-3-642-11690-2

[6] Tang, Y., Atkins, E. M., and Sanner, R. M., "Emergency Flight Planning for a Generalized Transport Aircraft with Left Wing Damage," *AIAA Guidance, Navigation, and Control Conference 2007*, Vol. 5, AIAA Paper 2007-6873, 2007. https://doi.org/10.2514/6.2007-6873

[7] Watts, R., Tsiotras, P., and Johnson, E., "Pilot Feedback for an Automated Planning Aid System in the Cockpit," *AIAA/IEEE Digital Avionics Systems Conference—Proceedings*, Inst. of Electrical and Electronics Engineers, New York, 2009, pp. 5.B.2-1–5.B.2-13. https://doi.org/10.1109/DASC.2009.5347472

[8] Watts, R., Claus Christmann, H., Johnson, E. N., Feigh, K., and Tsiotras, P., "Development and Evaluation of an Automated Path Planning Aid," *Journal of Aircraft*, Vol. 49, No. 6, 2012, pp. 1774–1785. https://doi.org/10.2514/1.C031513

[9] Meuleau, N., Plaunt, C., and Smith, D. E., "Emergency Landing Planning for Damaged Aircraft," *ICAPS-08 Scheduling and Planning Applications Workshop (SPARK)*, Assoc. for the Advancement of Artificial Intelligence, 2008.

[10] Meuleau, N., Plaunt, C., Smith, D. E., and Smith, T., "An Emergency Landing Planner for Damaged Aircraft," *Proceedings of the Twenty First Innovative Applications of Artificial Intelligence Conference*, Vol. 21, AIAA, Reston, VA, 2009, pp. 114–121.

[11] Meuleau, N., Neukom, C., Plaunt, C., Smith, D. E., and Smith, T., "The Emergency Landing Planner Experiment," *ICAPS-11*, Assoc. for the Advancement of Artificial Intelligence, 2011, pp. 60–67.

[12] Chen, T. L., and Pritchett, A. R., "Development and Evaluation of a Cockpit Decision-Aid for Emergency Trajectory Generation," *Journal of Aircraft*, Vol. 38, No. 5, 2001, pp. 935–943. https://doi.org/10.2514/2.2856

[13] Bainbridge, L., "Ironies of Automation," *Automatica*, Vol. 19, No. 6, 1983, pp. 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

[14] Strauch, B., "Ironies of Automation: Still Unresolved After All These Years," *IEEE Transactions on Human-Machine Systems*, Vol. 48, No. 5, 2018, pp. 419–433. https://doi.org/10.1109/THMS.2017.2732506

[15] Anonymous, "Aircraft Accident Report: Loss of Thrust in Both Engines After Encountering a Flock of Birds and Subsequent Ditching on the Hudson River, US Airways Flight 1549," National Transportation Safety Board TR NTSB/AAR-10/03 PB2010-910403, Washington, D.C., 2010.

[16] Van Baelen, D., Ellerbroek, J., (René) 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

[17] Borst, C., Sjer, F. A., Mulder, M., Van Paassen, M. M., and Mulder, J. A., "Ecological Approach to Support Pilot Terrain Awareness After Total Engine Failure," *Journal of Aircraft*, Vol. 45, No. 1, 2008, pp. 159–171. https://doi.org/10.2514/1.30214

[18] Borst, C., Mulder, M., and Van Paassen, M. M., "Design and Simulator Evaluation of an Ecological Synthetic Vision Display," *Journal of Guidance, Control, and Dynamics*, Vol. 33, No. 5, 2010, pp. 1577–1591. https://doi.org/10.2514/1.47832

[19] 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

[20] van Oort, E. R., Chu, Q. P., and Mulder, J. A., "Maneuver Envelope Determination Through Reachability Analysis," *Advances in Aerospace Guidance, Navigation and Control*, Springer, Berlin, 2011, pp. 91–102. https://doi.org/10.1007/978-3-642-19817-5_8

[21] Lombaerts, T., Schuet, S., Wheeler, K., Acosta, D. M., and Kaneshige, J., "Safe Maneuvering Envelope Estimation Based on a Physical Approach," *AIAA Guidance, Navigation, and Control (GNC) Conference*, AIAA Paper 2013-4618, 2013. https://doi.org/10.2514/6.2013-4618

[22] Asadi, D., Sabzehparvar, M., Atkins, E. M., and Talebi, H. A., "Damaged Airplane Trajectory Planning Based on Flight Envelope and Motion Primitives," *Journal of Aircraft*, Vol. 51, No. 6, 2014, pp. 1740–1757. https://doi.org/10.2514/1.C032422

[23] Atkins, E. M., Portillo, I. A., and Strube, M. J., "Emergency Flight Planning Applied to Total Loss of Thrust," *Journal of Aircraft*, Vol. 43, No. 4, 2006, pp. 1205–1216. https://doi.org/10.2514/1.18816

[24] Smaili, M. H., Breeman, J., Lombaerts, T. J. J., and Joosten, D. A., "A Simulation Benchmark for Integrated Fault Tolerant Flight Control Evaluation," *Collection of Technical Papers—AIAA Modeling and Simulation Technologies Conference, 2006*, AIAA Paper 2006-6471, Vol. 1, 2006, pp. 563–585. https://doi.org/10.2514/6.2006-6471

[25] "Aircraft Accident Report—El Al Flight 1862," Netherlands Aviation Safety Board, Aircraft Accident Rept. 92-11, 1992, https://reports.aviation-safety.net/1992/19921004-2_B742_4X-AXG.pdf.

[26] Smaili, M., and Mulder, J. A., "Flight Data Reconstruction and Simulation of the 1992 Amsterdam Bijlmermeer Airplane Accident," *Modeling and Simulation Technologies Conference*, AIAA Paper 2000-4586, Aug. 2003. https://doi.org/10.2514/6.2000-4586

[27] Norouzi, R., Kosari, A., and Sabour, M. H., "Evaluating the Effects of Control Surfaces Failure on the GTM," CoRR, Vol. abs/1905.0, 2019, pp. 1–17, http://arxiv.org/abs/1905.09794.

[28] Urnes, J. M., "Dynamic Flight Envelope Assessment and Prediction," *AIAA Guidance, Navigation and Control Conference and Exhibit*, AIAA Paper 2008-6983, 2008.

[29] Keller, J., McKillip, R., and Kim, S., "Aircraft Flight Envelope Determination Using Upset Detection and Physical Modeling Methods," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2009-6259, 2009, pp. 1–21. https://doi.org/10.2514/6.2009-6259

[30] Tang, L., Roemer, M., Ge, J., Crassidis, A., Prasad, J., and Belcastro, C., "Methodologies for Adaptive Flight Envelope Estimation and Protection," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2009-6260, 2009, pp. 1–14. https://doi.org/10.2514/6.2009-6260

[31] Pandita, R., "Dynamic Flight Envelope Assessment with Flight Safety Applications," Ph.D. Thesis, Univ. of Minnesota, Minneapolis, MN, 2011.

[32] De Visser, C. C., "Global Nonlinear Model Identification with Multivariate Splines," Ph.D. Thesis, Delft Univ. of Technology, Zutphen, The Netherlands, 2011.

[33] Lombaerts, T., Schuet, S., Wheeler, K., Acosta, D., and Kaneshige, J., "Robust Maneuvering Envelope Estimation Based on Reachability Analysis in an Optimal Control Formulation," *2013 2nd International Conference on Control and Fault-Tolerant Systems (Systol)*, Inst. of Electrical and Electronics Engineers, New York, 2013, pp. 318–323. https://doi.org/10.1109/SysTol.2013.6693856

[34] Asadi, D., Sabzehparvar, M., and Talebi, H. A., "Damaged Airplane Flight Envelope and Stability Evaluation," *Aircraft Engineering and*

Aerospace Technology, Vol. 85, No. 3, 2013, pp. 186–198.
https://doi.org/10.1108/00022661311313623

[35] Lu, P., Van Eykeren, L., van Kampen, E., de Visser, C. C., and Chu, Q. P., "Double-Model Adaptive Fault Detection and Diagnosis Applied to Real Flight Data," *Control Engineering Practice*, Vol. 36, March 2015, pp. 39–57.
https://doi.org/10.1016/j.conengprac.2014.12.007

[36] Lu, P., van Kampen, E., de Visser, C. C., and Chu, Q. P., "Nonlinear Aircraft Sensor Fault Reconstruction in the Presence of Disturbances Validated by Real Flight Data," *Control Engineering Practice*, Vol. 49, April 2016, pp. 112–128.
https://doi.org/10.1016/j.conengprac.2016.01.012

[37] Maciejowski, J. M., and Jones, C. N., "MPC Fault-Tolerant Flight Control Case Study: Flight 1862," *IFAC Proceedings Volumes*, Vol. 36, No. 5, 2003, pp. 119–124, http://www-control.eng.cam.ac.uk/Homepage/papers/cued_control_7.pdf.
https://doi.org/10.1016/S1474-6670(17)36480-7

[38] Edwards, C., Lombaerts, T., and Smaili, H., "Fault Tolerant Flight Control," *Lecture Notes in Control and Information Sciences*, Vol. 399, Springer, Berlin, 2010, pp. 47–89.

[39] Smaili, H., Breeman, J., Lombaerts, T., and Stroosma, O., "A Benchmark for Fault Tolerant Flight Control Evaluation," *4th European Conference for Aerospace Systems (EUCASS)*, EDP Sciences, 2011, pp. 333–346.
https://doi.org/10.1051/eucass/201306333

[40] Ruijgrok, G. J. J., *Elements of Airplane Performance*, Delft Univ. Press, Delft, The Netherlands, 2009, Chap. 3.

[41] Mulder, J. A., Chu, Q. P., Sridhar, J. K., Breeman, J. H., and Laban, M., "Non-Linear Aircraft Flight Path Reconstruction Review and New Advances," *Progress in Aerospace Sciences*, Vol. 35, No. 7, 1999, pp. 673–726.
https://doi.org/10.1016/S0376-0421(99)00005-6

[42] Vicente, K. J., and Rasmussen, J., "Ecological Interface Design: Theoretical Foundations," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 22, No. 4, 1992, pp. 589–606.
https://doi.org/10.1109/21.156574

[43] Borst, C., Flach, J. M., and Ellerbroek, J., "Beyond Ecological Interface Design: Lessons from Concerns and Misconceptions," *IEEE Transactions on Human-Machine Systems*, Vol. 45, No. 2, 2015, pp. 164–175.
https://doi.org/10.1109/THMS.2014.2364984

[44] Van Paassen, M. M., Borst, C., Ellerbroek, J., Mulder, M., and Flach, J. M., "Ecological Interface Design for Vehicle Locomotion Control," *IEEE Transactions on Human-Machine Systems*, Vol. 48, No. 5, 2018, pp. 541–555.
https://doi.org/10.1109/THMS.2018.2860601

[45] Rasmussen, J., "Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-13, No. 3, 1983, pp. 257–266.
https://doi.org/10.1109/TSMC.1983.6313160

[46] Landman, A., van Oorschot, P., van Paassen, M. M., Groen, E. L., Bronkhorst, A. W., and Mulder, M., "Training Pilots for Unexpected Events: A Simulator Study on the Advantage of Unpredictable and Variable Scenarios," *Human Factors*, Vol. 60, No. 6, 2018, pp. 793–805.
https://doi.org/10.1177/0018720818779928

[47] Cao, A., Chintamani, K. K., Pandya, A. K., and Ellis, R. D., "NASA TLX: Software for Assessing Subjective Mental Workload," *Behavior Research Methods*, Vol. 41, No. 1, 2009, pp. 113–117.
https://doi.org/10.3758/BRM.41.1.113

[48] Hart, S. G., and Staveland, L. E., "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in Psychology*, Vol. 52, No. C, 1988, pp. 139–183.
https://doi.org/10.1016/S0166-4115(08)62386-9

[49] Lombaerts, T. J. J., Smaili, M. H., Stroosma, O., Chu, Q. P., Mulder, J., and Joosten, D. A., "Piloted Simulator Evaluation Results of New Fault-Tolerant Flight Control Algorithm," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 6, 2009, pp. 1747–1765.
https://doi.org/10.2514/1.44280

[50] Borst, C., Visser, R. M., van Paassen, M. M., and Mulder, M., "Exploring Short-Term Training Effects of Ecological Interfaces: A Case Study in Air Traffic Control," *IEEE Transactions on Human-Machine Systems*, Vol. 49, No. 6, 2019, pp. 623–632.
https://doi.org/10.1109/THMS.2019.2919742

[51] Hollnagel, E., "Context, Cognition and Control," *Co-Operative Process Management, Cognition and Information Technology*, Taylor & Francis, London, 1998, pp. 27–52.

[52] Comans, J., Borst, C., van Paassen, M. M., and Mulder, M., "Risk Perception in Ecological Information Systems," *Proceedings of 17th International Symposium of Aviation Psychology*, Wright State Univ., Dayton, OH, 2013, pp. 436–441.

E. Atkins
Associate Editor