The Aeronautical Journal (2026), 1–25
doi:10.1017/aer.2026.10161

RESEARCH ARTICLE

Principles for 'intelligent assistant' systems in future flight deck design: autonomous action integration to reduce pilot workload

Declan Saunders[1], James Blundell[1], Wen-Chin Li[1], Peter Beecroft[2], Linghai Lu[1], Wojciech Tomasz Korek[1] and Wenbing Shi[1]

[1]Safety and Accident Investigation Centre, Cranfield University, Bedford, UK
[2]Rolls-Royce plc, UK
Corresponding author: Wen-Chin Li; Email: wenchin.li@cranfield.ac.uk

Received: 22 August 2025; **Revised:** 9 March 2026; **Accepted:** 9 March 2026

Keywords: autonomous task sequencing; future flight deck design; human-machine interaction; intelligent systems; workload

Abstract

The introduction of advanced automation and human-artificial intelligence (AI) teaming is expected to permit more efficient use of airspace in the face of increasing air transport demand. Additionally, the development of next-generation aircraft to support net-zero has introduced more complexity into the future flight deck and informational requirements. This study evaluates a design for an 'intelligent assistant' system that could share tasks with the pilot during engine failure and pilot incapacitation events, promoting greater reliance on system interaction as workload increases. Four professional pilots were split into two groups to perform six and eight scenarios, respectively. The aim was to identify the task-related information for the designed system to promote transparency to the pilots. Three modalities varied across each scenario (visual, auditory and physical) to evaluate the combination of modality to increase pilot monitoring and interaction with the system. Analysis of participant feedback indicated key limitations to existing human-machine-interaction design, with current operational procedures creating disparity between the system and pilots' authority to handle the scenario. Additionally, the use of audio narration was negatively received by participants, primarily due to the potential overlap between other audio stimuli, masking the perception of task-critical audio prompts and delaying critical flight tasks from being performed. Design considerations were generated for future 'intelligent assistant' systems, with further research required to understand the effect of each modality on pilot reliance on these 'intelligent assistant' systems.

Nomenclature

AI	artificial intelligence
ANC	aviate, navigate, communicate
ATC	air traffic control
ECAM	electronic centralised aircraft monitoring
ENG	engine
FMS	Flight Management System
FSS	Future Systems Simulator
HCD	human-centric design
HCI	human-computer interaction
HMI	human-machine interaction
HUD	head-up display
IA	intelligent assistant
LH2	liquid hydrogen
PF	pilot flying
PM	pilot monitoring

© The Author(s), 2026. Published by Cambridge University Press on behalf of Royal Aeronautical Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.
https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

SA	situational awareness
SAF	sustainable aviation fuels
SiPO	single-pilot operations
SOP	standard operating procedures
TEMP	temperature
THR	throttle
UCD	user-centred design

1.0 Introduction

The development of complex automated systems for the flight deck has created serious human-computer interaction (HCI) concerns in the aviation industry, specifically related to the 'black box' paradigm surrounding these automated systems. The black box paradigm, more commonly associated with artificial intelligence (AI), is the concealment of internal logic sequencing from a system/AI before it presents an output, creating distrust and unease over the reliability and accuracy of the system to perform safety-critical tasks [1–3]. The pilot is expected to actively engage with these complex systems and monitor the progress of any tasking delegated to the system through indirect inputs such as control dials and Flight Management System (FMS) inputs. Development and integration of these AI-powered systems is expected to become commonplace in subsequent years, necessitating human-AI- teaming (HAT), to support the pilots in a more dynamic and rapidly evolving industry [4, 5].

The introduction of 'intelligent assistant' (IA) systems to the flight deck is being explored as a means to enable pilots to assimilate the complexity of AI-powered flight information whilst mitigating corresponding detriment to pilot workload and situation awareness (SA) [1, 6]. The advent of alternative aviation fuels (LH2, SAF, biofuel) will require pilots to handle more complex and novel fuel management systems, creating a gap between legacy information, procedures and proposed operations [3, 7]. IA flight systems are seen as one option for future flight decks, increasing the authority of flight systems to act both independently and collaboratively with pilots to support achievement of mission goals [8–10], although concerns regarding pilot trust and reliance on these systems require further study [11–13].

At a human-machine interaction (HMI) design level, promoting appropriate trust and reliance upon IA systems on the flight deck will be denoted by the degree of transparency the system exhibits to the flight crew. One means to achieve transparency is to leverage multimodal sensory integration of task-related information to avoid overloading pilot cognitive capacity [14–16]. One notable study investigated the optimisation of head-up displays (HUDs) for pilots during taxi tasks using task-related HUD symbology alongside active stick, delivering haptic feedback cues. The interface design used user-centred design (UCD) principles to achieve a task-related organisation of information whilst minimising the opportunity for HUD 'clutter' effects. This enables pilots to switch from active to passive system engagement, such as manual control to system monitoring [17]. The findings suggest that combining the sensory modality of information can support pilot efficiency and safety during safety-critical tasking, indicating the importance of the UCD process to be integrated early into the development of complex systems for the flight deck. However, the growing complexity of airspace utilisation will necessitate a new operational environment for flight scheduling and cooperative airspace sharing [18, 19]. The operational challenges to the industry will likely introduce IA systems to the flight deck to support the precision flying required. However, this will create an HMI challenge related to the transparency of these systems when making actions and/or recommendations on the flight deck [3, 20, 21].

Other studies by Li et al. [22] investigated the effectiveness of touchscreen inceptors as a control mechanism for future flight decks, looking at the visual behaviour differences between pilot flying (PF) and pilot monitoring (PM) duties during landing scenarios. Investigating the division of visual attention between PF to PM duties to propose future single pilot operations (SiPO), the findings indicate that a new HMI will need to be considered for future flight deck design, regardless of SiPO or legacy crewing practices, to support effective scanning behaviour and performance on task. The distribution

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

of information will likely need to adapt to the pilot's visual behaviour, requiring adaptive HMI that can display this information at critical phases of flight [23].

Flight crews contribute critically to aviation safety through their ability to collaborate and draw on multiple feedback loops when responding to abnormal or emergency situations [14, 24, 25]. This collaborative capacity enables pilots to diagnose problems in real time by integrating their knowledge, experience and sensory cues while filtering and prioritising large volumes of information [26, 27]. These intuitive, often unquantifiable aspects of human cognition allow pilots to make rapid judgements based on a sense that 'something is not quite right', supporting a heightened level of SA that has repeatedly proven valuable in safety-critical contexts [6, 25, 28]. Such capabilities remain unmatched by current automated or autonomous systems, which lack the flexible, context-sensitive decision-making required in high-workload environments [2, 29, 30].

Recent regulatory work has begun to outline how more advanced cockpit automation should be integrated into future operations. The European Union Aviation Safety Agency (EASA) position on HAT, as articulated in the Human AI teaming Knowledge and Understanding (HAIKU) project [1], emphasises that IAs must be designed to collaborate with pilots rather than replace them. The document highlights the need for transparency, controllability and shared SA, noting that future flight deck systems will increasingly act as cognitive teammates capable of supporting pilots during high-workload or degraded-performance scenarios. EASA's framework also describes a progression of autonomy levels, from advisory systems to agents capable of taking initiative, provided they remain aligned with pilot intent and can transparently communicate their reasoning [1].

Several aviation-focused IA use cases described in the HAIKU project further illustrate how such systems may operate on the flight deck. These include flight deck IAs that help a single pilot recover from a startle-induced performance drop by directing attention to the most relevant instruments, and IAs that support crews in complex re-routing decisions during severe weather or airport disruptions [1, 5]. These examples demonstrate how IAs can provide real-time cognitive support, reduce workload and enhance decision-making while keeping pilots firmly 'in the loop'.

The HAIKU project concluded with the recommendation that future aviation IA use cases are needed, particularly use cases involving more complex IAs with the capability to independently initiate tasks. Such use cases include rare but high-risk scenarios – such as pilot incapacitation – that underscore the need for IA systems capable of assuming safety-critical tasks when required [31, 32]. Here, the IA must be able to reduce the remaining pilot's workload, enabling them to reallocate cognitive resources to the strategic decision-making required to safely manage and land the aircraft. Incapacitation may also occur when the other pilot is temporarily absent from the flight deck (e.g. during in-flight rest or a personal break), requiring the system to independently manage safety-critical tasks without immediate human oversight. When the pilot returns, the system must be able to communicate the rationale behind its actions and support a smooth transition of control back to the human operator [5, 33]. Ensuring transparency in the system's behaviour and decision-making processes is therefore essential for maintaining pilot situational awareness and understanding of the system's intentions and outcomes when it acts autonomously [14, 32]. Together, these regulatory developments and emerging use cases highlight a clear need to understand how IA should interact with pilots, how it should communicate its reasoning and how control transitions should be managed – providing the motivation for the present study.

This paper describes the development of an IA system, one designed to autonomously complete the engine shutdown checklist in response to a high engine oil temperature event. Four professional pilots evaluated the interaction logic of the system for the task and provided user satisfaction scores to determine the suitability for the purpose. The novelty of the study is the application of autonomous tasking to checklist actions with minimal pilot intervention to investigate the usability of such a system to handle a critical flight task. This paper describes the early development and evaluation of an IA system.

The development of the IA system and scenario setup is described in Section 2. Next, the results and interview analysis from each group are explained in Section 3. Finally, the recommendations of the study and design principles for future development of IA flight systems are presented in Section 4, with conclusions and future scope of research in Section 5.

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press



Figure 1. *(a) FSS configuration and location of five customisable large screens and two smaller screens; (b) experimental layout in X-Plane 12 configuration.*

2.0 Methodology

2.1 Simulator facility

This research was conducted using the Future Systems Simulator (FSS), developed by Cranfield University and Rolls-Royce. The FSS is a highly adaptable, fixed-base flight simulator designed for rapid prototyping of current and future flight deck configurations [34]. It features aircraft control and display systems that present information on seven customisable touchscreens, including two side screens and a large overhead panel display (Fig. 1(a)). Additionally, there are two physical side sticks and a motorised dual-engine throttle lever.

For this study, the X-Plane 12 Airbus 330 flight and virtual flight deck model (Fig. 1(b)) was configured to communicate (via User Datagram Protocol (UDP)) with a Unity-based application that visualised a MATLAB-controlled 'intelligent' system prototype (Fig. 1(b) and Fig. 2). The prototype presents traditional checklist information to the pilot along with additional cues related to current and planned actions of the 'intelligent' system.

2.2 Intelligent system design

The study was designed to clarify the task requirements of an 'intelligent' system in an engine high oil temperature scenario while in cruise flight, to identify the information required, for pilots, for the 'intelligent' system to conduct its tasks transparently. Initially, using the current pilot sample as subject matter experts (SMEs), a hierarchical task analysis (HTA) of the scenario was performed to identify the

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

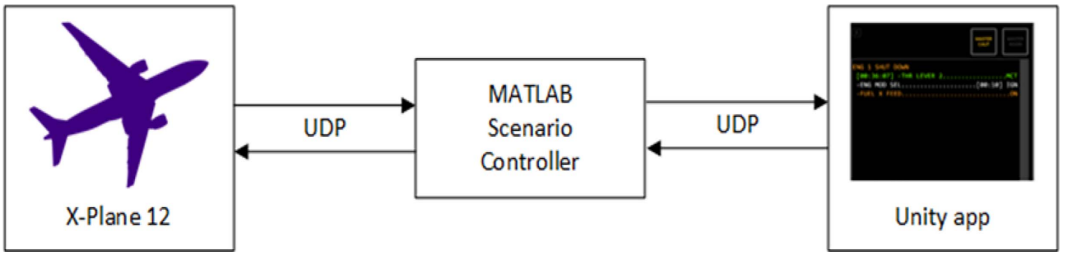


Figure 2. *The X-Plane – Unity communication interface, with the 'intelligent' system running through a custom MATLAB script.*

current composition of PF, PM and aircraft tasks. Tasks were further analysed to define the modality that communication for cross-checking and confirmation of task completion occurred. An extract of a tabulated version of the HTA can be seen in Table 1, which highlights omitted actions due to PM incapacitation (red) and tasks required by the system introduced (green).

SMEs commented that the task requirements of the 'intelligent' system should focus exclusively on the checklist actions, leaving the high-level actions and decisions to the pilot. If the system is required to make an action that would affect the flight's operation, the system should actively engage the pilots in the decision-making process through cross-checking and annunciation of intended actions. For example, switching an engine's master switch to 'OFF' was deemed too significant to be solely allocated to the 'intelligent' system. Scenario setup for the study would consider two separate conditions: (1) fully autonomous checklist completion with pilot supervision (out-seat), and (2) semi-autonomous checklist completion with pilot intervention (in-seat). The task-requirements and task-specific information to the pilots in both conditions would be identified, which can support future development of autonomous tasking on the flight deck. This would also conform to human-in-the-loop design and should reduce the cognitive load of the remaining pilot when monitoring the system and handling their respective tasks [6, 20]. The SMEs also provided the estimated timing of actions and time-to-completion for checklist tasks, allowing for the development of time-driven actions for the system to perform in the absence of pilot intervention.

It was important to design an 'intelligent' system display that could communicate effectively with the pilots to bring them into the loop and build SA of the system, in advent of an engine failure event and autonomous actions to secure the engine. The effect of modality on this interaction between the pilots and system was an important area of research for this study to evaluate the effectiveness of communication loops during high workload tasking [17, 33]. The considered modalities were: (1) visual alerting and colour coding, (2) audio narrative of actions and time remaining and (3) back-driven throttles. The system would perform the tasks of the PM after an initial delay time (5–10 s intervals), allowing pilot cross-checking before an action was performed and to simulate traditional checklist handling. The study was concerned with the information requirements from the system at this stage, as opposed to the interaction between the pilot and the 'intelligent' system, understanding the appropriate delay and modality to support pilot SA building of the system and flight environment, with the authority to stop the system at any point of the scenario. The design of the system display, seen in Fig 3(a) and 3(b), was a low-fidelity model of the system integrated into the FSS environment, looking at the information requirements of the pilot from the system to promote transparency of the system to support SA building of actions and consequences in flight.

Once the aircraft model is loaded in X-Plane 12, the MATLAB script starts the experiment sequence. It waits for a manual or time-based trigger to start inducing high oil temperature in the left engine (via X-Plane's DataRef functionality) and waits for an X-Plane's high oil temperature signal to send a warning and a relevant checklist signal to the Unity app. Then, the Unity app plays a warning signal and shows the checklist. While the automation scenario is in progress, the MATLAB script sequentially controls X-Plane's data (such as the engine throttle lever or an engine master switch) and updates the Unity

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

Table 1. *Short extract of HTA for engine 1 high oil temperature, with PF and PM tasks identified. Black text indicates PF tasks, with red text representing PM tasks that are replaced/required by the 'intelligent' system in green text. The communication method to confirm each phase was represented: visual (V), audio (A) and physical/haptic (P)*

Aircraft	Modality of communication	Pilot flying	Modality of communication	Pilot monitoring	Modality of communication
'Caution'	Communicate (A, V)	–	–	–	–
		Preliminary ANC		Preliminary ANC	
		Cancel 'Caution' Action	Communicate (A, V)	–	–
		Request PM to confirm 'Caution'	Communicate (A)		
Present ENG (1) OIL HI TEMP on ECAM	–	Read ENG (1) OIL HI TEMP on ECAM	Communicate (A)	Read ENG (1) OIL HI TEMP on ECAM	Communicate (A)
Present ENG Information	Communicate (V)	Confirm ENG (1) OIL HI TEMP on ENG status page	Communicate (A)	Confirm ENG (1) OIL HI TEMP on ENG status page	Monitor
Present THR (1) TO IDLE on ECAM	Communicate (V)	Read THR (1) TO IDLE on ECAM	Communication (A)	Read THR (1) TO IDLE on ECAM	Communication (A)
		Declare intention to THR (1) TO IDLE	Communication (V, A)	–	–
		Confirm intention THR (1) TO IDLE	Communicate (A)	Confirm intention THR (1) TO IDLE	Communicate (A)
		Action THR (1) TO IDLE	Aviate, Communicate (V)	–	–
Present ENG Information	Communicate (V, A)	Observe ENG Information	Monitor	Observe ENG Information	Monitor

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

Table 1. *(Continued)*

Aircraft	Modality of communication	Pilot flying	Modality of communication	Pilot monitoring	Modality of communication
Present ENG MASTER (1) TO OFF on ECAM	Communicate (V)	Read ENG MASTER (1) TO OFF on ECAM	Communicate (A)	Read ENG MASTER (1) TO OFF on ECAM	Communicate (A)
		Confirm ENG MASTER (1) TO OFF on ECAM	Communicate (A)	–	–
		Action ENG MASTER (1) TO OFF	Aviate, Communicate (A, V)	Action ENG MASTER (1) TO OFF	Aviate, Communicate (A, V)
'Caution'	Communicate (A, V)	–	–	–	–
Present ENG (1) SHUT DOWN on ECAM		–	–	–	–

Abbreviations: Aviate, Navigate, Communicate (ANC); Electronic Centralised Aircraft Monitoring (ECAM); Engine (ENG); High (HI); Temperature (TEMP); Throttle (THR).

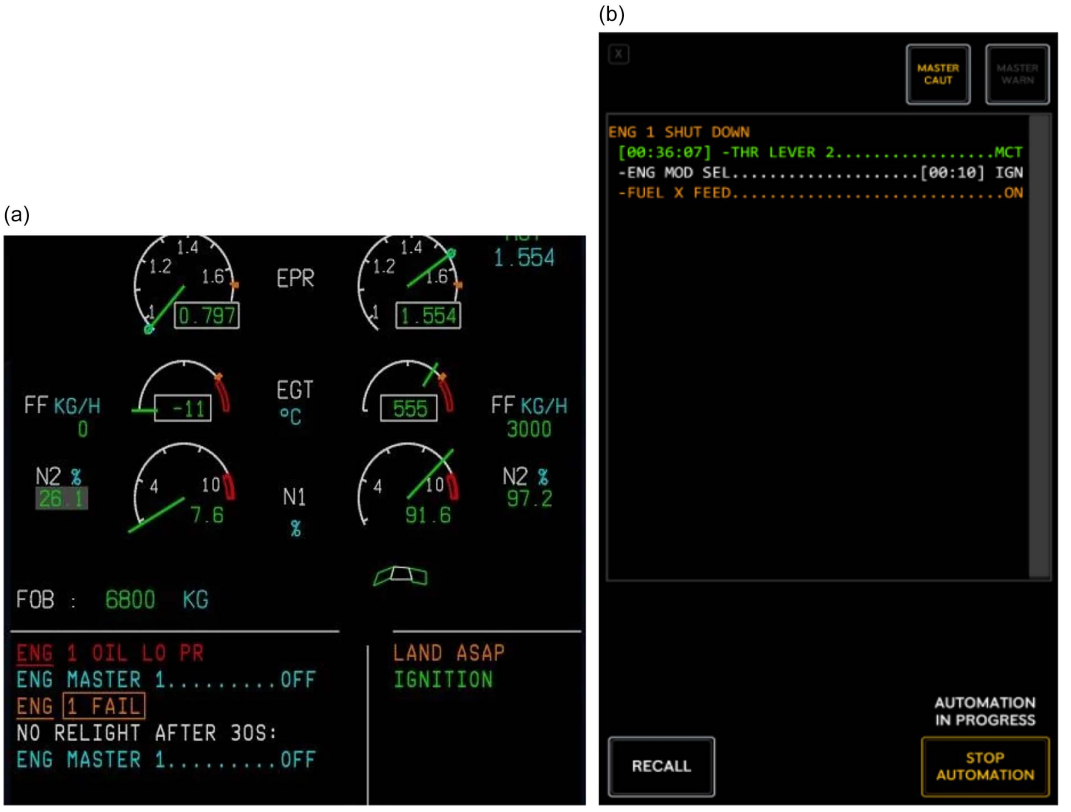


Figure 3. *(a) Comparison between legacy ECAM display with failure event and checklist actions; (b) the developed autonomous system display with checklist actions and time-driven tasking indicated.*

app with its progress; in case it is a baseline scenario (no automation), the MATLAB script monitors DataRefs in X-Plane to check if the pilot completed an action. If voice cues are enabled, Unity plays the relevant audio recording, and if the physical throttle cues are tested, it moves the Arduino-based throttle lever. The Unity app updates the progress of each checklist item automatically: completed items appear green, incomplete items are amber, and the item currently under automation is in white, with a timer shown that allows the pilot to interrupt it before the automation acts (Fig. 3(b)). The outside world view is provided via an additional X-Plane instance on a separate PC, generating the image for the three projectors.

Figure 3(a) represents the legacy engine management display with checklist actions for managing an engine 1 low oil pressure event. Colour vision is facilitated by three types of cone photoreceptors in the human eye, capable of capturing short, medium and long wavelengths. The average human being is capable of perceiving wavelengths between 400 nm (violet) and 700 nm (red), with spectral sensitivity peaking at 555 nm (yellow to green) and is able to differentiate more than 100 hues [35, 36]. The use of colour within the cockpit warning and alerting system ranges from red to green to indicate the severity and urgency of crew intervention, which is supported by the visual acuity of human beings in colour detection among a cluttered visual scene. On legacy displays, the colour-coding for information is as follows: failure items appear in red, crew awareness without immediate action required appear in yellow, normal operation items appear in green, remarks and information to guide the crew appear in white, and actions to be carried out appear in blue [37].

The developed system display, shown in Fig. 3(b), utilises the fully digital FSS configuration to offer a separate display that represents the tasks of the system and a countdown timer to each action as seen

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

in Table 1. Colour coding was updated in the current research, given the removal of one pilot due to incapacitation, and to study the optimum strategies for colour-coding displays to reduce pilot workload with an 'intelligent' system conducting pilot tasking independently [26, 35]. The Unity app updates the progress of each checklist item automatically: completed items appear in green, incomplete items are in amber, and the items currently under autonomous tasking are in white, with a timer shown that allows the pilot to interrupt it before the system acts (Figs 3(a) and 3(b)). Pilots also have the option to 'recall' previous actions of the system if they were not present on the flight deck, which would timestamp each action and display it to pilots for transparency behind autonomous tasking in real time. This functionality was discussed among the four pilots of the study, specifically, whether this would become a standard operating procedure (SOP) when returning to the flight deck after scheduled rest, and if the information would be sufficient to support awareness building into the scenario.

2.3 Pilot sample

There were four pilots who participated in the research. The pilot participants were split between military pilots, a commercial and test pilot, and a private pilot (2:1:1, respectively). The rank of all participants was Captain, with an average flying experience of 5,475h ($SD = 2185.61$). The majority of pilots (3/4) had experience flying Airbus fleet types (A330, A350, etc). The experiment was approved by Cranfield University Research Ethics System (CURES). Data collection, analysis and handling were in accordance with the UK Data Protection Act 2018.

2.4 Procedures and scenario

This study was designed to identify the information requirements for a transparent 'intelligent' system, which could autonomously complete tasks in place of an incapacitated pilot, with opportunity for the functional pilot to intervene if required. An engine shutdown checklist was chosen for the study to demonstrate how the system can help prevent unnecessary damage to the engine after an engine high oil temperature warning. The system must also be capable of communicating its actions and rationale to the pilot, under exceptional circumstances, if necessary, to demonstrate its applicability to the flight deck. Sensory modality was controlled as one method of increasing pilot ability to assess and monitor the autonomous tasking during two conditions: (1) in-seat and monitoring the autonomous system actions in a semi-autonomous condition, and (2) out-of-seat and understanding the actions done by the autonomous system in a fully autonomous condition. In both conditions, the pilots would be alone on the flight deck, simulating the worst-case scenario of a pilot incapacitation event. The rationale for this was to support the design of a system that can cooperatively work and communicate to a solo pilot, maintaining a manageable workload where the pilot is still able to exert their operational authority over the system.

To this end, the study split participants into two separate groups: (1) control group performing scenarios 1–3 and 5–7, and (2) test group performing all scenarios, as shown in Table 2. Group one would be exposed to each modality separately to provide high-level comments on the effectiveness of these modalities on the flight deck to communicate with the pilot during periods of high workload. Group two would be exposed to each modality separately and then combined in scenarios 4 and 8, to investigate the combinations of modality to support pilot SA building in both conditions mentioned above, and what this would look like on the flight deck. All experiments were conducted on the Cranfield FSS, with participants configured as a solo pilot, after a PM incapacitation event on the flight deck. A 30-min briefing session was provided to each group, explaining the condition of the scenario and the intended outcome of the scenario. Pilot actions during the scenario were monitored. Participants signed the consent form before being invited to familiarise themselves with the flight simulator environment. All scenarios were presented to the participants in the cruise segment of the flight, near London Heathrow airport (ICAO: EGLL).

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

Table 2. *Experimental scenario design to investigate the information requirements of an 'intelligent' tasking system, using sensory modality to communicate to pilots on the system's actions and progress*

	Scenario	Modality of communication	Description
1	High engine oil temperature during cruise flight at EGLL	Visual	The scenario started in cruise flight, with a short delay prior to alerting the engine high oil temperature (left engine) and automated engine shutdown. The scenario was intended to monitor the participants' reception towards the automated engine shutdown procedure during high engine oil temperature. Only visual aids on the automated checklist guided the participant on the progress of the automation.
2	High engine oil temperature during cruise flight at EGLL	Visual & audio	Same as scenario 1 The scenario was intended to monitor the participants' reception towards the automated engine shutdown procedure during high engine oil temperature, with visual and audio aids to complete the checklist and follow the automation.
3	High engine oil temperature during cruise flight at EGLL	Visual & physical	Same as scenario 1 The scenario was intended to monitor the participants' reception towards the automated engine shutdown procedure during high engine oil temperature, with visual and physical aids to complete the checklist and follow the automation.
4	High engine oil temperature during cruise flight at EGLL	Visual, audio & physical	Same as scenario 1 The scenario was intended to monitor the participants' reception towards the automated engine shutdown procedure during high engine oil temperature. All three cognitive aids were provided to participants during scenario (visual, auditory and physical).

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

Table 2. *(Continued)*

	Scenario	Modality of communication	Description
5	High engine oil temperature, return to cockpit during cruise at EGLL	Visual	The scenario started with participant absent from cockpit while in cruise flight. When the participant enters cockpit, a high engine oil temperature (left engine) has occurred with an automated response to secure the engine. The scenario was intended to study the identification and comprehension of the automation actions when returning to cockpit after the engine shutdown. Only visual aids on the automated checklist display, with recall functionality and timestamp for legacy actions.
6	High engine oil temperature, return to cockpit during cruise at EGLL	Visual & audio	Same as scenario 5 The scenario was intended to study the identification and comprehension of the automation actions when returning to cockpit after the engine shutdown. Visual and auditory aids were available to participate to aid comprehension of automation actions and progress in the checklist when returning to cockpit.
7	High engine oil temperature, return to cockpit during cruise at EGLL	Visual & Physical	Same as scenario 5 The scenario was intended to study the identification and comprehension of the automation actions when returning to cockpit after the engine shutdown. Visual and physical aids were available to participate to aid comprehension of automation actions and progress in the checklist when returning to cockpit.
8	High engine oil temperature, return to cockpit during cruise at EGLL	Visual, audio & physical	Same as scenario 5 The scenario was intended to study the identification and comprehension of the automation actions when returning to cockpit after the engine shutdown. All three cognitive aids were provided to participants during the scenario (visual, auditory and physical).

2.5 Post-scenario interviews

After completing each of the eight scenarios, pilots took part in a 20-min (approximate) unstructured interview with a member of the human factors team to provide feedback on the information requirements and expectations of the system to support pilot monitoring and SA building. The unstructured interview promoted exploratory questioning and was best fitted to the aim of the study. The focus of the post-scenario interviews was fixed on five areas of the study, which would generate areas for future study. These five areas were: (1) Autonomous Checklist, (2) Audio Usage/Assistant, (3) Simulator & Scenario, (4) Pilot Awareness, and (5) Operation & Procedures. If these areas were not discussed, prompting by the human factors team highlighted these areas to generate discussion and potential design considerations for subsequent analysis.

2.6 Data analysis

Data was analysed using NVivo 14 software. Post-scenario interview feedback across all scenarios listed in 2.4 Procedures and Scenarios were collated in NVivo and analysed based on their respective areas of concern, for this study: (1) User Experience, (2) Accountability of System, and (3) Expectation vs Reality. Data was coded into categories based on the above-mentioned areas of concern. The participants were required to provide their subjective thoughts and satisfaction to maintain SA on the procedure of engine 1 shutdown and generate design recommendations on supporting elements when in SiPO conditions.

3.0 Results

Thematic analysis was used to systematically identify and interpret themes and codes within a subjective dataset, consisting of pilots' interview feedback and experiences when interacting with the system. First-layer coding involved the transformation of pilot comments into textual data, which addressed the key design recommendations of the study. These comments were separated based on the two conditions listed in 2.2 Intelligent System Design to identify how pilot experience contributed to specific design recommendations (Figure 4). These findings could help identify how the expectations of an autonomous tasking system would differ in these different operating environments, and what level of assurance each operator would expect. For example, the system promoted time-driven actions, which are more commonly associated within the military domain, as discussed with SMEs during the development of the system. However, the commercial pilots and private pilots are taught to assess and identify issues before responding with actions, so the time-driven tasking would have conflicted with their mental model of the procedure for engine shutdown in these scenarios, indicated with more codes. Once completed, the second-layer coding involved layering the data into themes and codes to visually understand where the distribution of feedback is weighted towards and what areas needed urgent attention based on the pilot's analysis. As seen in Figs 5 and 6, the generation of themes was based on the areas of focus for the study, broken down into sub-themes to identify the root cause of the feedback from pilots. The codes related to the individual feedback from the pilots corresponded to the individual themes. The main results of the paper were divided into five categories: (1) Autonomous Checklist Display, (2) Audio Usage, (3) Simulator & Scenario, (4) Pilot Awareness and (5) Operations & Procedures. These five categories allowed for consolidation of pilot feedback into key areas of development, based on the high-level analysis in Figs 5 and 6.

3.1 Group one interview analysis

First-layer coding was applied to data, to categorise by priority and urgency of redesign for the purpose. Data was categorised based on the pilot's location when starting the scenario to identify the change of information requirements in both conditions. Second-layer coding categorised this data further into three areas seen in Fig. 5, to reflect how the design supported/limited their monitoring capacity of the system. The recommendations provided by the participants incorporated all scenarios presented, as seen in Table 2.

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

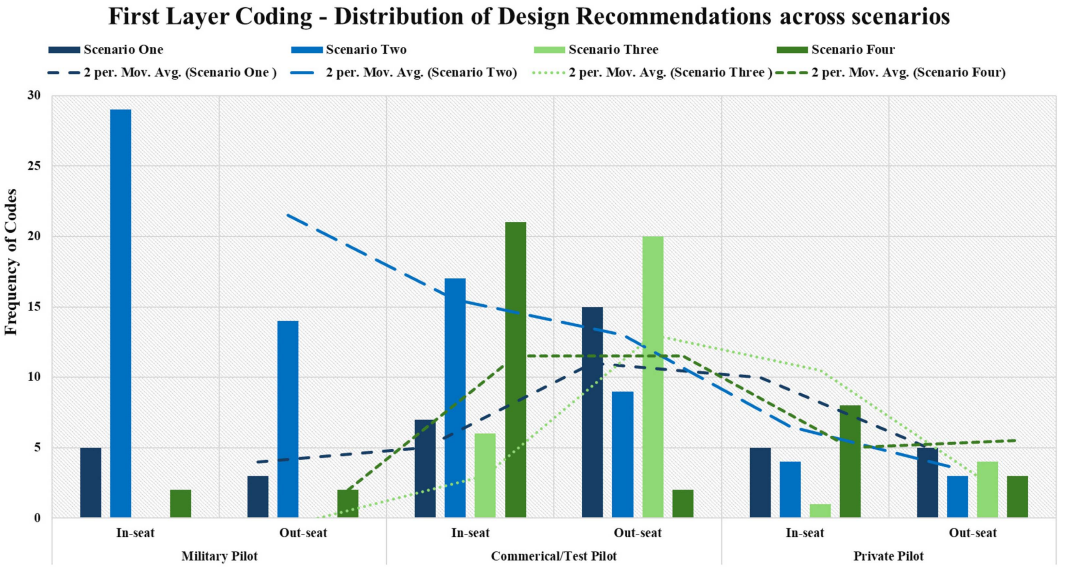


Figure 4. *First-layer coded analysis of design recommendation comments based on background experience of the participant. Scenarios 1–4 are indicated in both conditions (in-seat and out-seat) for all participants, as discussed in Table 2. A two-point moving average represents the variation in comment frequency based on experience.*

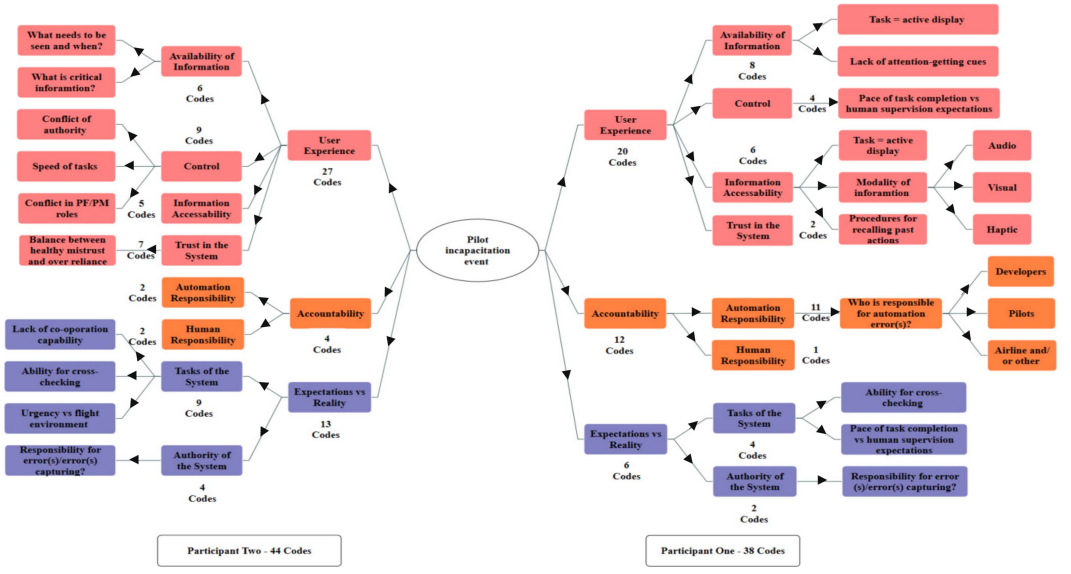


Figure 5. *Second layer coded analysis of design recommendation comments for participants one and two, categorised into three areas of interest: (a) User Experience (pink), (b) Accountability (orange) and (c) Expectation vs Reality (purple).*

3.1.1 Autonomous checklist display

Participant one commented on the lack of saliency of information when autonomous actions are taking place, specifically related to the ability to cross-check with the 'intelligent' system: 'Cause there's no other signal to me that the action has been completed. Whereas if I was doing it myself with another pilot, we would be going right throttle number 1 confirmed move'. Primary tasks were delayed to secondary

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

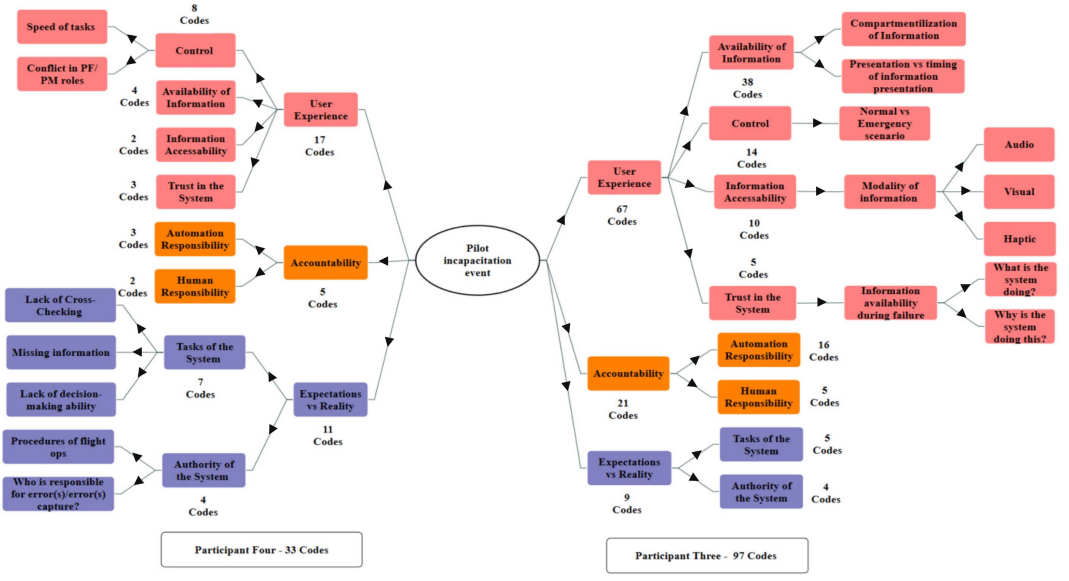


Figure 6. *Second layer coded analysis of design recommendation comments for participants three and four, categorised into three areas of interest: (a) User Experience (pink), (b) Accountability (orange), and (c) Expectation vs Reality (purple).*

tasks, while monitoring the system's actions, with a healthy mistrust of 'intelligent' systems prioritising the attention of the participant. This comment is an example of expectations vs reality; the system was performing actions of which the pilot was unaware or was unable to maintain sufficient awareness over. Participant two commented that the display becomes over-saturated with information, making it difficult to scan and assimilate information quickly:

> 'You're expending a certain amount of your capacity to watch what the computer is doing and try and figure out what it's doing because it's not telling you what it's doing. . . Now I can't be PF because I'm reading the ECAM and I'm looking to see what the computer is doing.'

Further comments referenced the colour-coding of textual information, which became easily saturated on the small display in low-light conditions of the FSS environment. This comment related to user experience and expectations vs reality because the system was not communicating effectively with the pilot for quick comprehension over the system status.

Additionally, comments from participant two addressed how the philosophy between human-automation teaming needs to be further developed in future scenarios, with respect to the responsibility and accountability of the system's role in handling critical tasks. This particular comment related to the accountability of the system. The 'intelligent' system would handle part of the checklist without pilot intervention, and then provided a partly completed checklist for the pilot to finish, consisting of critical actions that 'intelligent' systems could not handle (Air Traffic Control (ATC) communication, confirming tasks, moving fuel switch to 'OFF'): 'And that's because it doesn't tell you when it's doing the drills. . . I don't know when I'm stepping in and then I'm picking up half a drill now because half a drill is done'. Pilots commented on the need for cross-checking systems for both 'intelligent' systems and pilot actions to enable a collaborative working environment, as opposed to an 'action and response' environment.

3.1.2 Audio usage

Participant one commented that audio usage on the flight deck should be reserved for high-level tasks and outcomes of the 'intelligent' system, as opposed to congesting the sensory modality with checklist

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

actions: 'It will automatically shut the engine down and I want it to tell me after the fact this has happened just so you're aware, this has really happened'. The participant became concerned that ATC communications could become saturated if utilised with the system readouts. This comment relates to expectation vs reality of the system, which prioritised certain information considered less relevant to the task. Additionally, participant two expressed similar concerns as participant one, although favoured an audio assistant that could replace some tasks of PM, providing specific high-level verbalisation and chimes to indicate flight condition and overall health, as opposed to checklist reading and autonomous task sequencing. Associated tones/audio with verbal stimuli would benefit pilot awareness and monitoring of an 'intelligent' system, although it should only be triggerable upon pilot request: 'It hasn't said to me like I'm actioning ECAM thrust, watching with one eye on it. . . It needs to verbalise what it's doing, not what's on ECAM, but what it's doing. . . So you can cross-check what it's doing against the ECAM'. Reliable and intuitive information communication should be prioritised through this already-congested stimulus, to promote awareness and prevent over-saturating the pilots.

3.1.3 Simulator and scenario

Participant one commented that the scenarios conducted did not warrant the level of autonomous action, with longer delay durations expected when pilot input is required. Instead, the delay promoted more monitoring of the system, delaying critical actions that pilots would have preferred to occur before the checklist actions. These tasks included: ATC communication, diverting and route planning, diagnostic and clean-up of the flight deck, and general risk analysis of the failure condition vs the condition of the aircraft and environment: 'So we want to take action. It will automatically shut the engine down and I want it to tell me like after the fact like this has happened just so you're aware, this has really happened'. This is related to user experience for the system, which could not provide sufficient timing to enable pilot satisfaction when handling the time-driven tasks to the scenario conducted.

Participant two commented that the autonomous actions should be permissible with pilot actions, keeping pilots in the loop of the system and enabling faster pilot input. The concerns raised by the participant addressed the lack of cohesion between autonomous actions and the pilots' independent actions, which then occupies attentional resources to monitor and confirm the systems actions before performing their respective duties: 'I've got to start descending it before I finish the drill. . . Where is normally the PF would sort all that out while the PM runs a drill'. This is related to the accountability of the system, as there was a conflict in the authority of actions and tasking between the system and the pilot. Redesigning the FSS clutch pack would enable more control to the pilots, allowing pilots to remain on the controls without disrupting the autonomy tasking.

3.1.4 Pilot awareness

Participant one made no specific comments related to pilot awareness of the 'intelligent' system.

Participant two commented that there was a lack of attention-getting cues for cross-referencing autonomous actions with pilot oversight; promoting division between the two systems (pilot and autonomy) and creating distrust among pilots, expending a percentage of visual attention to monitor the progress of autonomy: 'You're expending a certain amount of your capacity to watch what the computer is doing and try and figure out what it's doing because it's not telling you what it's doing. . . I'm also concerned if it's going to move the right switch'. This comment related to the user experience and accountability of the system; the pilot felt the system was 'hiding' decisions from them and was concerned with the overall authority of the system. There were additional comments related to the split attention of the pilot when monitoring the autonomous action progress, and potential risk of misidentifying or lapses in observation that could adversely affect the safety of a task, such as landing: 'But now I'm completely split (listening to audio and monitoring checklist) and it'll be very easy to get lost in a process and miss something like the fact that I hadn't armed the altitude descent'. This comment related to the user experience when monitoring the system, which was considered confusing with many different stimuli simultaneously. Recommendations concluded that a redesign of the warning indicators

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

to represent the urgency of tasks would help promote awareness and pilot intervention, critical if the philosophy of SiPO remains with pilot intervention required for the future flight deck.

3.1.5 Operations and procedures

Participant one commented that the urgency of action was not warranted by the failure case presented, with plenty of altitude and time to resolve the engine high oil temperature. However, the advent of timed actions with limited cross-checking between system and pilot created unease, primarily around the authority and acknowledgement of consequences to each action: 'The one thing that did make me feel uncomfortable is like. . . no one is *cross* checking. But if we compute a cross-check itself, and if I disagreed, I got no control'. This relates to expectation vs reality of the system; the scenario did not warrant the level of authority in the system and thus was not a suitable test to understand pilot satisfaction with the system. Instead, the scenario tested the concept of autonomous tasking with pilot supervision and what the informational requirements were. Subsequent trials should look at different phases of flight and how 'intelligent' system dependency/usage changes in these situations, additionally, with more serious or cascading failures to prompt 'intelligent' systems' usefulness in these situations. Participant two commented that cross-checking procedures prevent the inadvertent movement of critical switches, which promotes a safety-related procedure that is removed in favour of autonomous actions: 'Those two are critical switches, we'd never move a critical switch or a guard switch without confirmation first, so I'm checking it's not about to turn the wrong engine. . . So I've got nobody to confirm it, so hopefully I've got that correct'. This relates to accountability and expectation vs reality of the system; the system promotes time-driven tasks in favour of pilot intervention to complete the checklist, which opens a potential risk related to inadvertent actions in critical flight phases. Future studies should investigate the time required for pilot awareness of these critical switches being moved by 'intelligent' systems and generally propose guidelines and expectations for the authority of the system.

3.2 Group two interview analysis

Following coding analysis, the data were categorised by the priority and urgency of the comments, focusing on the high-level design recommendations for future scope and action. Participants' satisfaction and overall discussion on each element of the trial were valuable information that will support the continual development of flight deck integration for human-automation teaming. The distribution design recommendations are seen in Fig. 6.

3.2.1 Autonomous checklist display

Participant three commented that the pacing of autonomous actions was insufficient to promote monitoring capacity in pilots: 'So I wanna monitor what's going on, but this checklist is moving on regardless', with participant four confirming the impracticality of offering a longer delay to the expense of visual scanning behaviour in traditional two-crew aircrafts: 'The problem with having long delays is that somebody has to. . . Well, not has to, you're drawn to monitoring that to the expense of almost everything else'. Consensus among participants was that checklist actions would be secondary to reconfiguring the aircraft for the loss of thrust and an electrical generator, prioritising the descent and ATC coordination of support prior to securing the engine incident. This relates to user experience for the system, which prioritised system actions before pilot intervention to understand the level of authority for autonomous systems. The autonomous actioning of engine shutdown was generally appreciated for reducing the workload of the pilots, although the application of autonomous authority to action critical tasks (fuel cutoff switches) needs pilot monitoring and confirmation in all situations: 'That is a recommended checklist, but actually I'm going to exercise my command decision and say no. I'd rather leave that engine running'.

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

3.2.2 Audio usage

Participant three commented on the utilisation of audio on the flight deck, which was favoured for high-level critical actions (such as cross-checking actions, or aircraft deviation from allocated flight path), as opposed to checklist items and autonomous action progress. The concern is the communication between ATC and checklist audio would become saturated and diminish the pilots reception of either source: 'It's very calm and it's very calming, but if someone else was talking over the top of that. . .you were trying to coordinate a descent or something, you wouldn't hear what you'd say because your auditory channel is, you're gonna be focused on the radio call, not on that'. This comment was related to the user experience of the system, which was seen to wash out the effectiveness of audio for warnings and attention-getting on the flight deck when overused for trivial tasks. Participant four had a similar outlook on the utilisation of audio on the flight deck, although commented on the ability to activate/deactivate the audio verbalisation: 'I quite like the audio. It's actually quite nice to have a voice. . . But it would be nice if there was some like, don't talk kind of, I'm thinking about this and then go back. This was considered in expectation vs reality coding, as the pilot's ability to control and activate the audio was removed in favour of audio prompting to support SA building of autonomous tasking. It was considered that an audio tone or chime would be more suitable for indicating completion of autonomous actions than verbal audio, although it acknowledged that the risk of confusion might be reduced with the verbalisation of actions and time to completion, in the scenario of pilot incapacitation.

3.2.3 Simulator and scenario

Participant three commented again on the urgency of the scenario, with flight level and uncontested airspace at the point of failure unrealistic, to the rushed shutdown of the engine at the expense of other actions/tasks. One primary task missed by the system was to set maximum continuous thrust (MCT) before shutting down the failure engine, causing the aircraft to drift from its assigned altitude without communication/coordination with ATC: 'Because it took a long time for us to set MCT. . . and I'm now needing to drift down. I might need to coordinate; I need to descend'. Furthermore, the pacing of the autonomous actions was not sufficient for pilot monitoring and acknowledgement, with concerns that incorrect or rushed actions could further harm the safety of the flight: 'But I don't know that the cadence allows me to monitor what's happens before it moves on to the next thing, and I don't think there's a criticality in this scenario to move and to switch that engine off from that hurry'. These comments related to the user experience and expectation vs reality of the system; the system did not provide sufficient time for pilot intervention and actioning to support the actions required by the checklist task. The shutdown of an engine is a critical task that would normally require crew coordination and agreement before the shutdown of the engine, although there is no cross-checking capacity in the system to engage the pilot to prevent the action, except to turn the system off completely. Further, though not a factor for this present study, the overall authority and responsibility of an 'intelligent' system to handle checklist actions and engine shutdown when the pilot is not present on the flight deck is an area of concern, requiring further investigation to identify the urgency and operational practice to permit such tasking without human oversight.

Participant four commented that the saliency of system information was not sufficient for pilot awareness of the task sequence and progress of these actions by the system, with colour saturation of the colour palettes used not suitable for the conditions tested: 'So I wasn't sure whether that was to tell me that the scenario was being run or whether the system was doing something automatically'. The scenario was unfavourable to the solo pilot, as standard procedures would have one pilot handling checklist actions and one pilot flying the aircraft, whereas the scenarios conducted had the solo pilot swapping between two roles, creating uncertainty about their responsibility and duties to the flight. This comment relates to the user experience of the system, with the design of the colour sequencing not supporting attention-getting to critical areas on the flight deck. Several design considerations, such as audio chimes and colour-coded cross-checking of autonomous actions, could reduce the workload of the solo pilot in the event of pilot incapacitation and an engine failure scenario.

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

3.2.4 Pilot awareness

Participant three commented on the effectiveness of attention-getting cues for the autonomous checklist display, which were not sufficient at promoting active engagement with the task sequencing in the same manner as two-crew procedures. The passive pilot engagement made the participant uncomfortable with the authority of the system to handle critical tasks with minimal human supervision, prompting additional warning chimes or visual cues to draw attentional resources from pilots to the system: 'But to me, you know, a master warning would be more useful, especially for that to have the cavalry charge, like when you disconnect the autopilot, something significant just happened'. The design incorporated a visual timer to indicate time-to-completion of autonomous actions, and time for pilot input to be completed within, although was addressed by participant four that the saliency of this information was insufficient to inform the pilot of the meaning: 'It wasn't immediately apparent then, I thought hang on, it hasn't changed, the timer has finished, but it hasn't done it. And then I thought, oh, that's for me to do it'. These comments by participant three related to the user experience of the system, which failed to sufficiently satisfy the pilot to monitor the system's actions and time-to-completion.

Additional concerns were raised by participant three, who commented that the use of verbalisation to promote active pilot engagement to autonomous actions could actually be counterproductive, instead drawing attention away from more critical tasks, such as ATC communication and see-and-avoid practices: 'And that's another big bit that automation doesn't, is it becoming a distraction because it's, you're not part of it, so you then go well what's it doing? Why it's doing it and then you try to either catch up or operate because it doesn't get taught very well'. Additionally, participants commented that the effectiveness of verbal audio might not be sufficient in the congested and noisy environment, with audio tones and chimes proving a more significant attention-getting cue, without requiring significant attentional resources to assimilate the information. This comment related to the expectation vs reality of the system, which used audio for non-critical tasks that created confusion about the urgency of the tasks and prompted attention at the expense of other actions. Future studies should investigate the difference between audio usage on the flight deck and the effectiveness of pilot monitoring of autonomous actions during the scenarios conducted.

3.2.5 Operations and procedures

Participant three commented on the lack of operational procedures to support the scenario of the study, which increased the resistance of the pilot to the autonomous actions of checklist items. This resistance was generally associated with the authority and responsibility of both system and human in the safety of the flight, with reflection to the safety of the aircraft itself or the passengers: 'So at the moment you've got me deep into a procedure that I might not have started just yet because I wanted to do other things because you've not been there. You then have to work out what kind of happened'. Furthermore, the delegation of tasks between the system and pilot needs to be refined, as designed conditions had participants between two philosophies of control and response, which created further confusion on their ultimate role on the flight deck: 'You're between two checklist philosophies. Then I think, you're going to manually check off when you've done something, but because you've gone to the fully automated, fully monitored route, I think you're better off kind of staying there'. These comments related to the accountability of the system in operation, where no current operational procedures permit the use of autonomous tasking without pilot intervention. The suggested changes would separate the duties of both system and pilot, tailored to the strengths of each: (1) system for rapid completion and task accuracy, and (2) pilot for critical decision making and problem solving in real-time.

Participant four made a further comment that the scenario of return to cockpit (RnTC) would require specialised training and operational guidelines for pilots to rapidly build SA on the systems actions that have taken place in their absence and the rationale for these actions: 'If it were possible to have nobody in the cockpit, you could train pilots that the first thing they do when they get back into the cockpit is check the primary flight display to see if there's been any automatic action'. This comment related to the expectation vs reality of the system, which demonstrated the conflict between system transparency

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

and pilot SA building on concurrent tasks associated with an engine shutdown checklist. Further studies into the authority of an 'intelligent' system are required, focused on the pilot incapacitation event and the level of assistance to the pilot and overall flight as a result of this event.

4.0 Discussion

4.1 Analysis and findings from participant interviews

In the study, the effectiveness of an 'intelligent' system to perform the tasks of the PM when incapacitated was evaluated, with interview sessions with pilots to identify the critical task-related information that is required by the system to monitor and respond, if necessary. Although the integration of autonomous actions within the pilots' workload is seen to positively improve safety and operational capacity [15, 38], the use of autonomous actions is hindered by factors such as: natural mistrust in AI/autonomy, mode confusion and saturation of information during emergency events [25, 39, 40]. The study discovered that pilots' perception towards these 'intelligent' systems should further integrate the pilots in a decision-making and cross-checking capacity, to satisfy the delegation of safety-critical tasks to a non-human pilot. This conclusion is supported by research from O'Neill et al. (2022), who investigated the literature available on human-autonomy teaming and pilot requirements to form an effective team, although further studies into the delegation of safety-critical tasks need to be conducted [30]. Overall authority of 'intelligent' systems needs to be redesigned and examined, to complement current operational practices of cross-checking and diagnostic tasking before conducting safety-critical tasks, or to completely redesign and propose new operational practices that would favour the increased autonomous authority on the flight deck, with pilots in a supporting role [6, 39].

The distribution of codes into the themes for the current study's area of focus (Figs 5 and 6) can be seen in Figs 7 and 8, which reflect the individual feedback from each pilot for analysis and reflection towards the necessary changes to the system. Analysis identified that participants had concerns with their ability to switch between active and passive monitoring of the system, with a loss of SA of the system actions and pilot actions during the scenarios. The loss of SA to the system was combined with an overload of information that was not considered relevant in that situation, which delayed pilot decision-making and created further mistrust in the system to perform tasks without active pilot engagement. This conflict between the system authority and time-driven actions created an unsuitable environment for the pilots to apply their overall command authority on the flight deck, with any external intervention from the pilots causing the system to switch off completely [6]. The time-driven tasks were provided by the pilots during workshop events earlier in the systems' development; however, when applied to the scenario, they were insufficient to invite pilot intervention. The preference among pilots was the overriding authority of the human to disregard the checklist and exert their authority based on their judgement and requirements from the aircraft, such as to climb away from terrain with low energy on a single engine [1, 25]. The conflict between system-pilot authority during routine and emergency tasking needs to be defined into respective responsibilities on the flight deck and criticality of actions to flight safety and/or flight performance/health [2, 6].

The modality for attention-getting cues was controlled across the scenarios of the study, configured either independently or combined with other modalities to understand the effectiveness of pilot SA building when out of the loop and having to quickly monitor and respond to system actions [33, 40, 41]. It was reported that attention-getting cues did not promote active engagement with the pilots to the autonomous actions; instead, they promoted monitoring tasks that hindered other critical actions, such as descent management and ATC communication. The lack of active engagement increased the mistrust in the system, questioning the transparency of the system to invite the pilot to decision-making activities related to handling the engine shutdown procedure [2, 25]. The future design of 'intelligent' systems should promote more active engagement from the pilots, in the form of confirmation or safety-critical switch movement (one movement by the system and subsequent from pilot to activate/deactivate), to support continued awareness in the system while opening cognitive capacity for other duties on the flight deck, without affecting their respective authority and monitoring duties [5, 27].

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

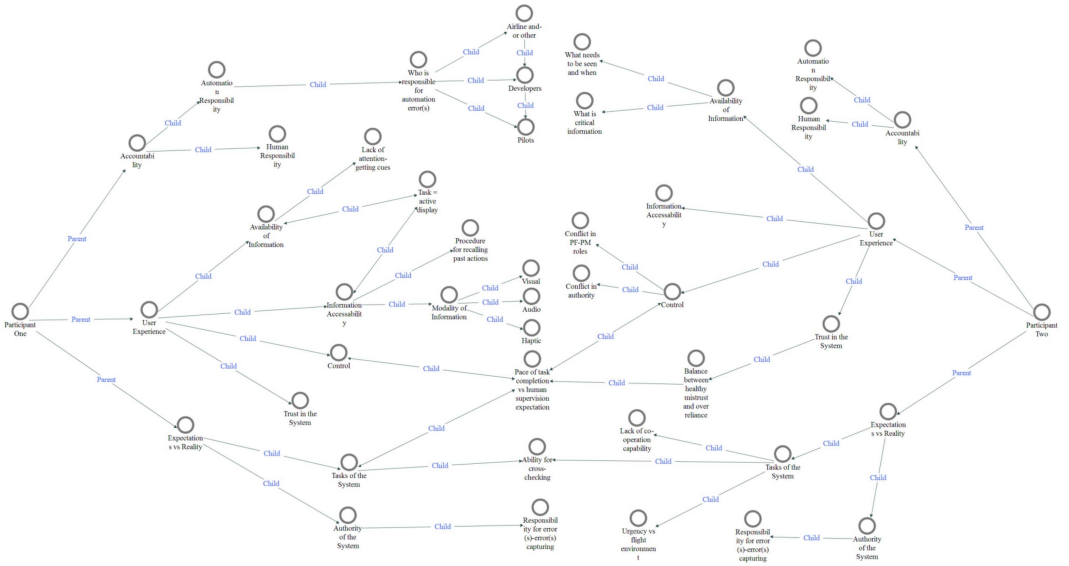


Figure 7. *Interaction map of participant one and participant two interview feedback, based on first-layer and second-layer coding. Parent codes indicate the three main concepts of the study with child codes a more detailed breakdown of these concepts related to interview feedback. Codes are connected either with one-way or two-way connections, demonstrating the connections and interactions between codes.*

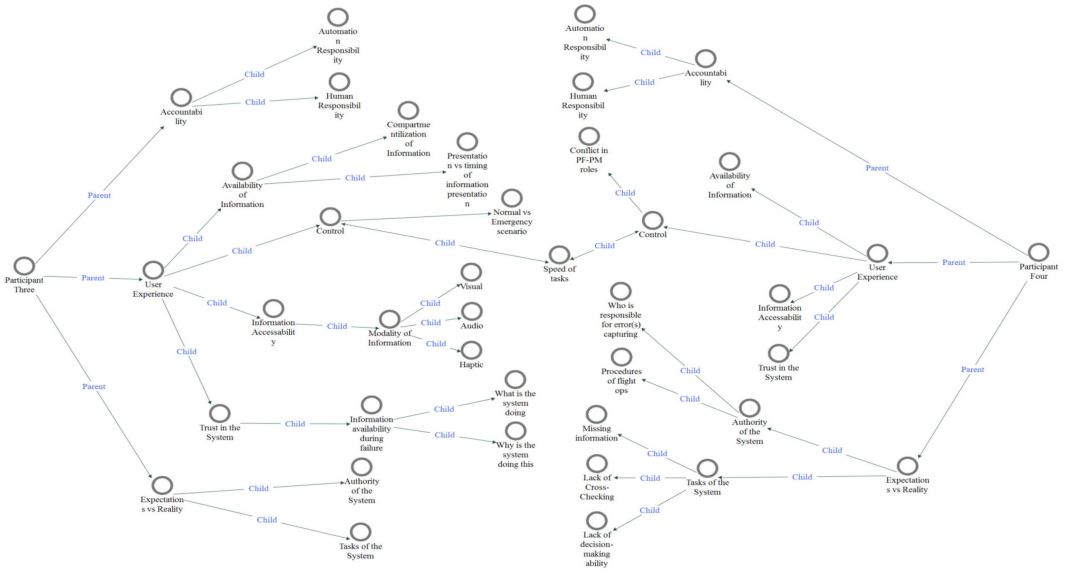


Figure 8. *Interaction map of participants three and four interview feedback, based on first-layer and second-layer coding. Parent codes indicate the three main concepts of the study with child codes a more detailed breakdown of these concepts related to interview feedback. Codes are connected either with one-way or two-way connections, demonstrating the connections and interactions between codes.*

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

Overall satisfaction of participants with the modality on the flight deck revealed a common interest in removing nonessential audio utilisation to the flight deck, with primary favour towards audio chimes and tones in respect to verbalisation. While an expected outcome of the study, the effectiveness of pilot attention-getting with audio verbalisation proved to support greater SA overall of the autonomous system, as demonstrated by interviews with the participants. This conflict represents the continual challenge of redesigning the flight deck to remove the workload of a secondary pilot in response to pilot incapacitation. It also highlights the need to promote a manageable workload for the remaining pilot by designing a system that can suitably operate with a single pilot [16, 30]. Despite this, there was significant interest in the prospect of autonomous actions to supplement the role of the pilot in both normal and emergency scenarios, although further work is required to identify the human-centric design (HCD) concepts required to support this integration into modern flight deck design. Primarily, there needs to be further studies into human-system collaboration and task-sharing to identify the cognitive capabilities of humans to share workload in the dynamic scenario, without compromising the overall SA of their core responsibilities.

This study successfully identified user satisfaction and the suitability of information modality for flight deck design under the exceptional circumstances of pilot incapacitation. While the scenario setup and designed system were relatively low fidelity, the study identified several HAT principles for future development of autonomous tasking on the flight deck. This includes the redesign of audio prompting to enhance pilot awareness and SA building during this scenario. Additionally, the study established the human-centric requirements for future HMI development to enable pilot supervision of autonomous tasking during these exceptional circumstances. Supporting this is the redesign of colour-coded enunciation to facilitate cross-checking capability on the flight deck.

4.2 Principles of HAT for 'intelligent' systems to the flight deck

During post-scenario interview analysis, it was discovered that the user-centred design process used for the 'intelligent' system development was not sufficient to support cognition in pilots, primarily associated with reduced capacity to monitor the system with designed time intervals for autonomous actioning [2, 42]. When developing this system, time intervals were designed to allow sufficient monitoring time reflective of two-pilot checklist tasking, but to enable the 'intelligent' system to secure the engine in a time-efficient manner to save its functionality. However, each participant expressed their concern that these time intervals would reduce pilot authority on the flight deck, as they would be unable to sufficiently process the intended actions and consequences before actions were performed, despite the attempts of the study to utilise sensory modality to reinforce the information of the system to the pilots.

Several key principles were taken from the study, with pilots commenting on their preferences and ideas for future design of these systems to integrate into their workflow on the flight deck. These were:

- When developing audio interaction between 'intelligent' systems and pilots, task-critical information should be prioritised. Unnecessary audio congestion can overload pilot cognitive capacity, with many pilots filtering out this information [14]. Audio narration should prioritise high-level information and be dynamic to the scenario and request of the pilot, ideally with an activation switch and preference of tone/pitch.
- When developing 'intelligent' systems on the flight deck, consideration of the overall authority and operational procedures to utilise the system should be given before testing. The exceptional circumstance of a pilot outside the cockpit when the event occurs will necessitate a higher authority system than if the pilot is on the flight deck during the event. The system should be dynamic and adaptive to the scenario, pilot input, and urgency of the failure event [1, 6]. This interaction is not understood completely in this study, although clear indications of future direction have been expressed.
- The choice of modality for an 'intelligent' system's transparency to pilots is dependent on the intended scenario, with no 'one size fits all'. Sensory modality should be combined and presented

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

to pilots for reception, as opposed to independent, verbose modalities that overwhelm pilot cognition [17]. While not identified in this study, a suitable combination of modality appears to combine visual and physical features over auditory narration, unless priority is given to a pilot for activation and is halted in favour of audio alerts/warnings on the flight deck, or for ATC communications.

4.3 Limitations of the study

Some of the limitations of the study became apparent during the scenarios with test pilots, where trials fell short of capturing the urgency and pacing of tasks as reflected in the autonomous checklist. Due to this, the pilots had significant comments on the system's initial delay time, as this creates an unnecessary monitoring task to identify when the 'intelligent' system would action the checklist and creates a delay in subsequent actions, such as diagnostic and communication tasks. Furthermore, pilots made significant comments to the initial conditions of the scenario, which had the flight in a non-maintainable altitude for single-engine operations, with no opportunity to manage and control the descent to a safe altitude, before performing the checklist and securing the engine, which was considered a secondary task, not primary. Again, the pacing of the system meant pilots felt discouraged on the utilisation of the autonomous system during this engine shutdown procedure, instead favouring manual handling and confirmation. Additionally, the status of the system did not provide satisfactory confirmation of autonomous actions and the consequences of these actions, leaving the pilots to pick up a 'half-completed checklist' and creating some dissociation from the aircraft and the flight deck. Finally, the relatively small sample size restricted the statistical significance of the results collected and was restricted to fixed-wing aircraft. The small sample size was due to a limited availability of pilots for the experiment; however, the study was focused on a proof-of-concept for an autonomous system to support pilot workload. The outcomes of the study can only be considered within the specific scenarios as tested, but have scope to support future development into autonomous systems for the flight deck, along with setting the principles of HAT on the flight deck.

Moving forward, improvements to the use of warning/alerting on the flight deck can help to capture pilot attention to autonomous actions, allowing time to derive and action the checklist when able. Further, proposing autonomous actions as a recommender system might improve the relationship between the human and system, giving the authority to the pilot to configure and activate the system, allowing for a shorter delay in tasking without compromising pilot awareness of actions. Consideration towards RnTC practices will need to consider the authority of 'intelligent' systems when the pilot is out-of-seat, and what is considered as flight critical, as opposed to trivial alerting that can wait for pilot oversight. Finally, the usability of audio verbalisation needs to be calibrated to be reflective of the operating environment, primarily the communication between ATC and the aircraft. Audio assistants that share the channel in pilot headsets could minimise the disruptive effect of verbal communication in an already congested environment, promoting an intuitive method of SA building on 'intelligent' systems even during high workload scenarios [16, 42].

5.0 Conclusion

The analysis of interview feedback found several key areas of improvement to the designed 'intelligent' system and autonomous checklist display, with further comments regarding the use of audio assistants on the flight deck. The issues are related to the human-automation interaction and the effectiveness of attention-getting cues to sustain monitoring performance in the event of a pilot incapacitation event or a failure event. The authority and distribution of tasking was considered an urgent area of improvement for future trials, where participants felt themselves dissociation and 'rushed' by the autonomous actions during the engine shutdown checklist. The delay in the autonomous actions was not sufficient for pilot supervision and prevention, as no time was permitted for scanning and diagnostic tasking of the causal factor of the engine's high oil temperature, or the operational constraints faced when the engine

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

is shut down. Participants suggested that a significantly longer delay between system actions would permit the cross-checking and monitoring roles of the pilots, ensuring the correct actions are performed. Alternatively, participants suggested an activation button for autonomous actions, which would allow the pilots to manually engage the system at their convenience, allowing time for diagnostic tasking and ATC communications to permit a descent associated with the loss of an engine in cruise.

The study developed an 'intelligent' system that could perform checklist tasks during an engine high oil temperature event. The study aimed to identify the informational requirements and modality for communication between the system and pilots on the flight deck. Four professional pilots took part in several scenarios in solo pilot configuration. Communication modality was varied across scenarios to determine the system requirements to enable pilot satisfaction when monitoring autonomous tasking during the scenarios. Based on the results and data analysis conducted, several key principles for 'intelligent' system design were generated. Future scope will further investigate the relationship between pilots and these 'intelligent' systems, to derive the responsibilities and role of the autonomous actions within the flight deck. Additionally, the authority and capability of autonomous action need to be carefully investigated to identify the requirements of a system to handle situations where pilot incapacitation would create an unacceptable workload increase for the remaining pilot. Further integration of 'intelligent' system authority and tasking on the flight deck needs to meet operational requirements that satisfy the pilot in their duties and promote a higher standard of safety for such exceptional circumstances. The work conducted can also be used to extend the discussion of future extended minimum crew operations (eMCO) operations, with greater support to the pilot to reduce workload and capture error(s), while maintaining the crucial human-in-the-loop design for 'intelligent' system collaboration with human operators.

Acknowledgments. The authors would like to thank Rolls-Royce for allowing this research study to be conducted on the Future Systems Simulator. The authors would also like to express gratitude to the pilots who took part in the experimentation and interview sessions, providing valuable insight into the area of study. This study was funded as part of the Powerplant Integration of Novel Engine System (PINES) project with Rolls-Royce.

References

[1] Kirwan, B. Human factors requirements for human-AI teaming in aviation, *Future Transp.*, 2025, **5**, (2), p 42. doi:10.3390/futuretransp5020042

[2] Luo, R., Du, N. and Yang, X.J. Evaluating effects of enhanced autonomy transparency on trust, dependence, and human-autonomy team performance over time, *Int. J. Hum.-Comput. Interact.*, 2022, **38**, (18–20), pp 1962–1971. doi:10.1080/10447318.2022.2097602.

[3] Kirwan, B., Charles, R., Jones, K., Li, W.-C., Page, J., Tutton, W. and Bettignies-Thiebaux, B. The human dimension in tomorrow's aviation system, CIEHF white paper, 2020. https://ergonomics.org.uk/resource/tomorrows-aviation-system.html

[4] Harris, D. Single-pilot airline operations: designing the aircraft may be the easy part, *Aeronaut. J.*, 2023, **127**, (1313), pp 1171–1191. doi:10.1017/aer.2022.110

[5] Brand, Y. and Schulte, A. Workload-adaptive and task-specific support for cockpit crews: design and evaluation of an adaptive associate system, *Hum.-Int. Sys. Integr.*, 2021, **3**, (2), pp 187–199. doi:10.1007/s42454-020-00018-8

[6] Schelble, B.G., Flathmann, C. and Mcneese, N. Towards meaningfully integrating human-autonomy teaming in applied settings, *HAI 2020 - Proceedings of the 8th International Conference on Human-Agent Interaction*, 2020, pp 149–156. doi:10.1145/3406499.3415077

[7] Adler, E.J. and Martins, J.R.R.A. Hydrogen-powered aircraft: fundamental concepts, key technologies, and environmental impacts, *Progress in Aerospace Sciences*, 2023, **141**, pp 1–30. doi:10.1016/j.paerosci.2023.100922

[8] Civil Aviation Authority. Make Monitoring Matter – Why pilot monitoring plays a key role in preventing and recovering from loss of control incidents, Monitoring Matters, UK Civil Aviation Authority, London, 2024. https://www.caa.co.uk/safety-initiatives-and-resources/safety-projects/monitoring-matters/make-monitoring-matter/

[9] Gil, M., Albert, M., Fons, J. and Pelechano, V. Designing human-in-the-loop autonomous cyber-physical systems, *Int. J. Hum. Comput. Stud.*, 2019, **130**, pp 21–39. doi:10.1016/j.ijhcs.2019.04.006

[10] Schmid, D., Vollrath, M. and Stanton, N.A. The system theoretic accident modeling and process (STAMP) of medical pilot knock-out events: pilot incapacitation and homicide-suicide, *Saf. Sci.*, 2018, **110**, pp 58–71. doi:10.1016/j.ssci.2018.07.015

[11] Wang, G, Li, M., Wang, M. and Ding, D.A. Systematic literature review of human-centered design approach in single pilot operations, *Chin. J. Aeronaut.*, 2023, **36**, (11), pp 1–23. https://www.sciencedirect.com/science/article/pii/S100093612300256X

[12] Ersin, K., Gundogdu, O., Kaya, S.N., Aykiri, D. and Serbetcioglu, M.B. Investigation of the effects of auditory and visual stimuli on attention, *Heliyon*, 2021, **7**, (7), p e07567. doi:10.1016/j.heliyon.2021.e07567
[13] Stewart, N. and Harris, D. Passenger attitudes to flying on a single-pilot commercial aircraft, *Aviat. Psychol. Appl. Hum. Fact.*, 2019, **9**, (2), pp 77–85. doi:10.1027/2192-0923/a000164
[14] Schlager, S.M., Abballe, A., Kinnison, M., Colter, J., Bryan, S. and Harbour, S.D. Coherent copilot: human versus synthetic voice for air combat performance, *2022 IEEE/AIAA 41st Digital Avionics Systems Conference (DASC)*, 2022, pp 1–6, doi:10.1109/DASC55683.2022.9925873. https://ieeexplore.ieee.org/document/9925873
[15] Parnell, K.J., Wynne, R.A., Plant, K.L. and Stanton, N.A. User-centred design and evaluation of future flight deck technologies, *Ergon. Hum. Fact.*, 2021a, **1**, pp 1–8. https://publications.ergonomics.org.uk/uploads/01_23.-edited.pdf
[16] Wang, Y., Pang, Y., Gorceski, S., Kostiuk, P., Mohen, M.T., Menon, P.K. and Liu, Y. A voice communication-augmented simulation framework for aircraft trajectory simulation, *IEEE Trans. Intell. Transp. Syst.*, 2021, **23**, (7), pp 7310–7320. https://ieeexplore.ieee.org/document/9390387
[17] Blundell, J., Collins, C., Sears, R., Plioutsias, T., Huddlestone, J., Harris, D., Harrison, J., Kershaw, A., Harrison, P. and Lamb, P. Low-visibility commercial ground operations: an objective and subjective evaluation of a multimodal display, *Aeronaut. J.*, 2023, **127**, (1310), pp 581–603. doi:10.1017/aer.2022.81
[18] IATA. Global Outlook for Air Transport, Highly Resilient: less Robust, 2023. https://www.iata.org/en/iata-repository/publications/economic-reports/global-outlook-for-air-transport—-june-2023/
[19] Faulhaber, A.K., Friedrich, M. and Kapol, T. Absence of pilot monitoring affects scanning behavior of pilot flying: implications for the design of single-pilot cockpits, *Hum. Factors*, 2020, **64**, (2), pp 278–290. doi:10.1177/0018720820939691
[20] Parnell, K.J., Wynne, R.A., Griffin, T.G.C., Plant, K.L. and Stanton, N.A. Generating design requirements for flight deck applications: applying the perceptual cycle model to engine failures on take-off, *Int. J. Hum.-Comput. Interact.*, 2021b, **37**, (7), pp 611–629. doi:10.1080/10447318.2021.1890488
[21] Chen, J.Y.C., Lakhmani, S.G., Stowers, K., Selkowitz, A.R., Wright, J.L. and Barnes, M. Situation awareness-based agent transparency and human-autonomy teaming effectiveness, *Theor. Issues Ergon. Sci.*, 2018, **19**, (3), pp 259–282. doi:10.1080/1463922X.2017.1315750
[22] Li, W-C., Korek, W.T., Liang, Y.H. and Lin, J.J.H. Touchscreen controls for future flight deck design: investigating visual parameters on human-computer interactions between pilot flying and pilot monitoring, *J. Aeronaut., Astronaut. Aviat.*, 2023, **55**, (2), pp 201–212. https://www.airitilibrary.com/Article/Detail/P20140627004-202212130001-00008
[23] Wang, Y., Li, W.-C., Korek, W.T. and Braithwaite, G. Future flight deck design: Developing an innovative touchscreen inceptor combined with the primary flight display, *Int. J. Ind. Ergon.*, 2024, **101**, p 103588. doi:10.1016/j.ergon.2024.103588
[24] Wickens, C.D., Gutzwiller, R.S. and McCarley, J.S. *Applied Attention Theory* (2nd ed.), CRC Press, Boca Raton, 2022. doi:10.1201/9781003081579.
[25] Air Line Pilot Association The Dangers of Single-Pilot Operations. ALPA White Paper, 2019. https://www.alpa.org/-/media/ALPA/Files/pdfs/news-events/white-papers/white-paper-single-pilot-operations.pdf?la=en
[26] Civil Aviation Authority CAP 737: Flight Crew Human Factors Handbook. UK Civil Aviation Authority, London, 2023. https://www.caa.co.uk/our-work/publications/documents/content/cap-737/
[27] Lyons, J.B., Sycara, K., Lewis, M. and Capiola, A. Human–autonomy teaming: definitions, debates, and directions, *Front Psychol* 2021, **12**, p 589585. doi:10.3389/fpsyg.2021.589585
[28] Etherington, T., Kramer, L., Bailey, R., Kennedy, K. and Stephens, C. Quantifying pilot contribution to flight safety for normal and non-normal airline operations, *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*, 2016, pp 1–14. https://ieeexplore.ieee.org/document/7778094
[29] Weiss, H., Liu, A., Byon, A., Blossom, J. and Stirling, L. Comparison of display modality and human-in-the-loop presence for on-orbit inspection of spacecraft, *Hum. Factors*, 2023, **65**, (6), pp 1059–1073. doi:10.1177/00187208211042782
[30] O'Neill, T., McNeese, N., Barron, A. and Schelble, B. Human–autonomy teaming: a review and analysis of the empirical literature, *Hum. Factors*, 2022, **64**, (5), pp 904–938. doi:10.1177/0018720820960865
[31] Schmid, D. and Stanton, N.A. A future airliner's reduced-crew: modelling pilot incapacitation and homicide-suicide with systems theory, *Hum.-Int. Sys. Integr.*, 2019, **1**, pp 27–42. doi:10.1007/s42454-019-00001-y
[32] Lim, Y., Gardi, A., Sabatini, R., Ramasamy, S., Kistan, T., Ezer, N., Vince, J. and Bolia, R. (2018). Avionics human-machine interfaces and interactions for manned and unmanned aircraft, *Prog. Aerosp. Sci.* 2018, **102**, pp 1–46. doi:10.1016/j.paerosci.2018.05.002
[33] Markovich, T., Honig, S. and Oron-Gilad, T. Closing the feedback loop: the relationship between input and output modalities in human-robot interactions, *Springer Proceedings in Advanced Robotics*, 2020, 12. doi:10.1007/978-3-030-42026-0_3
[34] Korek, W.T., Beecroft, P., Lone, M., Bragado Aldana, E., Mendez, A., Enconniere, J., Asad, H.U., Grzedzinski, K., Milidere, M., Whidborne, J., Li, W.-C., Lu, L., Alam, M., Asmayawati, S., del Barrio Conde, L., Hargreaves, D. and Jenkins, D. Simulation framework and development of the future systems simulator, *Aeronaut. J.*, 2024, **128**, (1330), pp 2754–2780. doi:10.1017/aer.2024.91
[35] Naifeh, N. and Kaufman, E.J. Color Vision. StatPearls [Internet], Treasure Island (FL), 2025. https://pubmed.ncbi.nlm.nih.gov/29261952/
[36] Carroll, J. and Conway, B.R. Color vision, in Barton, J.J.S and Leff, A. (Eds.), *Handbook of Clinical Neurology*. Vol. 178, Elsevier, Amsterdam, 2021, pp. 131–153. doi:10.1016/B978-0-12-821377-3.00005-2
[37] El Jouhri, A., Sharkawy, A., Paksoy, H., Youssif, O., He, X., Kim, S. and Happee, R. The influence of a color themed HMI on trust and take-over performance in automated vehicles, *Front Psychol*, 2023, **14**, pp 1128285. doi:10.3389/fpsyg.2023.1128285

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press

[38] Stanton, N.A., Harris, D. and Starr, A. The future flight deck: Modelling dual, single and distributed crewing options, *Appl. Ergon.*, 2016, **53**, pp 331–342. doi:10.1016/j.apergo.2015.06.019
[39] Xu, W. From automation to autonomy and autonomous vehicles: challenges and opportunities for human-computer interaction, *Interactions*, 2020, **28**, (1), pp 48–53. doi:10.1145/3434580
[40] Dziuban, C., Graham, C.R., Moskal, P.D., Norberg, A. and Sicilia, N. Blended learning: the new normal and emerging technologies, *Int. J. Educ. Technol. High. Educ.*, 2018, **15**, (1). doi:10.1186/s41239-017-0087-5
[41] Ruskin, K.J., Corvin, C., Rice, S., Richards, G., Winter, S.R. and Clebone Ruskin, A. Alarms, alerts, and warnings in air traffic control: an analysis of reports from the aviation safety reporting system, *Transp. Res. Interdiscip. Perspect.*, 2021, **12**. doi:10.1016/j.trip.2021.100502
[42] Seedhouse, E., Brickhouse, A., Szathmary, K. and Williams, E.D. Displays, in *Human Factors in Air Transport*, Springer International Publishing, Switzerland, 2020, pp 149–164. doi:10.1007/978-3-030-13848-6_9

Cite this article: Saunders D., Blundell J., Li W., Beecroft P., Lu L., Korek W.T. and Shi W. Principles for 'intelligent assistant' systems in future flight deck design: autonomous action integration to reduce pilot workload. *The Aeronautical Journal*, https://doi.org/10.1017/aer.2026.10161

https://doi.org/10.1017/aer.2026.10161 Published online by Cambridge University Press



https://dspace.org/

Principles for ‘intelligent assistant’ systems in future flight deck design: autonomous action integration to reduce pilot workload

Saunders, Declan; Blundell, James; Li, Wen-Chin; Beecroft, Peter; Lu, Linghai ...

2026-12-31
Cambridge University Press (CUP)
https://doi.org/10.1017/aer.2026.10161
Saunders D, Blundell J, Li W-C, et al., (2026) Principles for ‘intelligent assistant’ systems in future flight deck design: autonomous action integration to reduce pilot workload. The Aeronautical Journal, Available online 10 April 2026

Downloaded from a DSpace repository. https://dspace.org/