

RESEARCH ARTICLE



# Comparative analysis of touchscreen inceptors and traditional sidesticks on flight decks: flight performance, visual behaviours and situation awareness

Yifan Wang<sup>a</sup>, Wojciech Tomasz Korek<sup>b</sup>, James Blundell<sup>a</sup> and Wen-Chin Li<sup>a</sup> 

<sup>a</sup>Safety and Accident Investigation Centre, Cranfield University, Bedfordshire, UK; <sup>b</sup>Dynamics, Simulator and Control Group, Cranfield University, Cranfield, Bedfordshire, U.K

## ABSTRACT

The concept of touchscreen primary control device is a novel approach of touchscreen implementation. The objective of this study is to investigate differences in flight performance and attention allocation between a touchscreen inceptor and a traditional sidestick. Twenty-one participants flew four simulated instrument landing system (ILS) approaches – with the touchscreen inceptor or traditional sidestick – during flight scenarios where an aircraft attitude disturbance was either present or absent. Results demonstrated that participant performance scores were worse with the touchscreen inceptor compared to the sidestick during attitude disturbance scenarios. Interestingly, participants exhibited reduced attention to external visual cues with the touchscreen inceptor compared to the sidestick. In addition, use of the touchscreen inceptor resulted in lower performance and lower self-reported situation awareness. Overall, the touchscreen inceptor demonstrated poorer performance compared to the traditional sidestick, highlighting limitations in its current design that warrant cautious consideration and further investigation.

**PRACTITIONER SUMMARY:** Integrating the touchscreen inceptor into the PFD is an attempt to expand the application of the touchscreen in the flight deck and to explore the human-system integration mode of the future flight deck. Although the touchscreen inceptor cannot match the sidestick's performance, it has a potential impact on the pilot's situation awareness supply and attention allocation.

## ARTICLE HISTORY

Received 5 July 2024  
Accepted 22 January 2025

## KEYWORDS

Flight deck design; human-computer interactions; situation awareness; touchscreen inceptor; visual behaviours

## 1. Introduction

As the complexity of the flight deck increases, the traditional array of knobs, buttons and gauges is no longer suitable and has been transformed into the 'glass cockpit', an array of electronic (digital) flight instrument displays, typically large LCD screens (Nagabhushana and Sathyanarayana 2010). Due to the increasing intelligence of modern aircraft, the amount of information processed by flight computers has increased dramatically. For example, flight computers now handle vast amounts of sensor data from advanced avionics systems, including real-time weather updates, terrain mapping, and traffic collision avoidance systems (Singh, Jain, and Mishra 2023). Additionally, they manage complex autopilot functions, fly-by-wire controls, and integrate inputs from multiple systems to optimise flight efficiency and safety. Thus, the increasing quantity and complexity of

flight data place higher demands on the design of flight deck interfaces to improve human-computer interactions, enhancing pilots' ability to efficiently perceive, comprehend system state information, and act accordingly—an area where current flight deck interfaces are insufficient (Xie, Zhou, and Qu 2023). The application of touchscreens on flight decks is a potential solution to improving this integration and warrants future flight deck research.

Currently, the application of touchscreens in safety critical human-machine interaction systems, such as flight decks, is becoming increasingly widespread (Chu, Wu, and Xu 2021; Xue, Zhang, and Zhang 2022). Touchscreens themselves offer numerous advantages in systems that demand high human-machine interaction efficiency by providing system output and immediate user feedback within the same area where user input or selection occurs. Consequentially, touchscreen

**CONTACT** Wen-Chin Li  [wenchin.li@cranfield.ac.uk](mailto:wenchin.li@cranfield.ac.uk)  Safety and Accident Investigation Centre, Cranfield University, Martell House, Cranfield, Bedfordshire MK43 0TR, UK

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

interfaces can make interactions more intuitive and less mentally demanding – such as reducing the need for users to recall items (Tang, Weiss, and Stirling 2023; Xie, Zhou, and Qu 2023). This immediate feedback loop helps users stay focused on their current task without the distraction of having to remember previous actions or results. It enhances the overall user experience by making the interface easier to use and reducing cognitive load (Tang, Weiss, and Stirling 2023). However, most existing studies have focused on touchscreens as advanced input devices, primarily replacing traditional keyboards and buttons on the flight deck. Applications intended to support navigation or checklist management tasks have demonstrated the integration and efficiency benefits of touchscreens (Sadia et al. 2022; Watkins et al. 2018). Despite their proven benefits in supplementary tasks, the potential for touchscreens to serve as primary flight control devices—requiring real-time manipulation of aircraft flight dynamics—remains underexplored. This represents a key gap in current research that requires investigation into how the use of a touchscreen inceptor, compared to a sidestick inceptor, influences pilot manual flying performance, attention distribution, situational awareness, and mental workload. Our study addresses this gap by leveraging eye-tracking data to explore the differences in pilot behaviour and attention allocation between touchscreen and sidestick inceptors. Addressing this gap, our study aims to evaluate the effectiveness of touchscreens as primary flight control mechanisms in a simulated flight environment.

The implementation of glass cockpits, including touchscreens, also provides manufacturers the advantage of easily updating and modifying user applications running on such systems (Bhalla and Bhalla 2010). Screens can present critical information and actionable options in a clear and intuitive manner, informing users of the functions that can be performed. They can also alert users to potential issues in real-time, enabling quicker responses and proactive problem-solving. This significantly enhances overall operational efficiency and holds great potential to reduce workloads, decrease head-down scanning time, and improve safety (Noah, Li, and Rothrock 2017; Rouwhorst et al. 2017). However, the flight deck environment is complex, subject to factors like vibration, and demands high accuracy and response speed in operations. Due to the lack of physical feedback, touchscreens may impact users' perception of their own operation responses, leading to an over allocation of user attention and to the detriment of other task-relevant information (Wynne et al. 2021). Additionally, the ergonomic shortcomings of touchscreens may lead to issues such as increased physical

fatigue and accidental touches (Khoshnewiszadeh and Pool 2021). Because of these inherent challenges, touchscreens have been limited to supplementary roles in commercial aviation and their application as direct control mechanisms remains an open question.

### ***1.1. Touchscreen integration in flight decks: development, challenges, and applications***

Although previous studies have not specifically focused on using touchscreens as direct control input devices for aircraft, they provide valuable insights into the advantages and limitations of touchscreen use in the flight deck environment. These insights are crucial for understanding how the characteristics of touchscreens might impact their effectiveness as inceptors. To further understand the range touchscreen application scenarios on the flight deck, and enhance the development and the efficiency of human-system integration (HSI), numerous studies have explored the hardware and software requirements for touchscreens in specific environmental conditions - such as vibration resistance and high transmission speeds. For example, Tao et al. examined the effects of control-to-display gain in vibration environments and found control-to-display gain (GAIN) values around 1 supported maintenance of user accuracy in high vibration environments and decreased the precision requirement level (Tao et al. 2021). Cockburn et al. (2019) designed the braced touch for touchscreens to improve resistance to the effects of vibration, such as in high turbulence environments, and where more precise manual inputs were needed (Cockburn et al. 2019). Using Fitts' law to optimise the interaction process of a touchscreen can help improve the efficiency and ease of use of touchscreen interfaces in the flight deck. While Fitts' law quantifies input performance and allows for comparisons across different interface designs, the resulting improvements have potential to enhance hand-eye coordination by making interactions more intuitive and reducing the effort required to accurately target controls (Coutts et al. 2019; van Zon et al. 2020). These findings suggest that the design and optimisation techniques developed for general touchscreen applications could be crucial when adapting touchscreens for use as direct control devices.

Furthermore, researchers have worked on the user interface of touchscreen to improve HSI. The target size and spacing had significant effects on task performance using touchscreens in turbulence environments in flight deck (Lin and Lee 2022). Furthermore,

touchscreen location, size and screen brightness can affect the HSI efficiency on the flight deck (Dodd et al. 2014, 2019; S. H. Kim and Kaber 2014). Providing additional visual and auditory feedback for touchscreens (i.e. to offset the lack of touchscreen tactile feedback) can also support task performance and reduce error rates (Lawrenson et al. 2018; Wan, Prinnet, and Sarter 2017). Compatibility between the task demands and required gestural input is also critical to effective touchscreen design; Tasks such as tapping and entering numbers are less affected by distractions compared to drag-and-drop or sequential tasks. Simple gestures like tapping or typing require less cognitive effort and fine motor control, making them more resilient to environmental distractions (Alapetite et al. 2018; Coutts et al. 2019). These interface design principles, developed for conventional touchscreen tasks, are relevant when considering how touchscreens can be adapted as inceptors, particularly in terms of minimising errors and optimising user interaction under varying environmental conditions.

For flight deck design, touchscreens have been widely tried and tested for navigation tasks, replacing traditional checklists and traditional system control panels (Sadia et al. 2022). In the Gulfstream G500 and G600 symmetry™ flight decks, touchscreens have replaced the entire overhead panel (Watkins et al. 2018). These studies have enabled designers to further develop touchscreen design for the flight deck, thus realising the vision of improving the overall efficiency and integration of the flight deck. However, the potential for touchscreens to be used in more dynamic and critical functions, such as direct flight control, remains less explored. Our research seeks to fill this gap by investigating how these well-documented touchscreen characteristics can be leveraged and adapted to enable their effective use as primary control devices in the flight deck.

### **1.2. Prospects for touchscreen inceptors: aircraft control system development and future**

Aircraft have always been manoeuvred using a sidestick or control column, and few researchers have attempted to manoeuvre an aircraft using a touchscreen inceptor (W. C. Li, Wang, et al. 2022). The introduction of sidestick controllers, which have revolutionised the conventional yoke-based controls, caused a substantial shift in aircraft control systems in the aviation sector in recent years. In contemporary fly-by-wire aircraft, the sidestick—a small joystick usually situated to the side of the pilot's seat—has become the principal control interface. The Airbus A320 (1987)

became the first airliner to feature global real-time steering and side-stick control, and the sidestick became one of the most important elements of aircraft manoeuvring in the Airbus family to date (Favre 1994). The sidestick provides significant advantages in terms of spatial flexibility compared to other aircraft control devices. Its design allows the pilot more freedom of movement by being positioned to the side, which reduces physical obstructions and frees up valuable space in front of the pilot. This setup also offers better elbow support, enhancing comfort and reducing fatigue during long flights. Furthermore, the sidestick's compact and unobtrusive design contributes to a more efficient cockpit layout, allowing for better access to other controls and instruments, thereby improving overall pilot manoeuvrability and operational efficiency. However, it is worth noting that sidesticks have certain limitations compared to yoke-based controls. For instance, Field and Harris (1998) highlighted that sidesticks reduce the amount of information available to pilots regarding their co-pilot's control inputs. Unlike yokes, which allow pilots to observe their co-pilot's inputs as peripheral visual cues, sidesticks lack this visual feedback, potentially impacting coordination in multi-crew operations.

Recent research on active sidesticks has further enhanced their performance, with variable force gradients providing greater flexibility in aircraft operation. The feeling of force variation allows the pilot to directly feel the limitations of the flight envelope (X. Li, Li, and Liu 2018; Van Baelen et al. 2021). Moreover, active sidestick systems (like those used in the Gulfstream G500) address the lack of peripheral visual cues present by providing direct force feedback of the co-pilot's control inputs. This effectively substitutes visual cues with haptic feedback, enhancing coordination and awareness between pilots during multi-crew operations (Watkins et al. 2018). However, there have been very few attempts to integrate flight control functions into other input tools beyond sidestick and yokes; Not only because such a change would encompass multiple technical and safety requirements, but also because physical control inceptors leverage the human neuromuscular system that enable the perception of subtle force feedback cues that are pivotal to precise fine motor movement (Patel, Atashzar, and Tavakoli 2022; Wu and Li 2024). While current touchscreen devices lack the capability to exploit the human neuromuscular system in the way traditional physical controls can, they offer a more integrated and compact cockpit layout, which could provide operational efficiencies and warrant further exploration as an innovative approach to flight path control.

Replacing physical control inceptors with touchscreen technology allows for more dynamic and context-sensitive displays. For example, touchscreens provide a versatile and customisable platform that allows pilots to seamlessly interact with a variety of controls and displays. The intuitive nature of touchscreen interfaces has the potential to simplify cockpit layouts, reduce the need for physical switches, and enhance the overall experience. While these advantages are well-documented for higher-level input functions, it remains unclear how they will impact the use of touchscreens as inceptors. Our research also seeks to explore how these characteristics might influence the effectiveness of touchscreens when used as inceptors (Tao et al. 2021). But the use of touchscreens in flight control systems also brings up problems with muscle memory and haptic feedback (Robinson et al. 2012). Physical feedback provided by traditional controls, such as resistance felt by joysticks or sidesticks, contributes to a pilot's situation awareness. Ensuring that touchscreen interfaces maintain a balance between technological innovation and the retention of basic haptic cues is critical to their successful integration into the aviation industry. To understand participants' initial impressions towards the touchscreen inceptor, our previous study used subjective questionnaires and deviations between flight paths and ideal paths. Previous research found that, even though touchscreen performance was lower than sidestick performance, younger people were more accepting of the touchscreen inceptor than experienced pilot (Wang et al. 2024). Our study aims to build on previous research by using more objective evaluation techniques to understand the gap between touchscreens and sidesticks. Firstly, we expect to analyse whether there is a difference in flying performance using different inceptors (sidestick vs. touchscreen) under different experimental conditions (disturbance vs. non-disturbance). Based on the reduced neuromuscular/haptic feedback of touchscreen displays, the first hypothesis of the current study is 'H1a: Participant performance scores will be significantly higher when using a sidestick compared to a touchscreen inceptor'. Furthermore, the current study examines flying performance under different 'turbulence' conditions. In these instances, differences in flying performance between input modes is expected to be amplified. Hence, a separate interaction hypothesis will be tested: 'H1b: There will be a significant interaction between input mode and the presence of disturbance on participants' performance, with performance expected to decline more significantly when using the touchscreen inceptor under turbulence compared to the sidestick'.

### 1.3. Evaluation of touchscreen inceptors

The introduction of a completely new type of aircraft controller requires a thorough evaluation from a human factors perspective to understand its impact on HSI. A comprehensive evaluation of a touchscreen inceptor can be based on objective and subjective measurement of pilot's performance, and cognitive and physiological states during a task. In HSI studies, researchers commonly utilise subjective scales to evaluate users' perceived ease of use, acceptance and cognitive states (e.g. situation awareness). Subjective scales are low cost and simple to use (Hamden et al. 2020). Situation awareness rating technique (SART) is one of the most used rating scales for assessing situation awareness in HSI research, which contains three dimensions: supply, demand and understanding (Taylor 2017). However, subjective scales, like the SART, possess drawbacks in the form of being susceptible to the subjective feelings of the participants. Resulting in measurements that can be overly influenced by high levels of both between- and within-person variability, making it difficult to find systematic experimental effects. Our study builds on previous research by using the SART to assess the impact of the touchscreen inceptors on situational awareness (SA) in a flight context. Specifically, we aim to analyse whether there are significant differences in SA across different input modes (sidestick vs. touchscreen) and operational conditions (disturbance vs. non-disturbance). A second set of hypotheses was created to test the SA effects of a touchscreen inceptor. Based on the previous research (Wang et al. 2024), it was expected that situation awareness would be lower with a touchscreen inceptor. Therefore, 'H2a: SART scores will be significantly higher when using the sidestick compared to the touchscreen inceptor'. Likewise to performance, SA was examined under different 'turbulence' conditions leading to the following hypothesis: 'H2b: There will be a significant interaction between input mode and the presence of disturbances on SART scores, with SART scores expected to decline more significantly when using the touchscreen inceptor under disturbance conditions compared to the sidestick'.

To address the limitations of subjective metrics, the use of non-invasive physiological sensors can enable the real-time, indirect measurement of cognitive and physiological states, reducing the influence of personal factors. Eye tracking is currently a highly effective non-invasive tool for obtaining cognitive/attentional indices through the capture of fixation count, fixation duration, saccades, pupil diameter and blink measurements (Krejtz et al. 2018). Fixation is typically associated with

information processing and depth of cognitive processing (Callaway, Rangel, and Griffiths 2021; Dehais et al. 2017). Changes in fixation location can indicate where attention is distributed, while the length of fixation durations to different locations indicate the relative importance of the information existing within the geographical space and the number of cognitive resources needed to process that information. Longer fixation duration and more fixation counts indicated higher visual attention was allocated in certain area (Hebbar, Pashikar, and Biswas 2022; Schwerd and Schulte 2022). Pupil diameter and blink rate have been found to correlate with workload levels and fatigue. High workload and fatigue have been shown to associate with pupil diameter increase in prior studies (Chen and Epps 2014; Duchowski et al. 2018; Zhou, Yang, and De Winter 2022). Blink count has been found in previous studies to show the opposite: it increases with increasing workload levels, yet decreases with increasing demands in tasks with high visual demands (Faure, Lobjois, and Benguigui 2016; Peißl, Wickens, and Baruah 2018).

Utilising eye tracking to evaluate interface design and to understand pilots' cognitive processes and attention distribution is an effective way to improve flight deck design efficiency (Blundell et al. 2023; Causse et al. 2019; W. C. Li et al. 2020). Combining different objective and subjective approaches for HSI analysis lowers errors due to the bias of a single assessment method, increases credibility, and allows for a more thorough investigation of the system's degree of perfection from many angles (Liu et al. 2021). In this study, we aim to use eye tracking to investigate how gaze behaviours differ between the touchscreen and sidestick inceptors under varying operational conditions. Specifically, we hypothesise that gaze patterns will reflect the distinct interaction demands of the two input methods. Therefore, we expect that 'H3a: Participants using the touchscreen inceptor will show significant differences in fixation count and duration on the Primary Flight Display (PFD) compared to those using the sidestick. Specifically, more fixations and shorter fixation durations are expected when using the touchscreen', whilst 'H3b: There will be a significant interaction between input mode and the presence of disturbances on gaze behaviours. Specifically, more fixations and shorter fixation durations are expected when using the touchscreen inceptor under disturbance conditions compared to the sidestick'.

For blink count, as a measure of mental workload, we hypothesise that 'H4a: Participants will exhibit higher blink frequencies when using the touchscreen inceptor compared to the sidestick' and 'H4b: There will be a significant interaction between input mode

and the presence of disturbances on blink count, with blink frequencies expected to be highest when using the touchscreen inceptor under disturbance conditions'.

## 2. Method

### 2.1. Participants

There were 21 participants, aged from 22 to 58 years old ( $M=31.67$ ,  $SD = 10.66$ ) with fix-wing flight hours from 0 to 300 ( $M=70.81$ ,  $SD = 92.96$ ) and simulator flight hours from 0 to 1000 ( $M=63.81$ ,  $SD = 216.01$ ). Participants included both pilots and non-pilots to account for the significant differences in familiarity with sidestick and touchscreen controls among participants. This approach was chosen to minimise the impact of varying levels of familiarity on the results and to reduce the potential bias that could arise from using only experienced pilots. To assess the potential influence of participants' prior flight experience on their performance, we conducted a simple t-test comparing the performance scores of participants with and without fixed-wing flight experience exceeding 100 hours (over 100h:  $N=7$ ; less 100h:  $N=14$ ). To check for the influence of participant experience, a 3-way ANOVA for performance was conducted between experience, input mode and scenario. The results revealed no significant interactions that included experience as a term ( $p>0.05$ ). Subsequently, all further analyses omitted experience as a factor. The research proposal was approved by the Cranfield research ethics committee before the experiment was conducted. All participants provided informed consent, which clarified that the objective of the experiment was to explore performance and visual behaviours when using a touchscreen inceptor. Participants were assured of their rights to withdraw themselves and their data from the trial at any stage. The data collected was stored in accordance with the British Psychological Society (BPS) Code of Ethics and Conduct and the Data Protection Act 2018. This ensured that all data handling procedures adhered to the ethical principles of respect, competence, responsibility, and integrity as outlined by the BPS. Additionally, participants were informed that their anonymised data was to be used for research purposes only.

### 2.2. Apparatus and material

#### 2.2.1. Future systems simulator (FSS):

This research was conducted using the Future Systems Simulator (FSS), developed by Cranfield University and Rolls-Royce (Korek et al. 2024). The FSS is a highly

adaptable, fixed-base flight simulator designed for rapid prototyping of current and future flight deck configurations. It features aircraft control and display systems that present information on up to four large, customisable touchscreens and two smaller side screens (Figure 1). The use of touchscreen controls instead of physical switches and levers allows for greater flexibility in testing experimental configurations. Although the FSS is not capable of producing physical vibrations, simulated vibrations were integrated into the study through software-generated signals.

### 2.2.2. Touchscreen inceptor

The touchscreen inceptor was integrated into the primary flight display (PFD) as a 'virtual stick' which



Figure 1. Future Systems Simulator layout.

enabled circular movements along vertical and horizontal axes within an operational red-line circle area (Figure 2). The interactive operational area's radius was 86 mm, with the virtual stick occupying a 17 mm diameter. The virtual stick area overlays the centre of the PFD and does not obscure important information such as altitude and speed. Participants could engage with virtual stick by touching and dragging their finger in any direction within the circular area. The virtual stick would follow the finger, which produced a linear control signal to the aircraft's model (ranging from  $-1$  to  $1$  in both vertical and horizontal axes). Tapping the area outside the virtual stick would not produce input. The stick had to be dragged from the centre first, simulating the sidestick's movement from the central position, and after releasing the finger, the stick would come back to the centre, simulating spring movement of the sidestick. Four transparent green areas, positioned within the outer axis of the operational red-line circle area, indicated the current movement status of the finger; When the finger moves in a certain direction, the corresponding green area becomes less transparent to provide feedback on the current level of input. Based on previous studies, experienced pilots tend to have biases against the design and input logic of touchscreen inceptors (Wang et al. 2024). To balance performance differences arising from these biases, we selected participants with varying experience levels for this experiment. The control logic of the touchscreen inceptor was designed to simplify the gesture-to-response correlation, promoting 'eyes out' flight. Commonly in touchscreen technology, upward motion correlates with upward movement, while

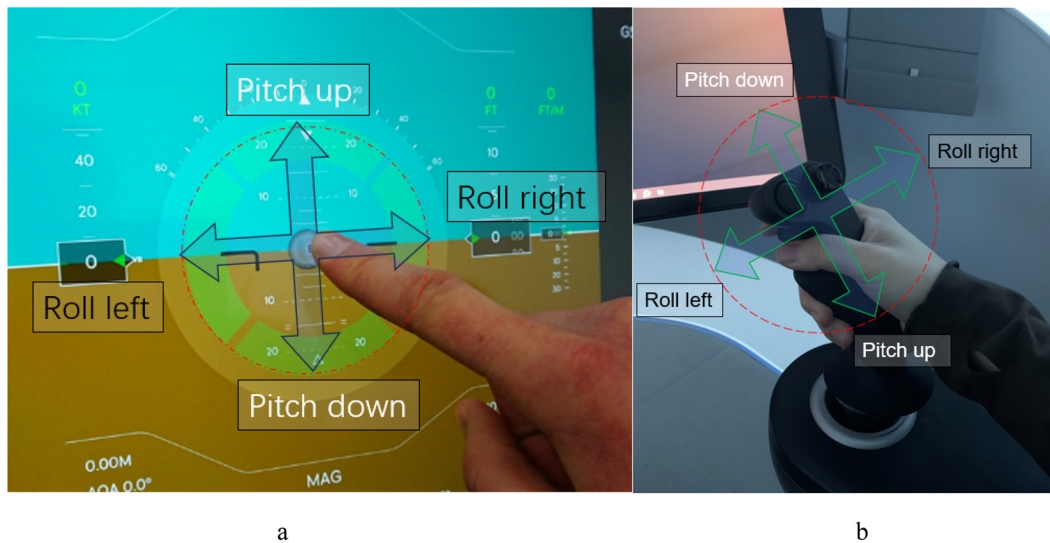


Figure 2. Input logic of touchscreen inceptor integrated with primary flight display (2a) and sidestick (2b). The interactive operational area of the touchscreen inceptor is shown by red dotted line. Dragging the stick to that boundary would stop it there, providing maximum control input to the model.

maintaining finger contact sustains control input (Asakawa, Dennerlein, and Jindrich 2017). The controller's vertical channel (pitch) was reversed based on the 'Input Equals Output' principle in Organic User Interfaces (Vertegaal & Poupyrev, 2008) and preliminary trials conducted within the authors' research group. These trials indicated that, for this prototype iteration, the inverted Y-axis design was the most intuitive option, as participants adapted more quickly and made fewer errors, despite some initial confusion among those with prior flying experience.

### 2.2.3. Performance score

Objective performance was measured using a composite performance score (PS) (Korek 2023). The performance score was based on flight path deviation, touchdown deviation, and downward gravitational force (G force). In the current study, focus was on flight path control so the flight path deviation attributing to 75% of the score and the touchdown deviation and downward G force sharing the remaining 25%. The equation of calculating the PS is:

$$PS = 100 - PP_{RMS} - (PP_{AccTD} + PP_{DesTD}) - PP_G$$

Where:

PS – performance score (landing with non-disturbance and disturbance).

$PP_{RMS}$  – Penalty points for deviance from baseline landing flight path, normalised to range 0 to 75.

$PP_{AccTD}$  – Penalty points for acceptable touchdown location error from baseline landing, normalised to range 0 to 5.

$PP_{DesTD}$  – Penalty points for desirable touchdown location error from baseline landing, normalised to range 0 to 5.

$PP_G$  – Penalty points for maximum downwards G compared with baseline landing, normalised to range 0 to 15.

Historically, researchers have utilised either spatial metrics (such as Root Mean Square Error - RMSE, Mean Average Error - MAE) or temporal metrics (such as time) independently. The scoring system introduced in this study permitted the incorporation of any combination of measures, with the flexibility to adjust the weights according to research requirements. This is analogous to the weighted NASA-TLX scale but for the measurement of objective performance. Additionally, defining a performance score with specified bounds ensured the exclusion of outliers. The participants in this study were selected (based on eye-tracking data quality) from a broader trial conducted by one of the authors, which provided a comprehensive analysis of

the weighting, bounds, and outlier detection used in the PS development. Acceptable touchdown location error was defined as being within 500 feet of the designated aim point on the runway, representing the threshold for adequate performance. In contrast, desirable touchdown location error was defined as being within 250 feet of the aim point, indicating superior performance and closer alignment with the intended landing zone (Mitchell 2019). Differentiating between acceptable and desirable locations allowed for a more granular evaluation of participant performance and facilitated the calculation of penalties based on proximity to the aim point. For an in-depth discussion on the methodology and rationale behind the PS calculation and weighting, please refer to Korek (2023). For the reproducibility, the recorded experimental data subset is publicly accessible as an [Appendix A](#).

### 2.2.4. Eye tracker

The research utilised the Pupil Labs eye tracker, comprised of a plastic frame, worn like glasses, that is fitted with a participant-facing 'eye' camera and forward-facing 'world' camera. The sampling rate of the eye tracker used in the study was 60 frames per second, with a resolution of 1280\*720 pixels. The eye camera was used to track the right eye of participants and collect data on pupil parameters, such as pupil diameter and blink rate. The data were pre-processed using the software Pupil Player developed by Pupil Labs. For the purpose of this research, fixations were defined as consecutive gaze points that fell within a 1.5° radius and lasted between 150ms and 800ms, which is consistent with previous research (W.C. Li, Wang, et al. 2022). The lighting situation during the experiment was a simulated cabin environment, with consistent illumination for all participants. The eye tracker was calibrated for each participant prior to the trial.

Predefined areas of interest (AOIs) were used for the analysis of eye movement parameters without adjusting the boundary. Gaze parameters selected for participant attention analysis were *proportion of fixation count* and *fixation duration* to two AOIs: Primary Flight Display (PFD) and out the window (OTW). These two AOIs were selected as the landing task primarily requires allocation of attention to the PFD and OTW (Foyle, Goodman, and Hooey 2003), thus differences in fixation count between these two areas can reflect hierarchical SA prioritises. Additionally, fixation duration was selected because it reflects the time spent on information acquisition and processing. Longer fixation durations may be correlated with the difficulty of obtaining information (Peißl, Wickens, and Baruah 2018).

Participant blink rate was recorded to reflect changes in participant mental workload. Blink frequency was positively correlated with the theoretical predictive value of workload. The reason for this being that blink frequency is related to visual load (Shuang, Xiaoru, and Damin 2014).

### 2.2.5. Subjective questionnaire

The situation awareness rating technique (SART) was used to subjectively assess SA issues related to touch-screen implementation within future flight decks. The SART is one of the most widely used subjective assessment for SA (Taylor 2017). SART consists of 10 seven-point questions divided across three dimensions: SART-D (Demand on Attentional Resources), SART-S (Supply of Attentional Resources), and SART-U (Understanding of the Situation). A SART total score can be calculated using the formula (Bolton, Biltekoff, and Humphrey 2022):

$$\text{SART total} = (\text{supply} - \text{demand}) + \text{understanding}$$

Higher supply indicates that the surrounding environment has the capacity and information to provide sufficient SA, whereas higher demand may indicate complexity and excessive information, and higher understanding may indicate that the information is easier to access and utilise. The questionnaire is attached in the [Appendix B](#).

### 2.3. Scenarios and procedure

To compare the performance of touchscreen and sidestick, there were two scenarios used for this study: landing with disturbance and landing with no disturbance. These two scenarios comprised of the same landing route, however in the disturbance scenario the participant was required to resist a disturbance during the landing process. The scenarios incorporated simulated turbulence as a disturbance. For each task, the disturbance signal was produced using a predefined sum-of-sines forcing function, and implemented through Lone's MATLAB function (Lone 2013; Lone and Cooke 2013). The code implements 'mkmsswp' function from System IDentification Programs for AirCRAFT (SIDPAC), which generates orthogonal multi-sine sweeps output (Morelli 2002, 2003). The amplitude of the disturbance was set at 0.4 degrees, with a frequency range between 0.01 Hz and 2.5 Hz. The sampling interval (dt) was 0.02 seconds, and the total duration of the signal (T) was 135 seconds. Only one signal (m=1) was generated for each scenario. The values for the power spectrum are normalised such that

the sum of each column equals one, ensuring a balanced distribution of power across the frequencies (Korek 2023). This signal emulated 'continuous turbulence' experienced by an aircraft, rather than screen vibration or full-motion simulation. It was designed to stimulate the participant's controller movements, enabling the measurement of their performance against a predefined goal or baseline. Different disturbance signals - which still maintained the same overall amplitude, frequency range, and duration properties - were utilised for trials with the sidestick and touchscreen to minimise between trial learning effects. Signals were consistent across participants to facilitate the comparative analysis in subsequent evaluations, as shown in [Appendix C](#).

Simulated vibrations were integrated into the experiment by summing the disturbance signal with the inceptor input. The resulting values were limited between the deflection range of the inceptors to ensure the inputs remained within the operations range of the control systems. Additionally, the disturbance signal was designed to end prior to the aircraft reaching the runway threshold; This provided participants sufficient time to stabilise the aircraft, increasing the likelihood that performance differences observed during the touchdown phase were due to the characteristics of the inceptors rather than due to residual effects of the disturbance signal (Lone 2013; Lone and Cooke 2013).

The participants were asked to perform the two scenarios using sidestick and touchscreen respectively, as shown in [Figure 3](#). To minimise differences in performance between using the touchscreen inceptor and sidestick being caused by differences in familiarity, we invited participants to take a 20-minute test flight to get used to the touchscreen inceptor and the sidestick before the actual experiment. The scenario starting point was set 5 miles away to the runway in a straight line at 1400 ft (below the glideslope path). The goal was to intercept and follow the ILS indicator, then switch to follow the glideslope according to the precision approach path indicator (PAPI) lights and land the aircraft. To eliminate the influence of other operations during the landing process on the comparison of inceptors, the aircraft began with an initial speed of 150 knots. Just before touchdown, the auto-throttle disengaged automatically, the spoilers extended, and the brakes were engaged. The weather was clear with no wind impacting the flight. To minimise the impact of the learning effect, the order in which inceptor types were used was randomised for each participant, and experiments were carried out with landings without disturbance followed by landings with disturbance.

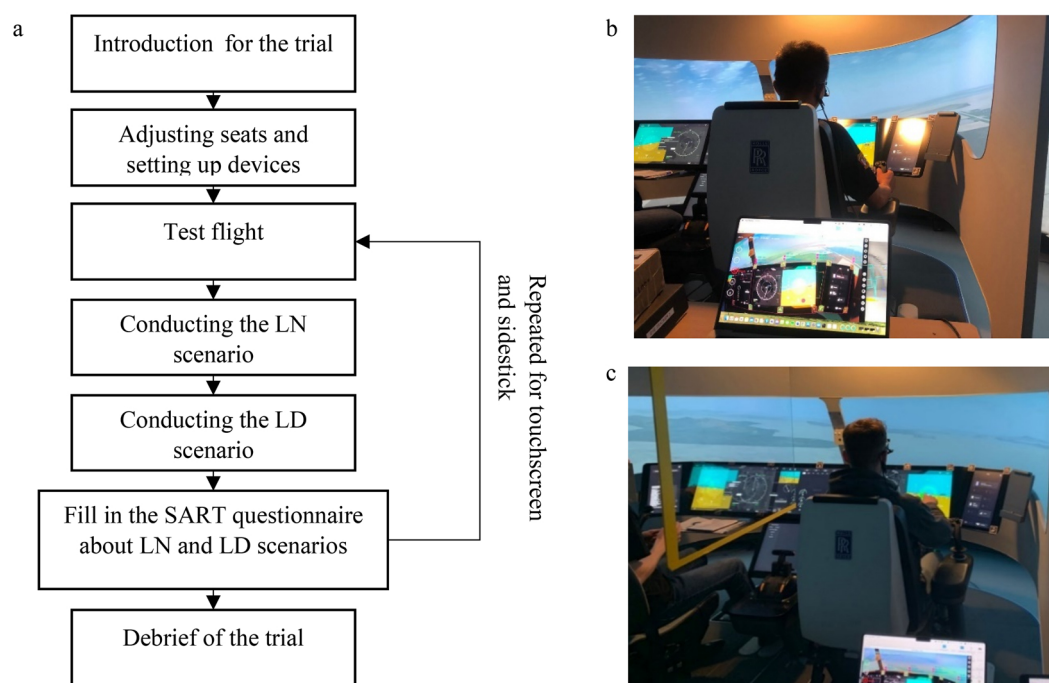


**Figure 3.** Experiment flow (a) and examples of participants performing landing with non-disturbance (LN) and landing with disturbance (LD) tasks using sidestick (b) or inceptor (c).

The experiment procedure was as follows: (1) participants provided their demographical information (5 minutes); (2) viewed a briefing video that elucidated the objective of the study and FSS layouts (10 minutes); (3) adjusted the seat and familiarised themselves with the touchscreen and sidestick in FSS (5 minutes); (4) wore and calibrated the eye tracker (10 minutes); (5) performed the test flight (15–20 minutes); (6) performed the ILS landing with disturbance and non-disturbance using two inceptors randomly and filled in the SART questionnaires (20 minutes); (6) debriefed and provided feedback about touchscreen design in flight deck (5–10 minutes). It took about 70 minutes to complete the trial for each participant.

## 2.4. Data analysis

The data for this experiment was analysed with Generalised Estimating Equations (GEE) using SPSS version 29.0.1 as the data exhibited graphical and statical signs of non-normality (high positive skew) and heterosphercity, which would undermine the robustness of traditional repeated measures ANOVA approaches. In comparison, GEE, and other longitudinal modelling methods such as linear mixed effects models, can accommodate various types of response distributions (e.g. symmetrical (normal) or asymmetrical (gamma)), and are equipped to efficiently deal with missing data. Making them particularly well-suited for data which

possess problematic parametric properties whilst preserving the flexibility of the analysis to accommodate complex interaction interpretations (de Melo et al. 2022) – unlike common non-parametric options. For this reason, GEE has been shown to provide more reliable results compared to ANOVA when small sample sizes are involved (Muth et al. 2016), a common attribute of human factors research involving costly simulation and subject matter experts (Blundell et al. 2023).

A total of 11 GEE models were estimated in this study – one per DV. Main effects and interactions were tested using the following IVs: *Inceptor* (2 levels: touchscreen/sidestick) and *Disturbance* (2 levels: present/absent). Based on the a priori power analysis for  $2 \times 2$  repeated measures using G\*power, the expected sample size for detecting effects with a medium effect size of  $f=0.25$  at a power of 0.80 is  $N=26$ . However, the current study's sample size ( $N=21$ ) is below this expectation, which may limit the detection of smaller effects. Since our data is continuous, identify link was considered. The identity link expected value of the response variable ( $\bar{\mu}$ ) is directly modelled as a linear combination of the predictors. To select the best fitting model, we tested the Gamma distribution and normal distribution with identity link. As a result, we chose normal distributions with identity link for proportion of fixation count on PFD and on OTW, and Gamma distributions with identity link for the rest of the models. These specifications were chosen because they

produced the lowest Corrected Quasi-likelihood under the Independence model Criterion (QICC), which means that they are the most suitable models (Zhao and Liang 2022). For QICC detail, see Appendix D. Robust variance estimation was chosen to improve the sensitivity of the analysis and provide robust standard errors. The Bonferroni method was applied to post-hoc analysis.

### 3. Result

#### 3.1. Sample characteristic

Descriptive results, means and standard deviations, for the different experimental conditions of each DV are presented in Table 1 below. Overall, a trend where the sidestick was superior to the touchscreen was observed across variables. Flying with the touchscreen, compared to the sidestick, led to worse manual flying performance, fewer and shorter fixations to the outside screen, higher task demand (measured via blinks and the SART-demand scale), and worse situation awareness (lower SART composite and SART-understanding scores), but slightly higher situation awareness supply.

A significant interaction effect of fixation duration on OTW, SART, and SART-D was identified. Therefore, a simple main effect analysis was conducted. The results, detailed in Table 2, include the B value, which indicates the strength and direction of the relationship, the 95% confidence interval (CI), showing the range within which the true effect likely falls, and the Wald Chi-Square, which tests the overall significance of the effect, along with the corresponding p-value.

#### 3.2. Interaction between inceptors and scenarios

##### 3.2.1. Pilot performance

As shown in the Figure 4, there was no significant interaction between inceptors and scenarios on pilot

performance score ( $B=7.17$ ,  $Wald = 1.94$ ,  $p = .16$ ) so a main effect analysis for *Inceptor* was performed. Results showed that touchscreen inceptor significantly led to reduced performance scores ( $B=-14.17$ ,  $Wald = 19.37$ ,  $p < .001$ ).

The G-forces recorded represent maximum downward forces upon touchdown and not during manoeuvring. The trial design ensured the disturbance signal ceased before participants approached the runway, allowing time to stabilise the aircraft. Notably, the maximum G-forces observed for touchscreen-controlled landings were significantly higher than those for side-stick landings. These high G-forces may stem from participants' inexperience and the reversed control logic of the touchscreen, leading to steeper descents or harder touchdowns. Automation features, such as auto-throttle disengagement and spoiler activation, also contributed to abrupt landings (Korek 2023). Interestingly, lower G-forces during touchscreen landings with disturbance may indicate a quick learning effect among participants.

##### 3.2.2. Eye tracking parameters

The GEE was conducted to analyse the gaze behaviours whilst using the two inceptors across landing scenarios based on the proportion of fixation count on the PFD. Figure 5 shows the proportion of fixation counts towards the PFD (5a) and OTW (5b). There was no significant interaction between inceptors and scenarios on proportion of fixation count on PFD ( $B = .06$ ,  $Wald = .54$ ,  $p = .46$ ) or to OTW ( $B=-.07$ ,  $Wald = 0.95$ ,  $p = .33$ ). However, a main effect of *Inceptor* for the PFD analysis was significant, which reflected a larger proportion of fixations allocated to the PFD when flying with the touchscreen ( $B = .13$ ,  $Wald = 6.35$ ,  $p = .01$ ). Similarly, a significant main effect of *Inceptor* for the OTW analysis revealed that fixations towards the OTW were lower with the touchscreen ( $B=-.14$ ,  $Wald = 9.91$ ,  $p < .01$ ).

**Table 1.** Mean and standard deviations (in parentheses) of data metrics for disturbance vs non-disturbances and sidestick vs touchscreen conditions.

| Metrics                      | Non-Disturbance |                | Disturbance     |                |
|------------------------------|-----------------|----------------|-----------------|----------------|
|                              | Sidestick       | Touchscreen    | Sidestick       | Touchscreen    |
| Performance                  | 81.94 (16.94)   | 63.90 (17.90)  | 67.69 (21.19)   | 60.21 (12.52)  |
| Proportion of Fixation Count |                 |                |                 |                |
| PFD                          | 0.79 (0.31)     | 0.90 (0.19)    | 0.63 (0.33)     | 0.79 (0.25)    |
| OTW                          | 0.18 (0.30)     | 0.07 (0.16)    | 0.33 (0.33)     | 0.15 (0.19)    |
| Fixation Duration (ms)       |                 |                |                 |                |
| PFD                          | 430.65 (123.94) | 396.66 (65.46) | 438.70 (149.20) | 394.07 (69.12) |
| OTW                          | 508.96 (107.28) | 369.59 (78.69) | 450.01 (109.86) | 397.26 (87.69) |
| Blink                        | 26.48 (23.45)   | 36.33 (27.83)  | 28.10 (28.09)   | 39.38 (25.77)  |
| SART                         | 24.38 (6.34)    | 20.52 (9.86)   | 18.52 (7.57)    | 17.33 (11.01)  |
| SART-D                       | 10.43 (4.03)    | 13.33 (4.32)   | 15.57 (4.40)    | 16.29 (5.12)   |
| SART-S                       | 19.00 (3.77)    | 19.76 (4.21)   | 19.95 (3.87)    | 20.52 (5.33)   |
| SART-U                       | 15.81 (3.12)    | 14.10 (4.85)   | 14.14 (4.45)    | 12.95 (5.35)   |

**Table 2.** GEE analysis B value and its 95% Wald Confidence Interval (CI), Wald Chi-Square, p value and Cohen's  $f^2$  of each data metric.

| Metrics                             | model             |                                        | B(95%CI)                  | Wald     | P      | $f^2$ |
|-------------------------------------|-------------------|----------------------------------------|---------------------------|----------|--------|-------|
| Performance                         | Gamma (identity)  | Interaction effect                     | 7.17 (-2.93, 17.28)       | 1.94     | 0.16   | 0.09  |
|                                     |                   | inceptor                               | -14.17 (-20.85, -7.86)    | 19.37*** | <0.001 | 0.92  |
|                                     |                   | scenario                               | -6.51 (-11.54, -1.49)     | 6.45*    | 0.01   | 0.31  |
| Proportion of Fixation Count<br>PFD | Normal (identity) | Interaction effect                     | 0.06 (-0.10, 0.21)        | 0.54     | 0.46   | 0.03  |
|                                     |                   | inceptor                               | 0.13 (0.05, 0.24)         | 6.35*    | 0.01   | 0.30  |
|                                     |                   | scenario                               | -0.13 (-0.21, -0.03)      | 7.56**   | <0.01  | 0.36  |
| OTW                                 | Normal (identity) | Interaction effect                     | -0.07(-0.20, 0.07)        | 0.95     | 0.33   | 0.05  |
|                                     |                   | inceptor                               | -0.14 (-0.23, -0.05)      | 9.91**   | <0.01  | 0.47  |
|                                     |                   | scenario                               | 0.11 (0.002, 0.22)        | 3.95*    | 0.05   | 0.19  |
| Fixation Duration<br>PFD            | Gamma (identity)  | Interaction effect                     | -22.96 (-66.72, 20.79)    | 1.06     | 0.30   | 0.05  |
|                                     |                   | inceptor                               | -56.19 (-87.29, -25.09)   | 12.54*** | <0.001 | 0.60  |
|                                     |                   | scenario                               | 4.82 (-14.77, 24.41)      | 0.23     | 0.63   | 0.01  |
| OTW                                 | Gamma (identity)  | Interaction effect                     | 86.61 (15.69, 157.54)     | 5.73*    | 0.02   | 0.27  |
|                                     |                   | Simple main effect of inceptors within |                           |          |        |       |
|                                     |                   | Non-Disturbance                        | -139.37 (-214.19, -64.54) | 13.33*** | <0.001 | 0.64  |
|                                     |                   | Disturbance                            | -52.75 (-110.82, 5.32)    | 3.17     | 0.08   | 0.15  |
|                                     |                   | Simple main effect of scenarios within |                           |          |        |       |
|                                     |                   | Sidestick                              | -58.94 (-116.42, -1.47)   | 4.04*    | 0.04   | 0.19  |
| Blink                               | Gamma (identity)  | Touchscreen                            | 27.67 (-21.22, 76.56)     | 1.23     | 0.27   | 0.06  |
|                                     |                   | Interaction effect                     | 1.43 (-5.00, 7.86)        | 0.19     | 0.66   | 0.01  |
|                                     |                   | inceptor                               | 10.53 (2.08, 18.97)       | 5.97*    | 0.01   | 0.28  |
| SART                                | Gamma (identity)  | scenario                               | 2.11 (-1.20, 5.41)        | 1.56     | 0.21   | 0.07  |
|                                     |                   | Interaction effect                     | 2.67 (0.61, 4.72)         | 6.47*    | 0.01   | 0.31  |
|                                     |                   | Simple main effect of inceptors within |                           |          |        |       |
|                                     |                   | Non-Disturbance                        | -3.86 (-7.54, -0.17)      | 4.20*    | 0.04   | 0.20  |
|                                     |                   | Disturbance                            | -1.19 (-4.99, 2.60)       | 0.38     | 0.54   | 0.02  |
|                                     |                   | Simple main effect of scenarios within |                           |          |        |       |
|                                     |                   | Sidestick                              | -5.86 (-8.85, -2.86)      | 14.71*** | <0.001 | 0.70  |
|                                     |                   | Touchscreen                            | -3.19 (-5.52, -0.86)      | 7.18**   | <0.01  | 0.34  |
| SART-D                              | Gamma (identity)  | Interaction effect                     | -2.19 (-3.48, -0.90)      | 11.06*** | <0.001 | 0.53  |
|                                     |                   | Simple main effect of inceptors within |                           |          |        |       |
|                                     |                   | Non-Disturbance                        | 2.90 (0.72, 5.09)         | 6.82**   | <0.01  | 0.33  |
|                                     |                   | Disturbance                            | 0.71 (-1.72, 3.15)        | 0.33     | 0.57   | 0.02  |
|                                     |                   | Simple main effect of scenarios within |                           |          |        |       |
|                                     |                   | Sidestick                              | 5.14 (3.65, 6.64)         | 45.46*** | <0.001 | 2.16  |
| SART-S                              | Gamma (identity)  | Touchscreen                            | 2.95 (1.47, 4.43)         | 15.32*** | <0.001 | 0.73  |
|                                     |                   | Interaction effect                     | -0.19 (-1.06, 0.68)       | 0.18     | 0.67   | 0.01  |
|                                     |                   | inceptor                               | 0.67 (-1.12, 2.46)        | 0.54     | 0.46   | 0.03  |
| SART-U                              | Gamma (identity)  | scenario                               | 0.86 (-0.27, 1.99)        | 2.24     | 0.13   | 0.11  |
|                                     |                   | Interaction effect                     | 0.52 (-0.38, 1.43)        | 1.30     | 0.25   | 0.06  |
|                                     |                   | inceptor                               | -1.43 (-3.37, 0.52)       | 2.06     | 0.15   | 0.10  |
|                                     |                   | scenario                               | -1.38 (-2.53, -0.23)      | 5.55*    | 0.02   | 0.26  |

a. The sidestick and Non-disturbance were set as the baseline for the GEE analysis, respectively.

b. \* $p < 0.05$  \*\* $p < 0.01$  \*\*\* $p < 0.001$ .

Figure 6 shows the average fixation durations towards the PFD (6a) and OTW (6b). No significant interaction between inceptors and scenarios on fixation duration on the PFD was found ( $B = -22.96$ , Wald = 1.06,  $p = .30$ ), though the main effect of *Inceptor* was significant – representing lower average fixation durations on the PFD with the touchscreen ( $B = -56.19$ , Wald = 12.54,  $p < .001$ ). For fixation durations to OTW there was a significant interaction between inceptors and scenarios ( $B = 86.61$ , Wald = 5.73,  $p = .02$ ). The simple main effect analysis showed that whilst fixation durations to OTW were longer with the sidestick compared to the touchscreen, the difference was only significant during the landings without the disturbance ( $B = -139.37$ , Wald = 13.33,  $p < .001$ ) and not for landings with the disturbance ( $B = -52.75$ ,

Wald = 3.17,  $p = .08$ ). Though this difference was close to significance ( $< .01$ ).

Figure 7 presents participant blink count results. As a measure of mental workload, there was no significant interaction between inceptors and scenarios on blink count ( $B = 1.43$ , Wald = .19,  $p = .66$ ). However, a main effect of *Inceptor* was observed whereby a higher frequency of blinks was found for trials with the touchscreen across landing conditions ( $B = 10.53$ , Wald = 5.97,  $p = .01$ ).

### 3.2.3. SART

Several GEE analyses were conducted to analyse the interactions between inceptors and scenarios on SART total score and its three dimensions: SART-D, SART-S,

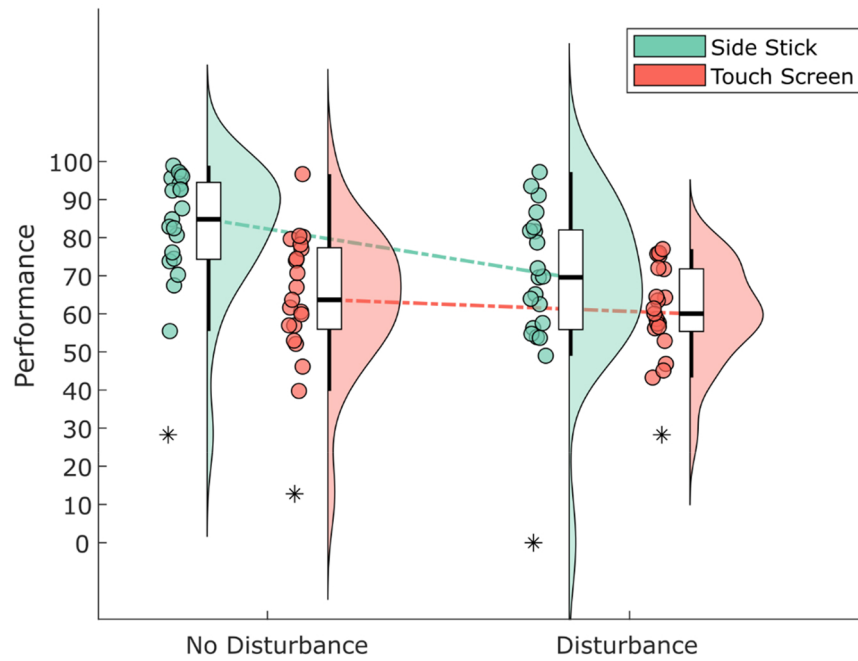


**Figure 4.** Participants' performance while interacting with the touchscreen inceptor and sidestick during landings without and with the disturbance.

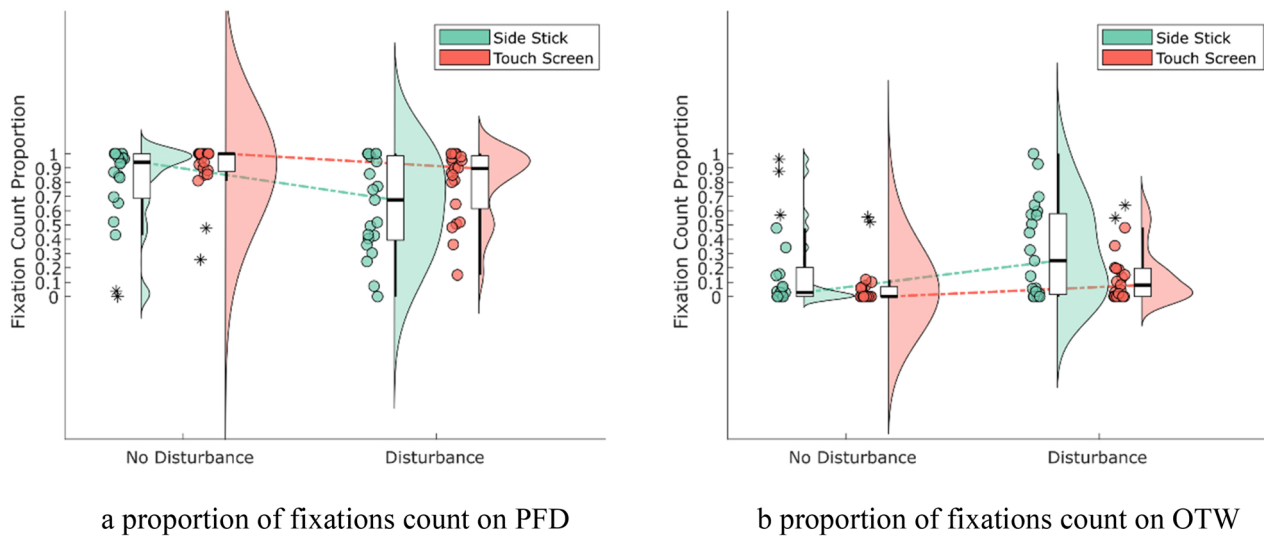


**Figure 5.** Participants' proportion of fixations count on the PFD and OTW while interacting with the touchscreen inceptor and sidestick during landings with and without the disturbance.

SART-U. **Figure 8** shows participants SART scores. There was a significant interaction between inceptors and scenarios on SART total score ( $B=2.67$ , Wald = 6.47,  $p = .01$ ). Simple main effect analysis showed that SART total scores were significantly lower with the touchscreen during non-disturbance scenario ( $B=-3.86$ , Wald = 4.20,  $p = .04$ ) but not in disturbance scenario ( $B=-1.19$ , Wald = .38,  $p = .54$ ). Similarly, for SART-D there was a significant interaction between inceptors and scenarios ( $B=-2.19$ , Wald = 11.06,  $p <$

.001). The simple main effect analysis showed that SART-D scores were higher with the touchscreen on non-disturbance scenario ( $B=2.90$ , Wald = 6.82,  $p < .01$ ), but not within disturbance scenario ( $B = .71$ , Wald = .33,  $p = .57$ ). In contrast, there was no significant interaction between inceptors and scenarios on SART-S ( $B=-.19$ , Wald = .18,  $p = .67$ ) and SART-U ( $B = .52$ , Wald = 1.30,  $p = .25$ ). Nor were there main effects of *Inceptor* for SART-S ( $B = .67$ , Wald = .54,  $p = .46$ ) or SART-U ( $B=-1.43$ , Wald = 2.06,  $p = .15$ ). Overall, SART

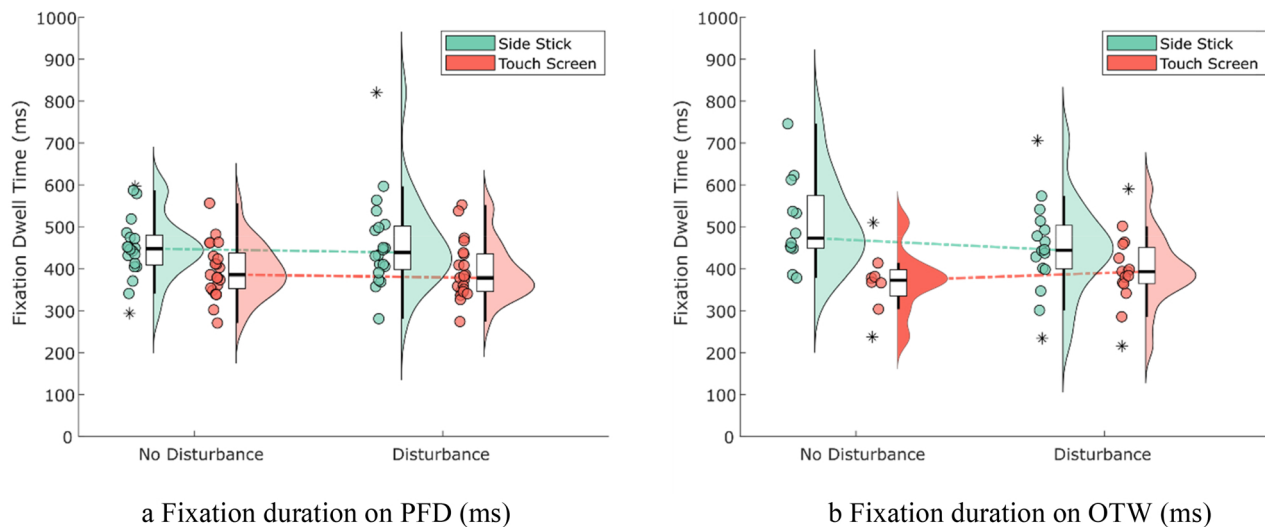


**Figure 6.** Participants' fixation duration on PFD and OTW (ms) while interacting with touchscreen inceptor and sidestick in landing with disturbance and landing with non-disturbance scenarios.

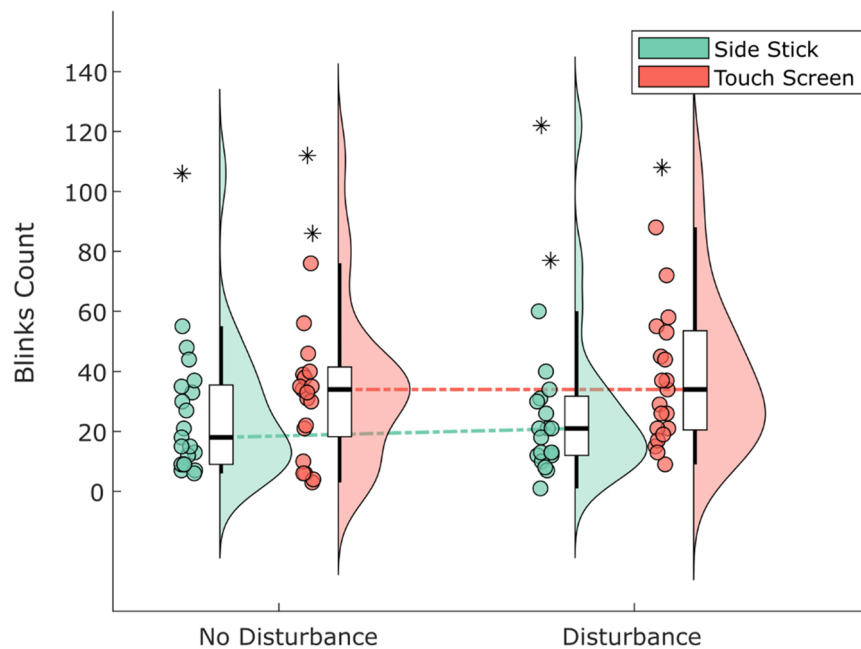


**Figure 7.** Participants' blink count while interacting with touchscreen inceptor and sidestick in landing with disturbance and landing with non-disturbance scenarios.

total scores were lower for touchscreens which was due to their higher SA demand requirement.

## 4. Discussions

### 4.1. Touchscreen's impact on pilot performance

Although using touchscreen aids can increase the integration of cockpit information (Kaminani 2011), it also presents difficulties for novel human-machine interactions. Since research on using touchscreens as an

inceptor is still in its early stages, it is crucial to comprehend how their use will affect pilot performance and situation awareness (SA). In the current study, participants' manual flying performance scores during an approach and landing task were significantly worse when using a touchscreen inceptor compared to a side-stick across different disturbance conditions, leading to the first hypothesis to be accepted. The performance score is derived from the interpolation calculation of the test pilot's flight trajectory and the standard trajectory. The lower performance observed when using a

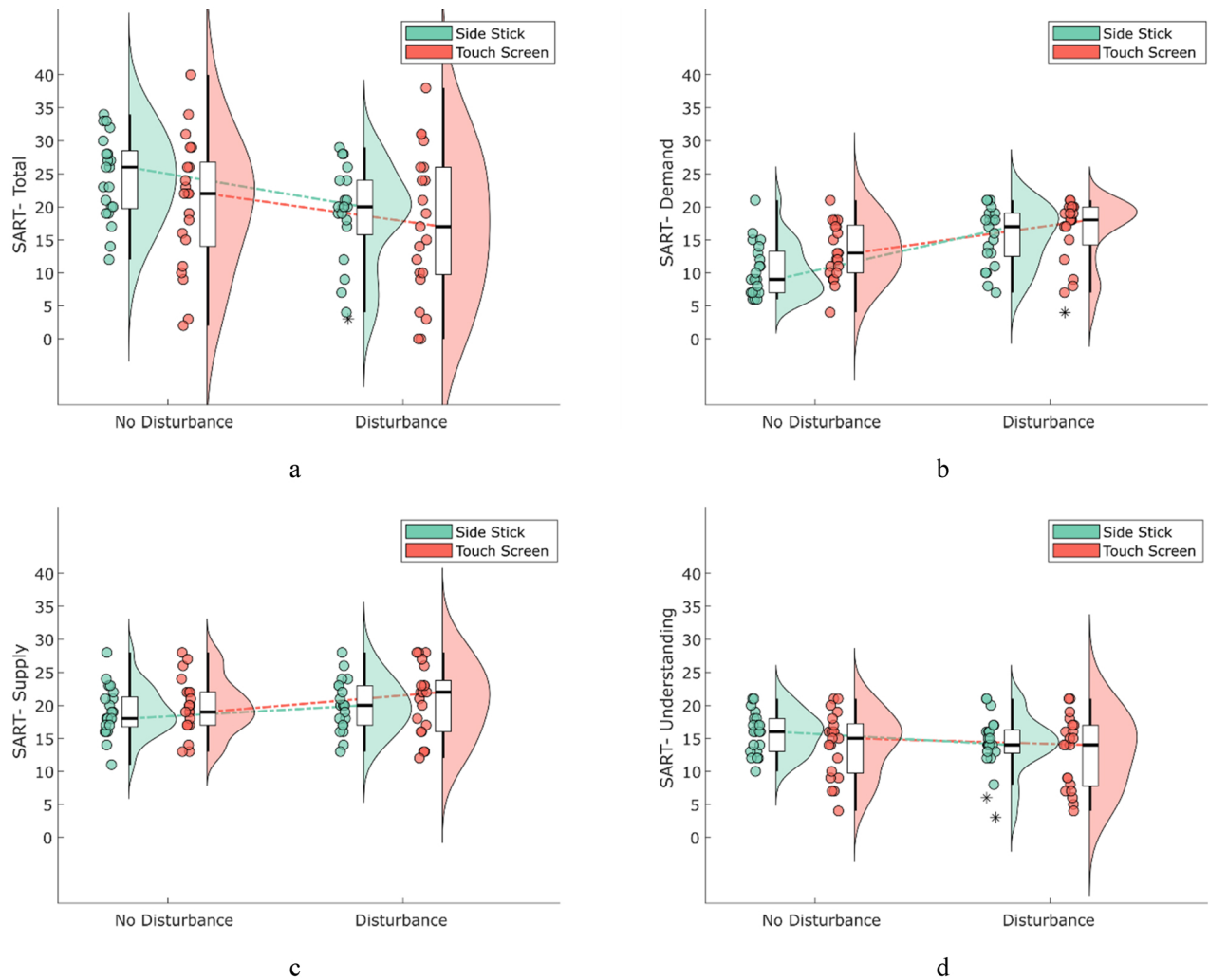


**Figure 8.** Participants' SART score and its three dimensions while interacting with the touchscreen inceptor and sidestick during landings with and without the disturbance.

touchscreen compared to a traditional sidestick can be attributed to several factors beyond simple familiarity as mentioned in previous work (Wang et al. 2024). Traditional sidesticks, having been refined over years of use, offering critical force feedback cues that are communicated through the human neuromuscular system which, in turn, aid precise fine motor control; A feature that is currently relatively lacking with touchscreen interaction. The absence of physical feedback on the touchscreen likely contributes to the decreased accuracy in pilot fine motor operations, a concern highlighted in previous studies (Wynne et al. 2021).

Moreover, the lack of arm or elbow support during touchscreen operation likely exacerbates fatigue, which can further impair performance (Khoshnewiszadeh and Pool 2021). In previous studies, participants have also reported that prolonged use of the touchscreen fatigued their arms, thus affecting precision (Wang et al. 2024). This suggests that fatigue, and the inherent design

limitations of touchscreens, are significant factors influencing their effectiveness as inceptors. In addition to the fatigue issues raised in previous studies regarding the use of touchscreens in the flight deck, this experiment goes a step further and uses eye tracking to uncover the reasons why a touchscreen inceptor performs poorly (discussed further in Section 4.2). Furthermore, whilst turbulence signals without simulated physical vibration were used in this experiment, similar results to previous experiments utilising physical vibration and turbulence were obtained: both turbulence signal and vibration had a significant negative effect on the quality of the aircraft manoeuvres (Alapetite et al. 2018; Dodd et al. 2014). To enhance the usability of touchscreens in the unstable vibrational flight deck environment, it is necessary to add supports to assist in stabilising the wrist and incorporate active built-in programs to mitigate turbulence (Cantu et al. 2021; K. Kim et al. 2021; Mobertz et al. 2018).

#### 4.1.1. G-force findings

Whilst not specifically examined within the hypothesis testing, the substantially higher G-forces observed during touchscreen-controlled landings suggest that participants had more difficulty achieving smooth touchdowns with the touchscreen inceptor compared to the traditional sidestick. This could partly be due to the touchscreen inceptor's control logic requiring upward finger movements to pitch the aircraft down, which is opposite to conventional aircraft controls. This reversed input may have caused confusion, leading to control errors such as overcorrection or delayed responses during the critical landing phase. However, it was observed that the inverted mapping usually caused more errors among participants with more flight experience compared to those who had never flown any aircraft, suggesting a familiarity bias. Furthermore, the novelty of the touchscreen interface, combined with limited tactile feedback, which can provide pilots with situation awareness information to maintain aircraft control without visual cues, could have impaired the necessary fine motor control during flare and touchdown (Cappello et al. 2020; McGrath, 1999). Participants unfamiliar with the touchscreen's sensitivity might have presented a challenge to modulating a precise descent rate, resulting in harder landings and higher G-forces. Additionally, the participant pool included non-pilots and individuals with varying levels of flight experience. Inexperience may have exacerbated difficulties in executing smooth landings with an unfamiliar control interface. The inclusion of automation features, such as auto-throttle disengagement and spoiler activation, could have introduced additional complexity, potentially contributing to abrupt touchdowns when using the touchscreen.

The weighting of G-force in the performance score was relatively low (15%) because the primary focus of the study was on flight path correction rather than landing finesse. However, the significant discrepancies in G-forces highlight the need to re-evaluate the weightings in future studies or to analyse G-force as a separate performance metric to fully capture the impact of touchscreen inceptors on landing quality and passenger comfort.

Interestingly, the lower G-forces recorded for touchscreen landings under disturbance conditions might suggest a quick learning curve or increased caution exercised by participants when handling disturbances. This could indicate that, with practice, operators might adapt to the touchscreen interface more effectively, mitigating some of the initial challenges observed.

#### 4.2. Visual behaviours different from using sidestick and touchscreen

##### 4.2.1. Gaze behaviours

The research findings supported H3, that participants using the touchscreen inceptor fixated less often out-the-window (OTW) compared to the sidestick (H3a). Furthermore, during disturbance conditions fixation durations to OTW were shorter for the touchscreen inceptor (H3b). Visual behaviour is highly correlated with attention distribution and pilot situation awareness (Yu et al. 2014). Monitoring participants' fixation locations using an eye tracker during the flight process can effectively analyse their attention distribution and compare it with their performance. In this study, our main interest about fixation counts and fixation duration is on the PFD and OTW because these two variables encompass most of the important information during an ILS landing (W. C. Li, Wang, et al. 2022). The data reveals that when using the touchscreen, pilots exhibited a significantly higher proportion of fixation on the PFD compared to when using the sidestick. This suggests that the touchscreen requires pilots to concentrate more on the PFD, potentially due to the need for more frequent monitoring and interaction with the digital interface. This increased focus on the PFD may indicate that pilots are compensating for the touchscreen's lack of tactile feedback by visually confirming their actions more often, which could lead to cognitive overload and reduced situation awareness. As shown in Figure 9 below, the heatmap example reveals that the participant's visual focus is primarily concentrated on two main areas: the Primary Flight Display (PFD) and the Out-the-Window (OTW). Additionally, the results also indicated that the proportion of fixation count on OTW was significantly higher in landing with disturbance than in landing with non-disturbance, suggesting that pilots tend to allocate more attention resources to OTW during more complex landing scenarios (Lounis, Peysakhovich, and Causse 2021; Rainieri et al. 2021). This shift is likely an adaptive strategy to ensure accurate alignment and descent by gathering critical visual information from outside the cockpit, especially when internal displays alone may not provide sufficient data for safe landing. In summary, the fixation count results allow for the drawing of two conclusions: 1) pilots need to spend more time on the PFD when using touchscreen, and 2) pilots need to get more information from the OTW when there is a higher disturbance.

Fixation duration indicates how focused an observer is in a certain region and how deeply the participants are processing information there (Einhäuser, Atzert,

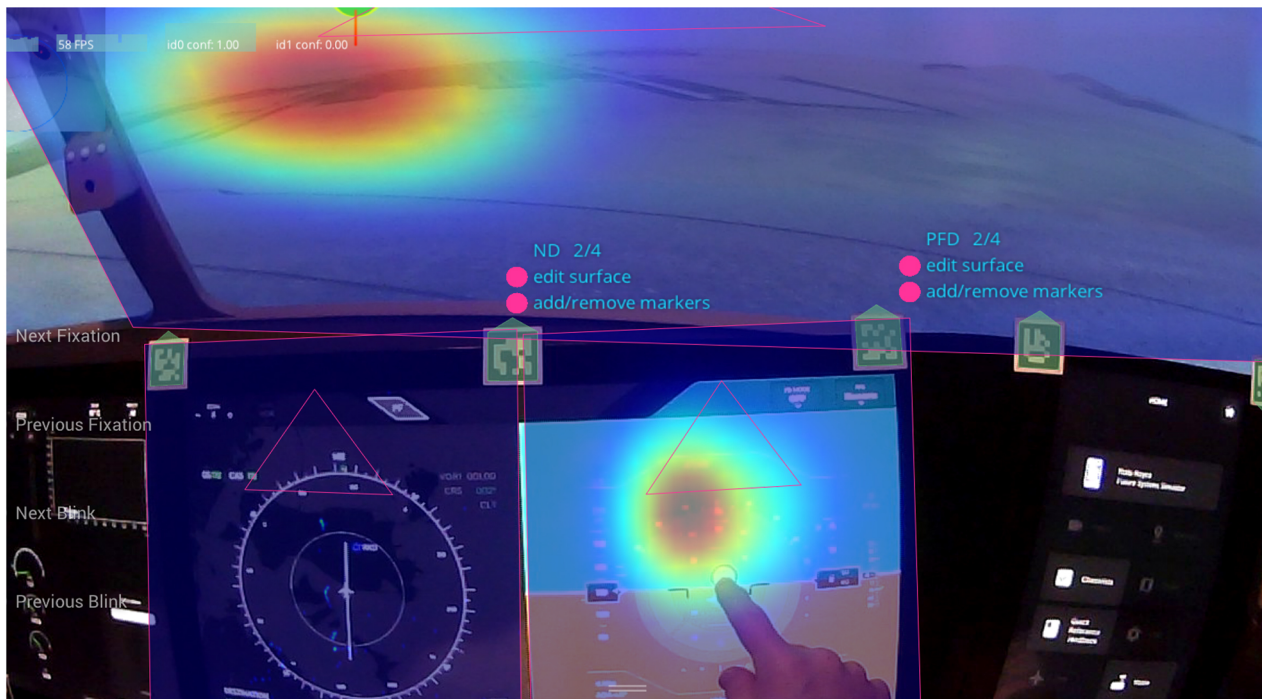


**Figure 9.** Example of participant visual heatmap during landing while using touchscreen inceptor across different AOIs.

and Nuthmann 2020; Mahanama et al. 2022). In general, the total average fixation duration when using touchscreen as inceptor was significantly lower than using sidestick. Fixation durations whilst using the touchscreen were significantly lower towards the PFD (56.19 ms) and OTW (139.37 ms during non-disturbance landings), suggesting that participants tended to scan more frequently rather than fixating on specific areas for extended periods when using the touchscreen. Although 50 ms may seem brief, it is sufficient for a saccade to shift the pilot's gaze by approximately 0.25 m on a 1-meter-distant display, based on established saccade velocity-amplitude relationships (Gibaldi and Sabatini 2021). This time period is adequate for transferring attention between different displays or key information areas. This increased scanning behaviour could be due to participants needing to be in constant contact with the PFD. During landing with disturbance scenario, the fixation durations to OTW were significantly higher than landing with non-disturbance, indicating that during less stable landings, participants need more time to process the information from outside.

There are several explanations for participants executing more frequent, shorter duration fixations to the PFD whilst using a touchscreen inceptor; 1) Visual search requirements: Given the lack of tactile feedback, touchscreen interaction may require more frequent visual searches need to find the correct control locations (W. C. Li et al. 2019); 2) Feedback and

mastery difficulty: Participants may find it challenging to ascertain the success of touchscreen operation compared to the operation of traditional sidesticks which provide immediate haptic feedback (Causse et al. 2019). This may have led to more frequent fixations to confirm their operations. 3) Cognitive workload: Participants' shorter fixation durations may reflect higher cognitive workload, and lower available spare attentional capacity, imposed by task requirements demanding visual attention be 'stretched' over a physical area, which in turn compromise participant situation awareness and task performance in general (Bednarik, Eivazi, and Vrzakova 2013). Differences in visual behaviour stem from the nature of interaction with each inceptor. The touchscreen required more frequent adjustments, leading to shorter fixations on the PFD, while the sidestick allowed for more intuitive control with fewer but longer fixations. However, regardless of the inceptor used during landing, participants needed to pay attention to the runway information in OTW. Due to the need for precise control during landing, the use of touchscreen increased task demands by requiring constant visual feedback from the PFD, which resulting in less frequent and shorter duration fixations to OTW. Overall, the touchscreen inceptor appears to have a negative impact on situational awareness (SA) in this context. Contrary to our expectations, direct aircraft control using a touchscreen may lead to suboptimal SA distribution, as increased focus on the PFD reduces critical attention to the OTW (PeiBl, Wickens,

and Baruah 2018; Ziv 2016). This finding highlights the need for caution in touchscreen use and design for aviation applications.

#### 4.2.2. *Blink count*

The blink count was only significantly impacted by input mode, indicating that H4 was accepted. Blink frequency has long been recognised as correlating with cognitive workload, whereby high workload may lead to a higher blink frequency (Pfleger et al. 2016), as our findings reaffirm. The results showed that participants produced more blinks when using touchscreen inceptor compared with using sidestick, indicating that participants experienced higher workload when using touchscreen inceptor (Chen and Epps 2014). This workload finding is in keeping with the current OTW fixation duration findings, as well as the SART demand scale results. However, blink frequency was not significantly different between the landing with non-disturbance scenario. This implies that the increased disturbance may not have a significant effect on workload since the blink count could reflect the workload level (Chen and Epps 2014; Pfeleger et al. 2016; Shuang, Xiaoru, and Damin 2014). However, there is some conflict with the performance results: participants' performance was significantly lower in the disturbance case than in the non-disturbance case. A possible reason for the decrease in performance may be the lack of situation awareness: combined with the distribution of fixation, the increased importance of external information in the disturbance situation further squeezed participants' attentional resources, resulting in a decrease in the participants' level of situation awareness.

#### 4.3. *Situation awareness affected by touchscreen and disturbance*

The Situation Awareness Rating Technique (SART), a reliable subjective measure of situation awareness, was employed in this study to assess participants' awareness during landings using different control interfaces. The results of SART total score revealed the hypothesis H2 was accepted: SART scores were significantly higher when participants used the sidestick compared to the touchscreen inceptor, but this effect was specific to non-disturbance landings (H2b). The reason why the touchscreen inceptor resulted in worse subjective SA stemmed from its higher attentional demand cost associated with its operation, confirmed by both the SART-demand results and by participant gaze behaviours. Consequentially, participants likely

experienced 'attention tunnelling', as the participants needed to obtain not only general flight information from the PFD, but also feedback after they had performed the flight controls (Choi, Ahn, and Seo 2020; Endsley, Bolte, and Jones 2003).

Although the performance of the touchscreen as an inceptor are currently unsatisfactory, the results of the SART's component show that the touchscreen inceptor has advantages in providing a good SA environment for pilots. Considering the three dimensions of SART, we can further analyse where the advantages and disadvantages of the touchscreen inceptor lie. Firstly, the interaction effects from the SART GEE results indicate that the increase in attentional demand, and the corresponding decrease in SA, driven by the presence of the disturbance was less severe when using the touchscreen compared to the side stick. In addition, a slight, yet not significant, improvement between trials was observed in SA supply for the touchscreen over the sidestick. The results indicate that whilst using the touchscreen inceptor SA supply levels were maintained between trials. These observations align with the aim of our research to explore the potential of touchscreen inceptors for efficient information integration. While previous studies have suggested that the highly integrated nature of touchscreens can positively influence both SA and pilot performance when used as advanced input devices (Dodd et al. 2014; van Zon et al. 2020), our findings suggest that touchscreen inceptors possess a degree of comparability, and even superiority, compared to traditional controls. The potential advantages in information integration require further experiments to validate and explore. However, based on the overall SART total and demand scores, the touchscreen inceptor is still in an immature stage. This new type of control is very unfamiliar to the participants and due to the lack of physical feedback (Watkins et al. 2018), the participants needed to spend more effort confirming discrepancy between their inputs and the actual outputs, which leads to higher demand scores and lower total SART scores compared to sidestick. Additionally, the reversed pitch control logic implemented in this study likely increased cognitive workload and contributed to control errors such as overcorrection or delayed responses, resulting in the high demand scores observed for the touchscreen inceptor. This study could not isolate the specific impact of the reversed pitch control logic from other factors, such as interface familiarity or tactile feedback, on pilot performance and situational awareness. However, future research could investigate the influence of input logic on pilots with varying levels of experience to better understand how conventional and reversed control schemes affect operational efficiency and cognitive

workload. In further optimising the touchscreen inceptor, greater consideration for providing feedback to enhance the pilot's perception of input-output feedback is required. An emerging research area addressing the lack of haptic feedback been studying the benefit of integrating electrovibration response signals within the touchscreen surface itself (K. Kim et al. 2021; Sadia et al. 2022; Vardar et al., 2017; Vuik, Pool, and Vardar 2024). Indeed, future research has already been proposed to examine of electrovibrational feedback on flightdeck touchscreen devices can be perceived under different turbulence conditions (Vuik, Pool, and Vardar 2024).

#### 4.4. Limitations and future directions

This research aimed to explore pilot manual flying performance, attention and situation awareness during an approach and landing task whilst using a touchscreen inceptor. However, there are several limitations in this study. First, due to the small sample size of participants, the generalisability of the experimental results to individuals with different levels of experience (such as highly experienced pilots, flight trainees, or non-pilots) is limited. While the preliminary results indicated no significant differences between experienced and non-experienced pilots, the study did not fully account for nuanced differences that may arise from familiarity biases or challenges related to the inverted mapping design. Future studies should incorporate flight experience as a factor in more comprehensive statistical models to explore its potential interaction with other variables. Second, the touchscreen inceptor setup is experimental in this study. While we chose to combine the touchscreen inceptor with the PFD, alternative display configurations (such as providing a unique touchscreen display panel), varying the touchscreen's responsiveness, or providing haptic feedback, could also influence performance. Additionally, the lack of physical support for the elbow or fingers, as explored in earlier studies (Asakawa, Dennerlein, and Jindrich 2017), may affect usability and performance and cause fatigue, suggesting the need for further ergonomic optimizations.

Another critical limitation of this study is the potential confounding between participants' unfamiliarity with the touchscreen inceptor and the reversed control logic implemented for pitch control. This design choice, while intentional, may have influenced performance outcomes, making it difficult to disentangle the specific contributions of these two factors (Wang et al. 2024). Future studies should aim to separate these effects, for example, by providing more extensive training to

participants on the touchscreen inceptor or by testing a control logic closer to conventional designs.

Additionally, the use of a non-moving simulator presents another limitation. While the experimental setup employed a pure interference signal to simulate disturbance, it does not account for performance under mixed vibration and disturbance conditions, or under pure vibration alone. The use of a touchscreen inceptor could be significantly hampered in the presence of turbulence or other vibrations, which would affect usability and control accuracy (Vuik, Pool, and Vardar 2024; Wynne et al. 2021). Future research should address this limitation by utilising motion-enabled simulators to more closely approximate real-world conditions and evaluate the touchscreen inceptor's performance in dynamic environments.

Future studies should explore various configurations and integration methods for touchscreen inceptors, potentially incorporating additional ergonomic features such as elbow or finger support to enhance performance. In particular, methods to minimise discrepancies between touchscreen controls and traditional sidesticks, such as electrovibrational feedback to simulate resistance forces, could be investigated to improve tactile response and control fidelity. Furthermore, the experimental setup may not fully reflect real-world applications, so experimenting with different integration approaches for touchscreen inceptors and evaluating their impact on usability will provide more insights into optimising these controls. Finally, incorporating more extensive eye tracking data could improve our understanding of gaze patterns and their effects on performance. Utilising machine learning and deep learning techniques for data analysis and prediction could offer deeper insights and more accurate assessments of pilot behaviour and performance. These methods have the potential to uncover more subtle patterns and relationships in the data that may not be apparent through traditional analyses. For example, using deep learning it is possible to combine eye tracking data with other data to achieve multimodal analyses, which can be used to predict pilot performance under different conditions, or to identify specific gaze patterns associated with higher cognitive load. Addressing these limitations in future research will contribute to a more thorough understanding of touchscreen inceptors and their effectiveness in aviation settings.

## 5. Conclusions

The findings of this research highlighted the potential difficulties of using touchscreen inceptors, whilst also demonstrating the benefit of using objective measurements - such as eye tracking - in experiments and its

potential for explaining the performance and situation awareness. As touchscreen technology matures and becomes more widely used in flight decks, exploring its scope of application, and analysing its advantages and disadvantages is an indispensable part of future flight deck design. The results support some of the research hypotheses, of particular interest is the significant effect of input mode on participants' overall manual flying performance, visual behaviours and situation awareness. The touchscreen inceptor in this research showed that currently the use of touchscreen as an inceptor is still very immature that is limited by: insufficiency of haptic feedback, current pilot training modes and familiarity with touchscreen, its potential for manual performance to be significantly more impaired by challenging flying environments (i.e. during a disturbance) and through fatigue which is more likely to occur when manoeuvring for long periods of time. Therefore, the use of touchscreen as the main inceptor is still far away. However, the potential information integration advantage of touchscreens, reflected by the resistance of subjective SA to environmental changes imposed by disturbance flying conditions, beckons the need for further research to understand the full scope of advantages the application of touchscreens can deliver within the flight deck environment. By analysing the different dimensions, it can be found that there is a significant relationship between visual behaviours and pilot performance, pilot situation awareness, and the use of eye tracking can be effective in analysing human system interaction efficiency. The multi-dimensional analysis approach used in this paper provides an effective analytical approach for future touchscreen application and evaluation studies in the flight deck.

## Acknowledgements

The authors would like to thank the PINES project Team in Cranfield. Especially Mudassir Lone, for his generous support during the project's development, and particularly Peter Beecroft from Rolls-Royce for supporting the research to be carried out in the Future Systems Simulator.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This research is co-financed by the European Union through the European Social Fund (grant number POWR.03.02.00-00-1029).

## ORCID

Wen-Chin Li  <http://orcid.org/0000-0002-8825-3701>

## References

- Alapetite, A., E. Möllenbach, A. Stockmarr, and K. Minakata. 2018. "A Rollercoaster to Model Touch Interactions during Turbulence." *Advances in Human-Computer Interaction 2018*: 1–16. doi:10.1155/2018/2698635.
- Asakawa, D. S., J. T. Dennerlein, and D. L. Jindrich. 2017. "Index Finger and Thumb Kinematics and Performance Measurements for Common Touchscreen Gestures." *Applied Ergonomics* 58: 176–181. doi:10.1016/j.apergo.2016.06.004.
- Bednarik, R., S. Eivazi, and H. Vrzakova. 2013. "A Computational Approach for Prediction of Problem-Solving Behavior Using Support Vector Machines and Eye-Tracking Data." London: Springer. doi:10.1007/978-1-4471-4784-8\_7.
- Bhalla, Mudit. R., and Anand. V. Bhalla. 2010. "Comparative Study of Various Touchscreen Technologies." *International Journal of Computer Applications* 6 (8): 12–18. doi:10.5120/1097-1433.
- Blundell, J., C. Collins, R. Sears, T. Plioutsias, J. Huddleston, D. Harris, J. Harrison, A. Kershaw, P. Harrison, and P. Lamb. 2023. "Multivariate Analysis of Gaze Behavior and Task Performance Within Interface Design Evaluation." *IEEE Transactions on Human-Machine Systems* 53 (5): 875–884. doi:10.1109/THMS.2023.3305715.
- Bolton, M., E. Bilteko, and L. Humphrey. 2022. "The Level of Measurement of Subjective Situation Awareness and Its Dimensions in the Situation Awareness Rating Technique (SART)." *IEEE Transactions on Human-Machine Systems* 52 (6): 1147–1154. doi:10.1109/THMS.2021.3121960.
- Callaway, F., A. Rangel, and T. L. Griffiths. 2021. "Fixation Patterns in Simple Choice Reflect Optimal Information Sampling." *PLoS Computational Biology* 17 (3): e1008863. doi:10.1371/JOURNAL.PCBI.1008863.
- Cantu, A., J. L. Vinot, C. Letondal, S. Pauchet, and M. Causse. 2021. "Does Folding Improve the Usability of Interactive Surfaces in Future Airliner Cockpits An Evaluation under Turbulent Conditions and Varying Cognitive Load." 32e *Conference Internationale Francophone Sur l'Interaction Homme-Machine, IHM 2021 - Actes de La Conference*. doi:10.1145/3450522.3451246.
- Cappello, L., W. Alghilan, M. Gabardi, D. Leonardis, M. Barsotti, A. Frisoli, and C. Cipriani. 2020. "Continuous Supplementary Tactile Feedback Can be Applied (and Then Removed) to Enhance Precision Manipulation." *Journal of Neuroengineering and Rehabilitation* 17 (1): 120. doi:10.1186/s12984-020-00736-9.
- Causse, M., F. Lancelot, J. Maillant, J. Behrend, M. Cousy, and N. Schneider. 2019. "Encoding Decisions and Expertise in the Operator's Eyes: Using Eye-Tracking as Input for System Adaptation." *International Journal of Human-Computer Studies* 125: 55–65. doi:10.1016/j.ijhcs.2018.12.010.
- Chen, S., and J. Epps. 2014. "Using Task-Induced Pupil Diameter and Blink Rate to Infer Cognitive Load." *Human-Computer Interaction* 29 (4): 390–413. doi:10.1080/07370024.2014.892428.
- Choi, M., S. Ahn, and J. Seo. 2020. "VR-Based Investigation of Forklift Operator Situation Awareness for Preventing Collision Accidents." *Accident Analysis and Prevention* 136: 105404. doi:10.1016/J.AAP.2019.105404.
- Chu, J., L. Wu, and J. Xu. 2021. "Touch Screen Application in Flight Deck for Civil Aircraft." In T. Z. Ahram & C. S. Falcão (Eds.), *Advances in Usability, User Experience, Wearable and*

- Assistive Technology: AHFE 2021* 275, 667–675. Cham: Springer. doi:10.1007/978-3-030-80091-8\_79/TABLES/2.
- Cockburn, A., D. Masson, C. Gutwin, P. Palanque, A. Goguy, M. Yung, C. Gris, and C. Trask. 2019. "Design and Evaluation of Braced Touch for Touchscreen Input Stabilisation." *International Journal of Human-Computer Studies* 122(February 2018): 21–37. doi:10.1016/j.ijhcs.2018.08.005.
- Coutts, L. V., K. L. Plant, M. Smith, L. Bolton, K. J. Parnell, J. Arnold, and N. A. Stanton. 2019. "Future Technology on the Flight Deck: assessing the Use of Touchscreens in Vibration Environments." *Ergonomics* 62 (2): 286–304. doi:10.1080/00140139.2018.1552013.
- de Melo, M. B., D. Daldegan-Bueno, M. G. Menezes Oliveira, and A. L. de Souza. 2022. "Beyond ANOVA and MANOVA for Repeated Measures: Advantages of Generalized Estimated Equations and Generalized Linear Mixed Models and Its Use in Neuroscience Research." *The European Journal of Neuroscience* 56 (12): 6089–6098. doi:10.1111/ejn.15858.
- Dehais, F., J. Behrend, V. Peysakhovich, M. Causse, and C. D. Wickens. 2017. "Pilot Flying and Pilot Monitoring's Aircraft State Awareness During Go-Around Execution in Aviation: A Behavioral and Eye Tracking Study." *The International Journal of Aerospace Psychology* 27 (1-2): 15–28. doi:10.1080/10508414.2017.1366269.
- Dodd, S., J. Lancaster, A. Miranda, S. Grothe, B. De Mers, and B. Rogers. 2014. "Touch Screens on the Flight Deck: The Impact of Touch Target Size, Spacing, Touch Technology and Turbulence on Pilot Performance." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 58 (1): 6–10. doi:10.1177/1541931214581002.
- Dodd, S., J. Lancaster, B. Demers, and S. Boswell. 2019. "Multi-Touch Touch Screens on the Flight Deck: The Impact of Display Location, Display Inclination Angle and Gesture Type on Pilot Performance." *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*. doi:10.1109/DASC43569.2019.9081723.
- Duchowski, A. T., K. Krejtz, I. Krejtz, C. Biele, A. Niedzielska, P. Kiefer, M. Raubal, and I. Giannopoulos. 2018. "The Index of Pupillary Activity." doi:10.1145/3173574.3173856.
- Einhäuser, W., C. Atzert, and A. Nuthmann. 2020. "Fixation Durations in Natural Scene Viewing Are Guided by Peripheral Scene Content." *Journal of Vision* 20 (4): 15. doi:10.1167/jov.20.4.15.
- Endsley, M. R., B. Bolte, and D. G. Jones. 2003. What is Situation Awareness? In M. R. Endsley (Ed.), *Designing for Situation Awareness: An Approach to User-Centered Design* (1st ed., pp. 17–33). London: CRC Press. doi:10.1201/9780203485088.
- Faure, V., R. Lobjois, and N. Benguigui. 2016. "The Effects of Driving Environment Complexity and Dual Tasking on Drivers' Mental Workload and Eye Blink Behavior." *Transportation Research Part F: Traffic Psychology and Behaviour* 40: 78–90. doi:10.1016/j.trf.2016.04.007.
- Favre, C. 1994. "Fly-by-Wire for Commercial Aircraft: The Airbus Experience." *International Journal of Control* 59 (1): 139–157. doi:10.1080/00207179408923072.
- Field, E., & Harris, D. 1998. A comparative survey of the utility of cross-cockpit linkages and autoflight systems' backfeed to the control inceptors of commercial aircraft. *Ergonomics*, 41(10), 1462–1477. doi:10.1080/001401398186216.
- Foyle, D. C., A. Goodman, and B. L. Hooey. 2003. NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays. <http://www.sti.nasa.gov>
- Gibaldi, A., and S. P. Sabatini. 2021. "The Saccade Main Sequence Revised: A Fast and Repeatable Tool for Oculomotor Analysis." *Behavior Research Methods* 53 (1): 167–187. doi:10.3758/s13428-020-01388-2.
- Hamden, E., D. Christopher Martin, B. Milliard, X. Al Su, Z. Li, J. Song, C. Zhang, L. Li, S. Xu, J. Zhang, Z. Wang, and Y. Lu. 2020. "The Comprehensive Evaluation Method of the Human Factors in the Flight Deck Based on the Physical and Behaviour Measurements." *Journal of Physics: Conference Series* 1549 (3): 032005. doi:10.1088/1742-6596/1549/3/032005.
- Hebbbar, P. A., A. A. Pashilkar, and P. Biswas. 2022. "Using Eye Tracking System for Aircraft Design – a Flight Simulator Study." *Aviation* 26 (1): 11–21. doi:10.3846/aviation.2022.16398.
- Kaminani, S. 2011. "Human Computer Interaction Issues with Touch Screen Interfaces in the Flight Deck." *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*. doi:10.1109/DASC.2011.6096098.
- Khoshnawisadeh, A., and D. M. Pool. 2021. "Mitigation of Biodynamic Feedthrough for Touchscreens on the Flight Deck." *International Journal of Human-Computer Interaction* 37 (7): 680–692. doi:10.1080/10447318.2021.1890490.
- Kim, K., J. H. Jeong, J. H. Cho, S. Kim, J. Kang, J. Ryu, and S. W. Lee. 2021. "Development of a Human-Display Interface with Vibrotactile Feedback for Real-World Assistive Applications." *Sensors (Basel, Switzerland)* 21 (2): 1–17. doi:10.3390/s21020592.
- Kim, S. H., and D. B. Kaber. 2014. "Examining the Effects of Conformal Terrain Features in Advanced Head-up Displays on Flight Performance and Pilot Situation Awareness." *Human Factors and Ergonomics in Manufacturing & Service Industries* 24 (4): 386–402. doi:10.1002/hfm.20501.
- Korek, W. T. 2023. *Research and Development of a New Touch-Screen based Inceptors Design for an Aircraft Control*. Silesian University of Technology]. [https://bip.polsl.pl/nadania\\_dr/wojciech-korek/](https://bip.polsl.pl/nadania_dr/wojciech-korek/)
- Korek, W. T., P. Beecroft, M. Lone, E. Bragado Aldana, A. Mendez, J. Enconniere, H. U. Asad, K. Grzedzinski, M. Milidere, J. Whidborne, W. C. Li, L. Lu, M. Alam, S. Asmayawati, L. Del Barrio Conde, D. Hargreaves, and D. Jenkins. 2024. "Simulation Framework and Development of the Future Systems Simulator." *The Aeronautical Journal* 128 (1330): 2754–2780. doi:10.1017/aer.2024.91.
- Krejtz, K., A. T. Duchowski, A. Niedzielska, C. Biele, and I. Krejtz. 2018. "Eye Tracking Cognitive Load Using Pupil Diameter and Microsaccades with Fixed Gaze." *PLoS One* 13 (9): e0203629. doi:10.1371/JOURNAL.PONE.0203629.
- Lawrenson, J. G., E. Graham-Rowe, F. Lorencatto, J. Burr, C. Bunce, J. J. Francis, P. Aluko, S. Rice, L. Vale, T. Peto, J. Presseau, N. Ivers, and J. M. Grimshaw. 2018. "Interventions to Increase Attendance for Diabetic Retinopathy Screening." *The Cochrane Database of Systematic Reviews* 1 (1): CD012054. doi:10.1002/14651858.CD012054.PUB2/MEDIA/CDSR/CD012054/IMAGE\_N/CD012054-CMP-002-01.PNG.
- Li, W. C., G. Braithwaite, T. Wang, M. Yung, and P. Kearney. 2020. "The Benefits of Integrated Eye Tracking with Airborne Image Recorders in the Flight Deck: A Rejected Landing Case Study." *International Journal of Industrial Ergonomics* 78: 102982. doi:10.1016/j.ergon.2020.102982.
- Li, W. C., J. Zhang, T. Le Minh, J. Cao, and L. Wang. 2019. "Visual Scan Patterns Reflect to Human-Computer Interactions on Processing Different Types of Messages in the Flight Deck." *International Journal of Industrial Ergonomics* 72: 54–60. doi:10.1016/j.ergon.2019.04.003.

- Li, W. C., Y. Wang, and W. T. Korek. 2022. "To be or Not to be? Assessment on Using Touchscreen as Inceptor in Flight Operation." *Transportation Research Procedia* 66 (C): 117–124. doi:10.1016/j.trpro.2022.12.013.
- Li, W.-C., P. Moore, J. Zhang, J. Lin, and P. Kearney. 2022. "The Impact of out-the-Window Size on Air Traffic Controllers' Visual Behaviours and Response Time on Digital Tower Operations." *International Journal of Human-Computer Studies* 166: 102880. doi:10.1016/j.ijhcs.2022.102880.
- Li, X., B. Li, and H. Liu. 2018. "Discussion on the Application of Active Side Stick on Civil Aircraft." *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10905 LNCS : 441–449. Springer Verlag. doi:10.1007/978-3-319-92046-7\_37.
- Lin, C. J., and Y. H. Lee. 2022. "Most Acceptable and Least Capable Target Sizes for Zoom-and-Click Tasks on Mobile Touchscreens- Implications for User Performance and Applications to Interface Design." *International Journal of Industrial Ergonomics* 87: 103251. doi:10.1016/j.ergon.2021.103251.
- Liu, X., Z. Liu, P. Q. Chen, Z. Y. Xie, B. J. Lai, B. Zhan, and J. R. Lao. 2021. "Human Reliability Evaluation Based on Objective and Subjective Comprehensive Method Used for Ergonomic Interface Design." *Mathematical Problems in Engineering* 2021(1): 1–16. doi:10.1155/2021/5560519.
- Lone, M. M. 2013. *Pilot Modelling for Airframe Loads Analysis*. Cranfield University. <https://dspace.lib.cranfield.ac.uk/handle/1826/10388>
- Lone, M. M., and A. K. Cooke. 2013. "Pilot-Model-in-the-Loop Simulation Environment to Study Large Aircraft Dynamics." *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 227 (3): 555–568. doi:10.1177/0954410011434342.
- Lounis, C., V. Peysakhovich, and M. Causse. 2021. "Visual Scanning Strategies in the Cockpit Are Modulated by Pilots' Expertise: A Flight Simulator Study." *PLoS One* 16 (2): e0247061. doi:10.1371/JOURNAL.PONE.0247061.
- Mahanama, B., Y. Jayawardana, S. Rengarajan, G. Jayawardana, L. Chukoskie, J. Snider, and S. Jayarathna. 2022. "Eye Movement and Pupil Measures: A Review." *Frontiers in Computer Science* 3: 733531. Frontiers Media S.A. doi:10.3389/fcomp.2021.733531.
- McGrath, B. J. 1999. *Tactile Instrument for Aviation*. Department of Aeronautical Engineering, Graduate School of Engineering, University of Sydney. <https://apps.dtic.mil/sti/citations/ADA531665>
- Mitchell, D. G. 2019. "Fifty Years of the Cooper-Harper Scale." *AIAA Scitech 2019 Forum, January*, 1–18. doi:10.2514/6.2019-0563.
- Mobertz, X. R. I., D. M. Pool, M. M. Van Paassen, and M. Mulder. 2018. "A Cybernetic Analysis of Biodynamic Effects in Touchscreen Operation in Turbulence." *AIAA Modeling and Simulation Technologies Conference*, 2018. doi:10.2514/6.2018-0115.
- Morelli, E. A. 2002. "System Identification Programs for Aircraft (SIDPAC)." *AIAA Atmospheric Flight Mechanics Conference and Exhibit*. doi:10.2514/6.2002-4704.
- Morelli, E. A. 2003. "Multiple Input Design for Real-Time Parameter Estimation in the Frequency Domain." *IFAC Proceedings Volumes* 36 (16): 639–644. doi:10.1016/S1474-6670(17)34833-4.
- Muth, C., K. L. Bales, K. Hinde, N. Maninger, S. P. Mendoza, and E. Ferrer. 2016. "Alternative Models for Small Samples in Psychological Research: Applying Linear Mixed Effects Models and Generalized Estimating Equations to Repeated Measures Data." *Educational and Psychological Measurement* 76 (1): 64–87. doi:10.1177/0013164415580432.
- Nagabhushana, S., and N. Sathyanarayana. 2010. *Lasers and Optical Instrumentation*. New Delhi: IK International Pvt Ltd.
- Noah, B., J. Li, and L. Rothrock. 2017. "An Evaluation of Touchscreen versus Keyboard/Mouse Interaction for Large Screen Process Control Displays." *Applied Ergonomics* 64: 1–13. doi:10.1016/j.apergo.2017.04.015.
- Patel, R. V., S. F. Atashzar, and M. Tavakoli. 2022. "Haptic Feedback and Force-Based Teleoperation in Surgical Robotics." *Proceedings of the IEEE* 110 (7): 1012–1027. doi:10.1109/JPROC.2022.3180052.
- Peißl, S., C. D. Wickens, and R. Baruah. 2018. "Eye-Tracking Measures in Aviation: A Selective Literature Review." *The International Journal of Aerospace Psychology* 28 (3–4): 98–112. doi:10.1080/24721840.2018.1514978.
- Pfleging, B., D. K. Fekety, A. Schmidt, and A. L. Kun. 2016. "A Model Relating Pupil Diameter to Mental Workload and Lighting Conditions." *Conference on Human Factors in Computing Systems - Proceedings*, 5776–5788. doi:10.1145/2858036.2858117.
- Rainieri, Giuseppe, Federico Fraboni, Gabriele Russo, Martin Tul, Andrea Pingitore, Alessia Tessari, and Luca Pietrantonio. 2021. "Visual Scanning Techniques and Mental Workload of Helicopter Pilots During Simulated Flight." *Aerospace Medicine and Human Performance* 92 (1): 11–19. doi:10.3357/AMHP.5681.2021.
- Robinson, T., G. Grabski, K. Yugawa, J. Jeffers, J. Green, and S. Humphrey. 2012. "Touchscreen Display Enhancements for Flight Deck Applications." *SAE International Journal of Aerospace* 5 (1): 151–156. doi:10.4271/2012-01-2116.
- Rouwhorst, W., R. Verhoeven, M. Suijkerbuijk, T. Bos, A. Maji, M. Vermaat, and R. Arents. 2017. "Use of Touch Screen Display Applications for Aircraft Flight Control." *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, 2017-September. doi:10.1109/DASC.2017.8102060.
- Sadia, B., A. Sadiq, M. Ayyildiz, and C. Basdogan. 2022. "Exploration Strategies for Tactile Graphics Displayed by Electrovibration on a Touchscreen." *International Journal of Human-Computer Studies* 160: 102760. doi:10.1016/j.ijhcs.2021.102760.
- Schwerd, S., and A. Schulte. 2022. "Experimental Assessment of Fixation-Based Attention Measurement in an Aircraft Cockpit." In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13307 LNAI. doi:10.1007/978-3-031-06086-1\_32.
- Shuang, L., W. Xiaoru, and Z. Damin. 2014. "A Quantitative Situational Awareness Model of Pilot." *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* 3 (1): 117–122. doi:10.1177/2327857914031019.
- Singh, C. H., K. Jain, and V. Mishra. 2023. "UAV-Based Terrain-Following Mapping Using LiDAR in High Undulating Catastrophic Areas." *Lecture Notes in Civil Engineering*, 304. doi:10.1007/978-3-031-19309-5\_3.
- Tang, J., H. Weiss, and L. Stirling. 2023. "A Comparison of Sensorimotor Assessment Accuracy between Augmented Reality and Touchscreen Environments." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 67 (1): 1532–1536. doi:10.1177/21695067231192547.

- Tao, D., J. Zeng, K. Liu, and X. Qu. 2021. "Effects of Control-to-Display Gain and Operation Precision Requirement on Touchscreen Operations in Vibration Environments." *Applied Ergonomics* 91: 103293. doi:10.1016/J.APERGO.2020.103293.
- Taylor, R. M. 2017. "Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design." In *Situational Awareness in Aerospace Operations (AGARD-CP-478)* (pp. 3/1 –3/17). Neuilly Sur Seine, France: NATO-AGARD.
- Van Baelen, D., M. M. van Paassen, J. Ellerbroek, D. A. Abbink, and M. Mulder. 2021. "Flying by Feeling: Communicating Flight Envelope Protection through Haptic Feedback." *International Journal of Human-Computer Interaction* 37 (7): 655–665. doi:10.1080/10447318.2021.1890489.
- van Zon, N. C. M., C. Borst, D. M. Pool, and M. M. van Paassen. 2020. "Touchscreens for Aircraft Navigation Tasks: Comparing Accuracy and Throughput of Three Flight Deck Interfaces Using Fitts' Law." *Human Factors* 62 (6): 897–908. doi:10.1177/0018720819862146.
- Vardar, Y., Güçlü, B., & Basdogan, C. 2017. Effect of waveform on tactile perception by electrovibration displayed on touch screens. *IEEE Transactions on Haptics*, 10(4), 488–499. doi:10.1109/TOH.2017.2704603.
- Vertegaal, R., & Poupyrev, I. 2008. Organic user Interfaces. *Communications of the ACM*, 51(6), 26–30. doi:10.1145/1349026.1349033.
- Vuik, J. D. A., D. M. Pool, and Y. Vardar. 2024. "Impact of Whole-Body Vibrations on Electrovibration Perception Varies with Target Stimulus Duration." <https://arxiv.org/abs/2404.18972v1>
- Wan, Y., J. C. Prinnet, and N. Sarter. 2017. "Visual and Auditory Feedback to Improve Touchscreen Usability in Turbulence." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61 (1): 89–93. doi:10.1177/1541931213601488.
- Wang, Y., W. C. Li, W. T. Korek, and G. Braithwaite. 2024. "Future Flight Deck Design: Developing an Innovative Touchscreen Inceptor Combined with the Primary Flight Display." *International Journal of Industrial Ergonomics* 101: 103588. doi:10.1016/j.ergon.2024.103588.
- Watkins, C. B., C. Nilson, S. Taylor, K. B. Medin, I. Kuljanin, and H. B. Nguyen. 2018. "Development of Touchscreen Displays for the Gulfstream G500 and G600 Symmetry™ Flight Deck." *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, 2018-September. doi:10.1109/DASC.2018.8569532.
- Wu, Q., and J. Li. 2024. "Analysis and Performance Evaluation of a Modular Finger-Mounted Haptic Device for High-Density Feedback." *Journal of Ocean University of China* 23 (6): 1583–1594. doi:10.1007/S11802-024-5845-0/METRICS.
- Wynne, R. A., K. J. Parnell, M. A. Smith, K. L. Plant, and N. A. Stanton. 2021. "Can't Touch This: Hammer Time on Touchscreen Task Performance Variability under Simulated Turbulent Flight Conditions." *International Journal of Human-Computer Interaction* 37 (7): 666–679. doi:10.1080/10447318.2021.1890492/SUPPL\_FILE/HIHC\_A\_1890492\_SM1106.ZIP.
- Xie, Y., R. Zhou, and J. Qu. 2023. "Fitts' Law on the Flight Deck: evaluating Touchscreens for Aircraft Tasks in Actual Flight Scenarios." *Ergonomics* 66 (4): 506–523. doi:10.1080/00140139.2022.2097318.
- Xue, H., Q. Zhang, and X. Zhang. 2022. "Research on the Applicability of Touchscreens in Manned/Unmanned Aerial Vehicle Cooperative Missions." *Sensors (Basel, Switzerland)* 22(1): 8435. (), doi:10.3390/S22218435.
- Yu, C. S., E. M. Y. Wang, W. C. Li, and G. Braithwaite. 2014. "Pilots' Visual Scan Patterns and Situation Awareness in Flight Operations." *Aviation, Space, and Environmental Medicine* 85 (7): 708–714. doi:10.3357/ASEM.3847.2014.
- Zhao, Q., and J. Liang. 2022. "A Q-Q Plot for Detecting Non-Multinormality Based on a Normal Characterization and the S-W Statistic." *Communications in Statistics - Theory and Methods* 51 (5): 1367–1378. doi:10.1080/03610926.2020.1761983.
- Zhou, F., X. J. Yang, and J. C. F. De Winter. 2022. "Using Eye-Tracking Data to Predict Situation Awareness in Real Time During Takeover Transitions in Conditionally Automated Driving." *IEEE Transactions on Intelligent Transportation Systems* 23 (3): 2284–2295. doi:10.1109/TITS.2021.3069776.
- Ziv, G. 2016. "Gaze Behavior and Visual Attention: A Review of Eye Tracking Studies in Aviation." *The International Journal of Aviation Psychology* 26 (3-4): 75–104. doi:10.1080/10508414.2017.1313096.

## Appendix

**Appendix A.** The Mean and standard deviation of participants' RMS, touchdown location error and max G-force when using sidestick and touchscreen in both landing with disturbance and landing with non-disturbance scenarios.

|                              | sidestick      |                |            | touchscreen    |                |              |
|------------------------------|----------------|----------------|------------|----------------|----------------|--------------|
|                              | RMS            | Location Error | G-force    | RMS            | Location Error | G-force      |
| landing with non-disturbance | 142.30(138.76) | 253.43(226.04) | 3.49(3.64) | 261.89(192.95) | 217.26(199.53) | 22.86(48.98) |
| landing with disturbance     | 255.91(289.16) | 287.18(298.52) | 5.39(4.89) | 300.61(138.04) | 342.76(340.36) | 10.96(7.55)  |

## Appendix B. Situation awareness rating technique

### Situation Awareness Rating Technique

#### **Demand**

instability of situation

low

☐ ☐ ☐ ☐ ☐ ☐ high

Variability of situation

low

☐ ☐ ☐ ☐ ☐ ☐ high

Complexity of situation

low

☐ ☐ ☐ ☐ ☐ ☐ high

#### **Supply**

Arousal

low

☐ ☐ ☐ ☐ ☐ ☐ high

Spare mental Capacity

low

☐ ☐ ☐ ☐ ☐ ☐ high

Concentration

low

☐ ☐ ☐ ☐ ☐ ☐ high

Division of Attention

low

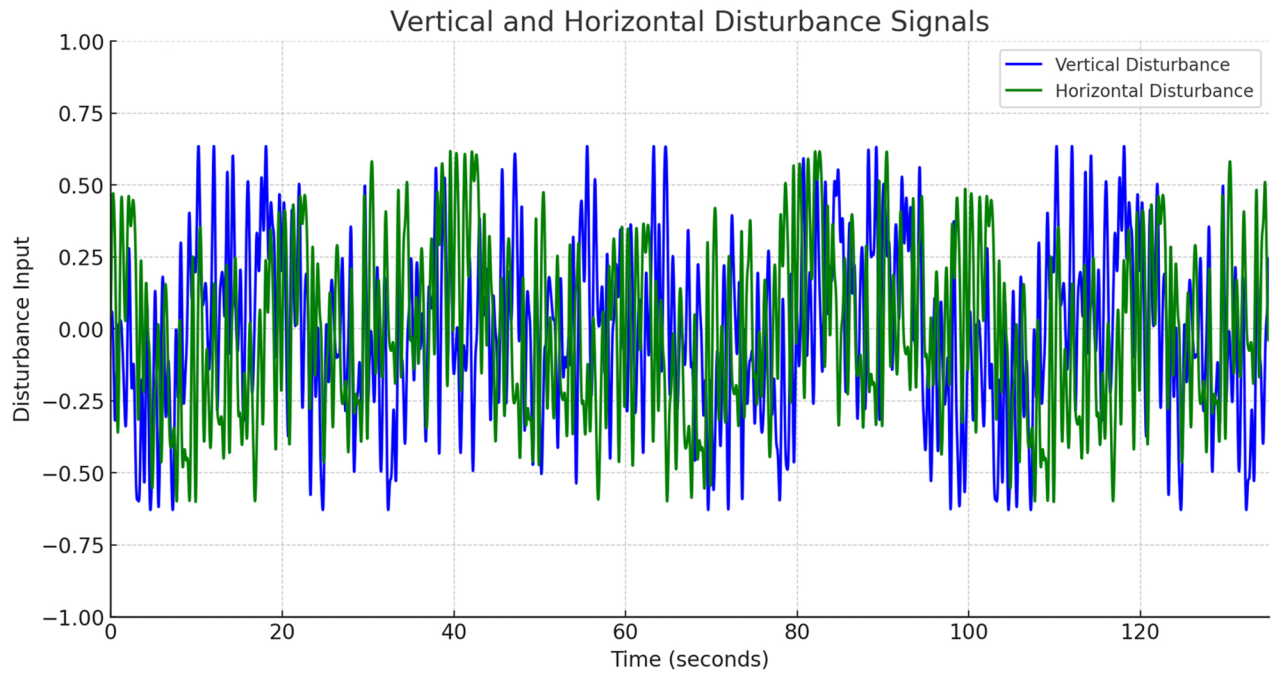
☐ ☐ ☐ ☐ ☐ ☐ high

#### **Understanding**

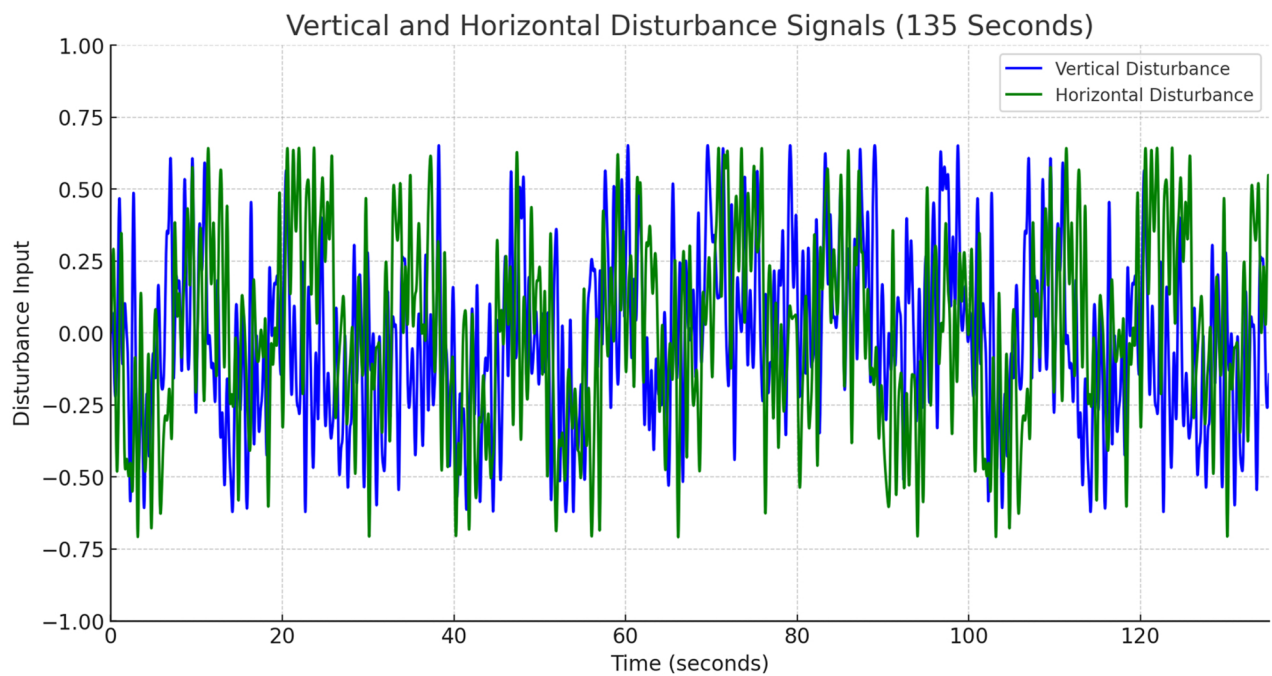
Information Quantity

low

☐ ☐ ☐ ☐ ☐ ☐ high

**Appendix C. The disturbance signal over time when using sidestick (a) and touchscreen (b)**

(a)



(b)

**Appendix D. Corrected quasi-likelihood under the independence model criterion for all models**

| Metric                       | Corrected Quasi-likelihood under the Independence model Criterion |                   |                  |
|------------------------------|-------------------------------------------------------------------|-------------------|------------------|
|                              | model                                                             | normal (identity) | Gamma (identity) |
| Performance                  |                                                                   | 24275.37          | 13.63            |
| Proportion of Fixation Count |                                                                   |                   |                  |
| PFD                          |                                                                   | 14.03             | 14.03            |
| OTW                          |                                                                   | 13.21             | 13.21            |
| FD                           |                                                                   | 499598.01         | 10.43            |
| PFD                          |                                                                   | 5512774.75        | 10.58            |
| OTW                          |                                                                   | 458668.66         | 10.47            |
| Blink                        |                                                                   | 55556.67          | 61.30            |
| SART                         |                                                                   | 6436.67           | 74.36            |
| SART-D                       |                                                                   | 1617.24           | 18.37            |
| SART-S                       |                                                                   | 1512.00           | 11.97            |
| SART-U                       |                                                                   | 1642.57           | 19.32            |



**DSPACE**

<https://dspace.org/>

**Comparative analysis of touchscreen inceptors and traditional sidesticks on flight decks: flight performance, visual behaviours and situation awareness**

**Wang, Yifan; Korek, Wojciech Tomasz; Blundell, James; Li, Wen-Chin**

**2026-02**

Taylor & Francis

<https://doi.org/10.1080/00140139.2025.2459305>

Wang Y, Korek WT, Blundell J, Li W-C. (2026) Comparative analysis of touchscreen inceptors and traditional sidesticks on flight decks: flight performance, visual behaviours and situation awareness. *Ergonomics*, Volume 69, Issue 2, February 2026, pp. 267-291