The Aeronautical Journal (2023), 1–23
doi:10.1017/aer.2022.81

REGULAR PAPER

Low-visibility commercial ground operations: An objective and subjective evaluation of a multimodal display

J. Blundell[1,*], C. Collins[2], R. Sears[2], T. Plioutsias[2], J. Huddlestone[1], D. Harris[1], J. Harrison[3], A. Kershaw[3], P. Harrison[3] and P. Lamb[3]

[1]Institute of Clean Growth and Future Mobility, Coventry University, Coventry, CV1 5FB, United Kingdom, [2]Faculty of Engineering, Environment and Computing, Coventry University, Coventry, CV1 5FB, United Kingdom and [3]BAE Systems Ltd, Rochester, ME1 2XX, United Kingdom
*Corresponding author. Email: james.blundell@coventry.ac.uk

Received: 9 May 2022; **Revised:** 8 August 2022; **Accepted:** 24 August 2022

Keywords: Head-up displays; Haptics; Multimodal; Ground operations; Performance; Workload

Abstract
Flight crews' capacity to conduct take-off and landing in near zero visibility conditions has been partially addressed by advanced surveillance and cockpit display technology. This capability is yet to be realised within the context of manoeuvring aircraft within airport terminal areas. In this paper the performance and workload benefits of user-centre designed visual and haptic taxi navigational cues, presented via a head-up display (HUD) and active sidestick, respectively, were evaluated in simulated taxiing trials by 12 professional pilots. In addition, the trials sought to examine pilot acceptance of side stick nose wheel steering. The HUD navigational cues demonstrated a significant task-specific benefit by reducing centreline deviation during turns and the frequency of major taxiway deviations. In parallel, the visual cues reduced self-report workload. Pilot's appraisal of nose wheel steering by sidestick was positive, and active sidestick cues increased confidence in the multimodal guidance construct. The study presents the first examination of how a multimodal display, combining visual and haptic cues, could support the safety and efficiency in which pilots are able to conduct a taxi navigation task in low-visibility conditions.

Nomenclature

ANOVA	Analysis of variance
ATPL	Airline transport pilots licence
CAVOK	Cloud and visibility OK
HITS	Highway in the sky
HMD	Head-mounted display
HMI	Human-machine-interfaces
HUD	Head-up display
MRT	Multiple resource theory
NASA	National aeronautics and space administration
NASA-TLX	NASA task load index
RMSE	Root-mean-squared-error
RVR	Runway visual range
SPSS	Statistical product and service solutions
UCD	User-centred design
VFR	Visual flight rules
VMC	Visual meteorological conditions

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

1.0 Introduction

Many operational concepts are emerging to improve airport terminal area operations within different international next generation air transport systems [1]. However, the precision navigational requirements of these concepts necessitate the provision of Visual Flight Rules (VFR)-like capabilities and safety irrespective of weather conditions for flight crews [2, 3]. Whilst this capacity is largely available to flight crews for take-off and landing in near zero visibility conditions, this capability is yet to be realised within the airport terminal area for taxiing and manoeuvring aircraft to and from active runways and gates. Furthermore, taxi clearances can be complex and are subject to change whilst taxiing is in progress. Working towards the goal of developing cockpit displays for future commercial aviation environments, the current study evaluates the performance and workload effects of a multimodal taxi symbology set within a human-in-the-loop taxi simulator trial.

Previous research has demonstrated that perspective visual displays capable of presenting task specific conformal and non-conformal information are able to enhance pilot performance, whilst suppressing a corresponding increase in mental workload. The most notably example from the aviation domain would be the presentation of a three-dimensional highway in the sky (HITS) to support precision landings in either fixed wing [4, 5] or rotary aircraft [6, 7]. The scientific literature on perspective displays that are designed to support aviation ground operations is far sparser. Past research findings emanating from three experiments conducted by the National Aeronautics and Space Administration (NASA) between 2007 and 2017 [8–10] have highlighted the benefits of supporting taxiing in low-visibility weather conditions using eyes-out symbology displays (i.e. head-up displays (HUDs) and head mounted displays (HMDs)). The research involved the presentation of task-relevant taxi symbology information that consisted of a ground speed, the current taxiway, next cleared taxiway, centreline markers, and virtual cones on the taxiway edge. Non-conformal taxi symbology showed the taxiway centreline, the landing gear and velocity trend line. Across these studies, pilots were better able to perform the taxi task (taxied faster and had reduced centreline deviation) when taxiing with the advanced taxi symbology (on either a HUD or HMD), compared to taxiing with a head-down moving map or paper charts, in reduced visibility conditions (night visual meteorological conditions (VMC) and day, 700feet (300m) RVR). However, despite these positive results no further research has investigated how a task specific taxiing symbology set might be able to improve the safety and efficiency of ground operations. In particular, to our knowledge, no study to date has provided a robust account of how a user-centred design (UCD)-based taxi symbology set that incorporates both conformal and non-conformal elements could benefit pilots.

The importance of optimising the presentation of task-related visual information through the process of UCD is highlighted by the potential detrimental effect of ‘clutter’. ‘Clutter’ in display design is defined as superimposing the pilots’ forward field of view with a quantity of information that is disruptive to the search of information in the environment. A key study in this area by Ververs and Wickens [11] demonstrated that HUD clutter adversely affected detection of both commanded changes on the instrumentation and aircraft in the environment. Such findings highlight the benefit of the UCD approach in developing complex human-machine interfaces (HMI) that are employed in safety critical environments. In the context of HUD design, a UCD approach is pivotal to achieving a HUD symbology set that optimises the balance between maximising task-related information whilst minimising detrimental HUD ‘cluttering’ effects.

Whilst most cockpit interfaces rely on communicating information via visual and/or audio modalities, within the past two decades a growing body of HMI research has explored the information processing benefits associated with presenting information via the haptic sensory channel, either through the proprioceptive (i.e. force feedback) and vibrotactile haptic sub-senses. This focus has largely been driven by Wicken’s multiple resource theory (MRT) [12, 13], which posits that humans have several semi-independent cognitive resources that serve different sensory modalities (e.g. visual, audio, or haptic), and that tasks requiring the use of different resources can often be effectively performed together. Furthermore, competition for the same resource can produce interference. A number of reviews of human-in-the-loop experiments [14–16] have lent credence to the potential multiple resource information processing advantage that proprioceptive and vibrotactile interfaces could provide. Studies from the surgical and teleoperations fields have found that supplementing visual feedback with haptic information can promotes a notable advantage in the enhancing the sensation of "presence". Having haptic information serving as a redundant cue in this fashion has been associated with an improvement in motor learning tasks, particularly when the trained task is complex [17]. In the aviation domain several studies by the Delft group have explored how haptic information can be used to enhance pilot situation awareness. For example, Ellerbroek *et al.* [18] described how haptic feedback, delivered through the sidestick, can communicate the intent of the aircraft's flight envelope protection systems to the pilot. In this instance, proprioceptive (i.e. stick stiffness) and vibrotactile cues (i.e. stick shaker) were used to alert the pilot to flight envelop boundary encroachment and provide protective feedback (e.g. shifting of stick centre). Results from the study found that haptic feedback improved performance without an associated increase in workload. Similar performance, situational awareness and workload benefits have been found in studies involving the use of haptic feedback cues to support path following and separation management of unmanned aerial vehicles [19, 20] and motorway driving [21]. However, no studies have investigated the benefit of haptic feedback provided to pilots in the context of manoeuvring aircraft within the airport terminal area. Likewise, no study has examined the value of haptic cues coupled with the presentation of visual HUD taxi navigational cues (i.e. haptic cues functioning as redundant taxi guidance cues).

In the current study the benefits of a UCD based multimodal taxi symbology set are examined in a human-in-the-loop taxi simulator trial. The study expands upon the findings from previous NASA research [8–10] by further maturing the adaptive symbology concept through the development and evaluation of HUD and active stick presented task-related information during low-visibility airport surface conditions. Pilots completed several different runway-to-gate and gate-to-runway routes at Munich Airport (EDDM) in low visibility conditions. Trials were designed to evaluate the performance and workload benefits of visual and haptic taxi navigational cues, both when presented in isolation and in combination with one another. In addition, the trials sought to examine the feasibility and pilot acceptance of using a side stick as the control surface for nose wheel steering, compared to traditional steering with a tiller. Given the greater versatility of potential haptic cues that can be delivered via the side stick compared to the tiller, a feasible side stick nose wheel steering concept would facilitate the future application of task-relevant haptic information within the taxiing task.

2.0 Methods

2.1 Participants

Twelve professional airline pilots, holders of an Airline Transport Pilots Licence (ATPL), participated in the study. The rank of participants included four first officers, two senior first officers and six captains. Participants' average flying experience was 19.92 years (SD = 15.03) and 5,793 hours on type (SD = 6,832). Fifty-eight percent of the participants (7/12) had experienced flying Airbus airliner types (e.g. A320/A330), the remainder flew Boeing airliner types. Participants possessed a mixture of short (8/12) and long haul (4/12) experience. Thirty-three percent of the participants had past experience of flying with a HUD (4/12). The experiment was approved by the Coventry University Ethics Committee and was in line with ethics guidelines as per the British Psychological Society.

2.2 Simulator facility

A generic fixed-wing fixed-based simulator running X-plane 11 Professional (Laminar Research) was used. The flight model of a Boeing 737 type aircraft was employed. The simulator was equipped with a 180- × 40- degree collimated projection system enabling equivalent real world depth perception for the participants that accurately simulated the experience of a HUD on an actual aircraft. Each participant

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

Figure 1. *BAE Systems Rochester simulator cockpit layout.*

was positioned in the captain seat position with the tiller located on the left, and the side stick located forward of the tiller (Fig. 1). A member of the research team, who possessed an ATPL, performed pilot monitoring duties during the trial. A custom, BAE Systems, data logging program was developed to interface with the X-Plane flight model and simulator environment in order to drive the HUD and stick guidance cues (60Hz), as well as retrieve relevant X-Plane data refs (4Hz sampling rate) for subsequent performance analyses.

2.3 UCD process

The multimodal taxi symbology set developed by BAE Systems was designed using UCD principles through consultation with subject matter experts. The UCD process was undertaken over the course of a six-month period that preceded the current human-in-the-loop taxi simulator trial. This included four full-day workshops with the subject matter expects to iteratively optimise both the intuitiveness of visual and haptic navigational cues, and the way in which the cues were coupled with one another.

The design of the symbology was in keeping with recent human factors interface design principles [22]. For example, the distance to the next cleared taxiway on the HUD was presented both numerically and as a radial, as opposed to the numerical only format employed within the previous NASA symbology set. Design of the side stick haptic cues focused on established appropriate cue onset parameters and cue signal properties. From a multimodal perspective, the UCD process emphasised the temporal coupling of visual HUD and haptic side stick navigational cues.

2.4 HUD symbology

Figure 2 demonstrates the current HUD symbology design. The non-conformal elements included a ground speed/throttle radial in the bottom-left corner of the display. At the top of the display was a raw data indicator showing the linear deviation of the nose wheel and main gear from the taxiway centreline (along with 10m deviation increment markers). On the right of the display was located a moving map and a distance to turn radial. Near the bottom of the display was a timer marking the time till the end of the taxi route. The conformal elements of the display included an overlay of the taxiway centreline denoting relevant routing information. Non-conformal depictions of the conformal symbology were also provided on the moving map display.

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

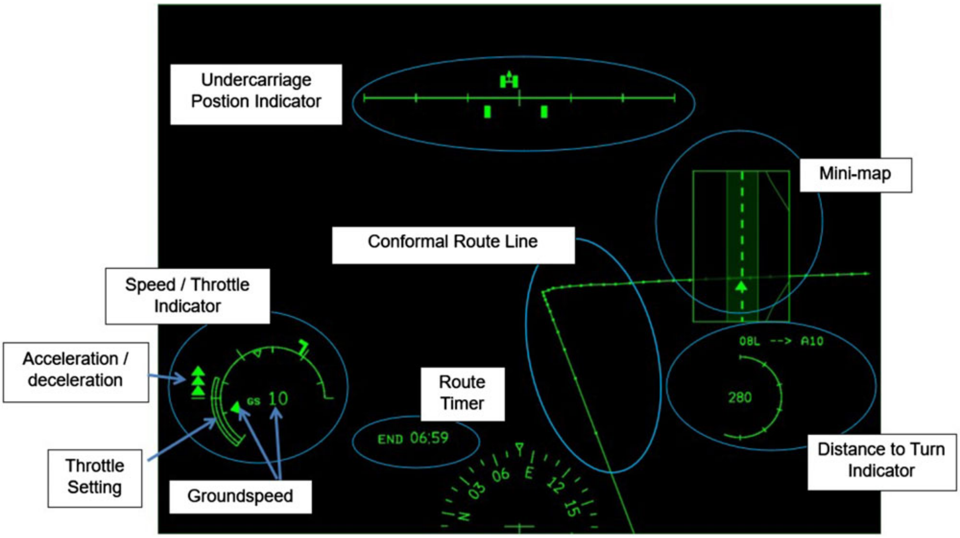

Figure 2. *Example of the taxi navigation head-up symbology. The central conformal taxiway line guidance depicts guidance to take the next exit off the runway onto taxiway A10.*

2.5 Input interfaces

Two proprioceptive guidance cues and one vibrotactile warning cue delivered through the simulator's side stick were evaluated in the trial.

The first proprioceptive cue was a dynamic guidance force that was applied laterally to protect against oversteer and nose wheel scrubbing during turning. The protective cue was created by adding a stick force profile using a spring resistance behaviour that changed as a function of groundspeed (Fig. 3). A soft stop, which activated above a groundspeed of 8Kn, was introduced to communicate nose wheel scrubbing boundaries to the participant. Furthermore, a later hard stop was positioned to coincide with a stick deflection position where excessive nose wheel turning was being demanded by the participant. A dead band was created to minimise overcorrection on the stick.

The second proprioceptive cue was a lateral nudge designed to both alert the participant to taxi centreline deviation and to provide directional corrective guidance. The cue used a discrete signal that produced a lateral force in the direction of the preferred stick movement to correct back to the taxi centreline. In addition, the signal type was categorised into two different signals: a minor and major nudge of differing scales (Fig. 4a). The minor and major nudges activated if the deviation of the nose wheel from the route line was greater than 1.5m and less than 7m or greater than 7m, respectively (Fig. 4b). The minor nudge type was not instigated during turn route segments.

The final cue was a longitudinal vibrotactile warning cue that was used to inform participants about the distance to a turn on the route (Fig. 5a). A continuous cue was initiated 160m prior to an upcoming turn with four vibrations/rumbles delivered through the side stick at 40, 80 120 and 160m before the turn should be applied (Fig. 5b).

2.6 Scenario and procedure

The experiment took place at BAE Systems Rochester, which is where the flight simulator test environment is based. Simulation scenarios were based on taxi routes at Munich airport (EDDM). Participants taxied in CAT-III visibility conditions along 5-minute (approximate) taxi routes. Based on consultation with subject matter experts two equivalent runway-to-gate EDDM taxi routes were constructed that consisted of a 120-degree, 90-degree, and S-Bend turns, and a series of straights (see Appendix A for

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

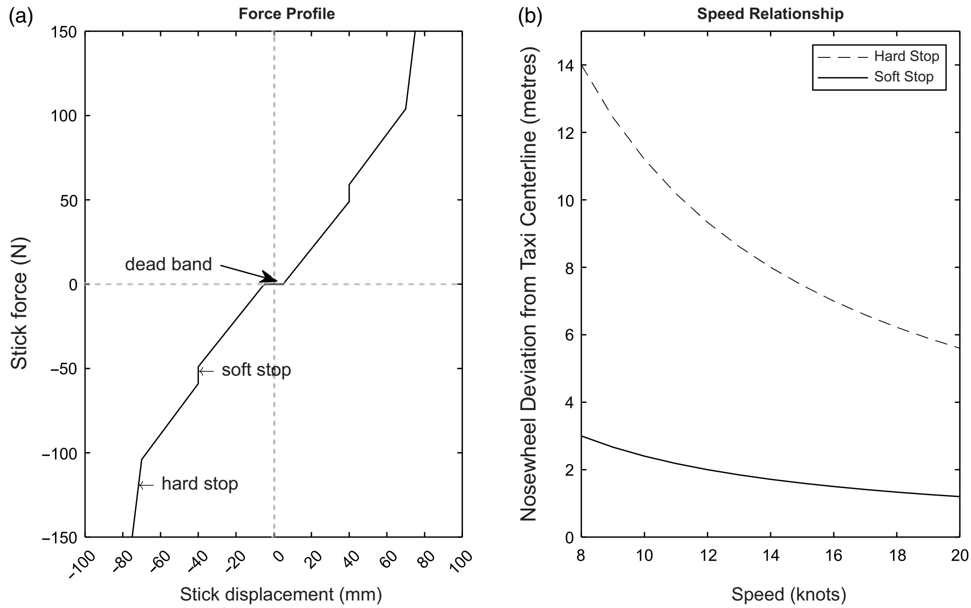

Figure 3. *Proprioceptive guidance forces: (a) force profile; (b) speed relationship.*

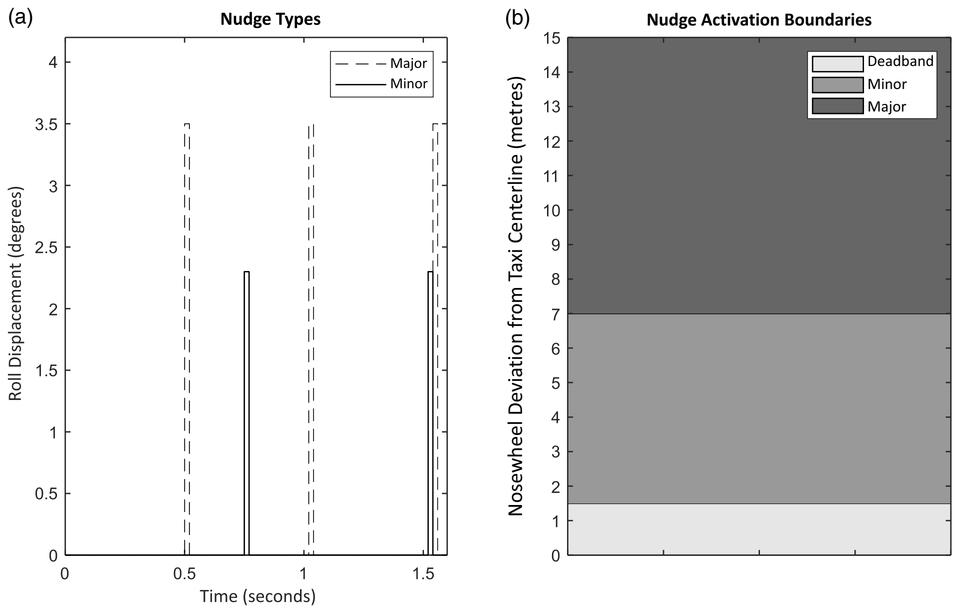

Figure 4. *Haptic route deviation warning forces: (a) nudge types; (b) nudge activation boundaries.*

an example). A further two gate-to-runway routes were created by reversing the above runway-to-gate routes (four total route alternatives).

The taxi trial lasted approximately 2.5–3h. Participants began the trial by receiving a video briefing to introduce them to the novel visual and haptic interface features. This was followed by a practice session in the simulator that allowed the participants to familiarise themselves with the simulator layout, the aircraft's manoeuvring capabilities, and the interface feedback behaviours. All participants completed

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

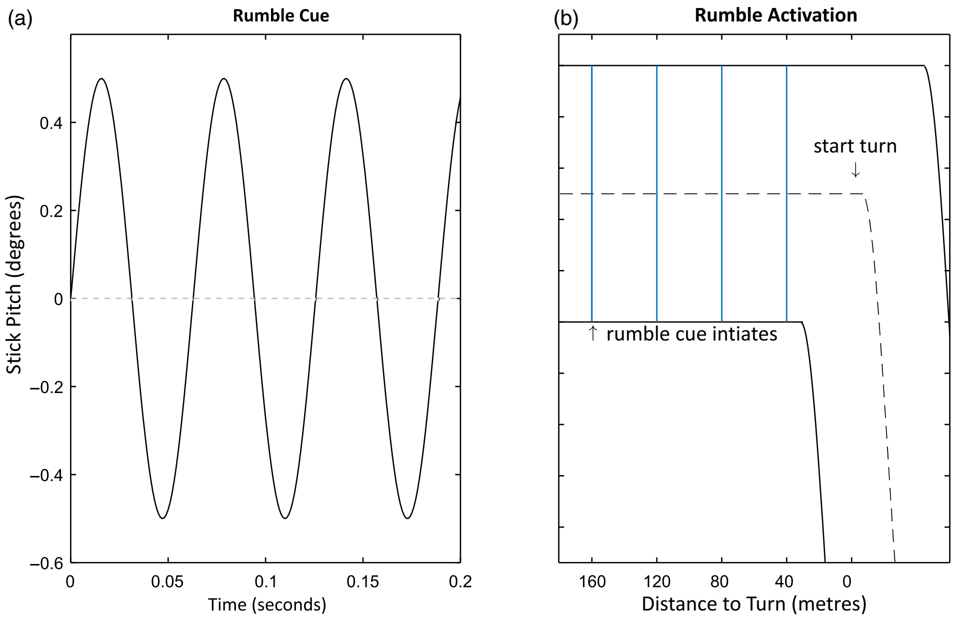

Figure 5. *Haptic turn warning forces: (a) rumble cue; (b) activation boundaries.*

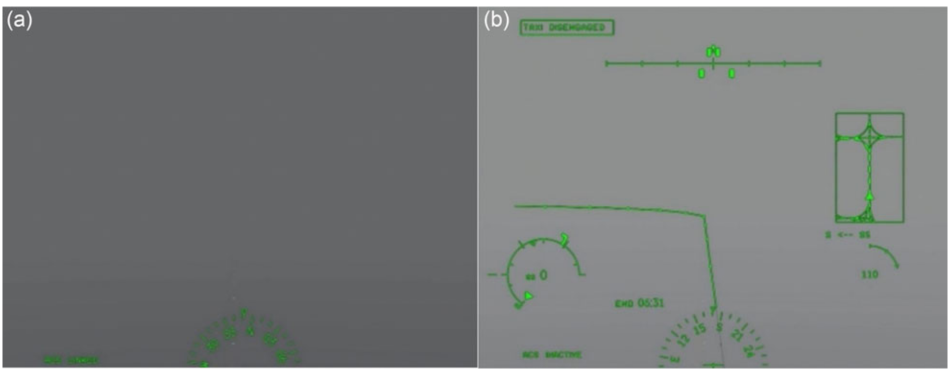

Figure 6. *Examples of the head-up symbology showing the CAT III conditions: (a) Without the designed taxi navigation symbology with traditional head-up symbology compass rose; (b) With the taxi navigation symbology.*

six taxiing trials to evaluate two experimental factors: HUD symbology presence (two levels: on/off (Fig. 6)) and nose wheel steering input mode (three levels: tiller/side stick/active stick). Three levels were chosen for the input mode factor to: (1) evaluate participants' transition from traditional tiller controls to the side stick (without haptic cues); and (2), to evaluate the benefit of the haptic cues. To differentiate between the two stick input conditions we denoted conditions with and without the haptics as active stick and side stick steering conditions, respectively.

The study followed a 2 × 3 repeated measures design with the order of conditions counterbalanced between participants. After each trial, participants completed a post-trial workload scale. A debrief usability questionnaire was completed at the end of the study.

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

2.7 Outcome measurements

2.7.1 Performance

Performance variables of interest were: (1) the frequency of major taxiway deviations, defined by occurrences where participants unintentionally departed the assigned route; (2) participant main gear lateral deviation from the taxi centreline in metres (MG); and (3) ground speed in Kn (GS). As well as comparing across HUD and input mode conditions, performance metrics were evaluated across four different route segment types: straights, 90-degree turn, 120-degree turn, and S-bend turn (chicane) segments.

2.7.2 Post-trial subjective workload questionnaire

Participants completed the NASA Task Load Index (TLX) [23] after every taxi trial. The TLX is a long-established scale designed to capture subjective workload ratings across six workload dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. Each workload dimension is measured on a scale from 0 to 20 [24],where higher ratings represent higher subjective workload. In the current study, raw workload scores from the six TLX dimensions were analysed instead of using the alternative dimension weighting method. This abridged approach was chosen due to the inconclusive evidence that dimension weighting improves the TLX's sensitivity [24].

2.7.3 Usability debrief questionnaire

After completing all six taxi trials participants completed structured debriefing questionnaire in order to acquire end-user feedback on the different interface elements that they experienced in the simulation (see Appendix B). Participant gave feedback using a seven-point Likert response scale. The questionnaire asked participants to rate the effectiveness of the three cue modality modes (visual alone, haptic alone, and multimodal) in supporting participants with the taxiing task. Open-ended questions were included that provided participants with space to expand upon the effectiveness ratings they assigned to the different modality modes. Furthermore, open-ended questions were provided that asked participants to describe their experience with using a side stick to taxi instead of a tiller.

2.8 Data analysis

Data was analysed using IBM SPSS™ 26. Parametric taxi performance (e.g. centreline deviation, ground speed) and subjective scale data (e.g. NASA-TLX) were analysed using repeated-measures general linear model ANOVA techniques. Performance main effects were the presence of the HUD guidance (*HUD* - two levels: off/on), the steering input mode (*Input* – three levels: tiller/side stick/active stick), and taxi route segment (*Route* – four levels: straight/90-degree turn/120-degree turn/s-bend turn) were investigated. NASA-TLX subjective workload main effects were *HUD*, *Input*, and *Dimension* (the six dimensions of the TLX). Significant main effects and interactions were confirmed using an alpha level of 0.05. Where appropriate, post-hoc pairwise comparisons were conducted using a Bonferroni adjustment.

Preliminary ANOVA and correlation analyses were applied to determine the respective extraneous influence that pilot aircraft type experience and years flight experience might exert upon the performance and workload outcome variables. The analysis found no significant differences or relationships ($p > 0.05$) between any of the pilot demographics and study's primary outcomes. Details of the demographic analysis are included in Appendix C.

Trial data that was not suitable to factorial analysis or that did not meet parametric requirements were subjected to traditional non-parametric hypothesis testing methods. This was the case for the frequency of major taxiway deviations and usability questionnaire responses.

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

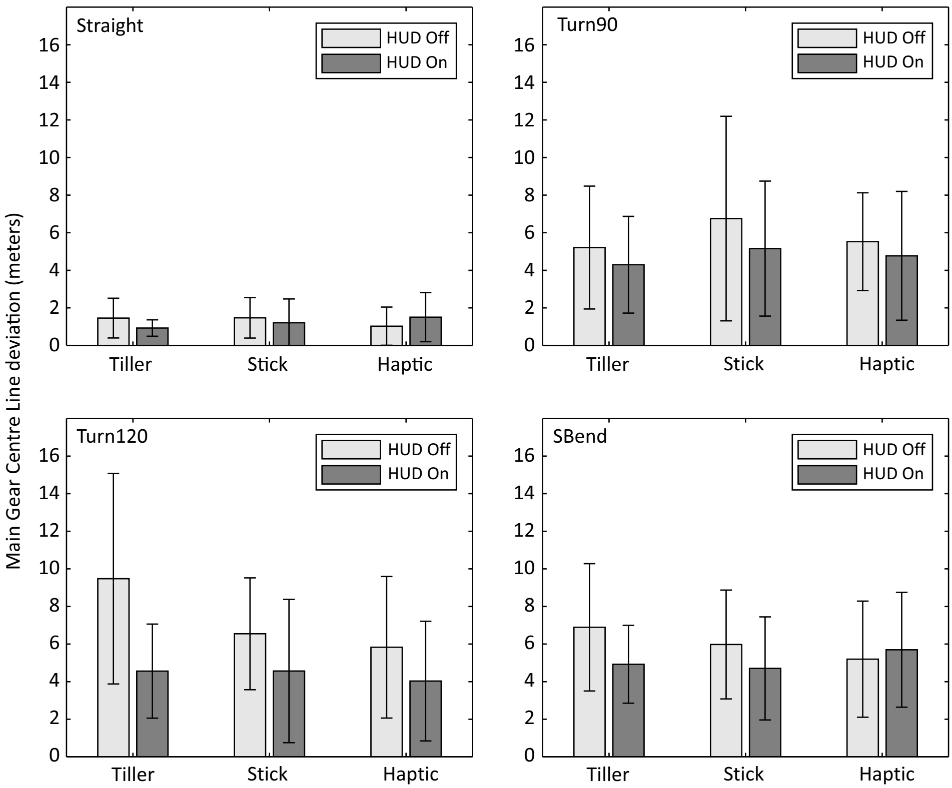

Figure 7. *Mean MG centre line deviation RMSE grouped by runway segment, input mode, and presence of the taxi navigation symbology. MG RMSE standard deviations are shown as error bars.*

3.0 Results

3.1 Major taxiway deviations

Across the 288 completed taxi route segments a total of 25 major centre line deviations that resulted in the participant vacating the edge of their current taxiway were committed. Fewer deviations occurred with the HUD guidance ($N = 7$) compared to without ($N = 18$). Across input conductions, deviations with the tiller were the most frequent ($N = 13$), side and active stick deviations were similar ($N = 5$ and 7, respectively). Statistical analysis was treated non-parametrically; differences between HUD conditions were significant (Wilcoxon's signed-rank test: $T = 2$, $z = -2.06$, $p < 0.05$). Differences between input conditions were not significant (Friedman's test: $X^2_F(2) = 2.17$, $p = 0.34$).

3.2 Main gear deviation (MG)

Mean RMSE MG route deviation performance, with standard deviations, is presented in Fig. 7. RMSE MG data from route segments where participants committed a substantial taxiway deviation were omitted from the data. The presence of the HUD reduced centre line deviation by a mean of 1.26m. Across the input modes participants were most accurate when taxiing with the active stick (Mean/SD RMSE deviation = 4.20/3.26), followed by the side stick (Mean/SD RMSE deviation = 4.53/3.67), with the largest deviations found when taxiing with the tiller (Mean/SD RMSE deviation = 4.72/3.87). For route segments, as expected, deviations were larger on the turns compared to the straight segments.

The two-way repeated-measures ANOVA for RMSE MG revealed a significant *HUD* by *Route* one-way interaction ($F(3, 33) = 4.59$, $p < 0.01$, $\eta_P^2 = 0.30$). Examination of the post-hoc pair-wise contrasts

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

explained that the significant interaction was due to the taxi HUD guidance exclusively reducing the MG RMSE on the 120-degree turns by 2.90 (95% CI: 0.86 – 4.94; $p < 0.01$) metres. In addition, there was a near significant main effect for *HUD* ($F(1, 11) = 4.79$, $p = 0.051$, $\eta_P{}^2 = 0.30$). Medium effect sizes were found for an interaction between *Input* and *Route* ($\eta_P{}^2 = 0.12$), and for the main effect of the *Input* ($\eta_P{}^2 = 0.07$). However, neither of these effects were significant and nor was the two-way interaction of *HUD*, *Input* and *Route*.

3.3 Ground speed (GS)

Overall, there were marginal differences between taxi GS between conditions with the HUD guidance (Mean/SD GS = 8.66/3.29) and without (Mean/SD GS = 8.35/2.89), irrespective of route segment type. In terms of input mode, GS was largely equivalent between the tiller and side stick input conditions, with lower GS on with the active stick (Mean/SD GS = 8.22/3.06). The complete set of GS means and standard deviations, grouped by *HUD, Input,* and *Route* are included in Table 1. The repeated-measures ANOVA for GS revealed no significant interactions or main effects for either *HUD* or *Input*.

3.4 Subjective workload (NASA TLX)

Participants self-reported their workload experience on each taxi trial using the NASA-TLX (scale 0–20). Overall, there appeared to be a clear effect of presenting participants with the HUD guidance in reducing self-reported workload (Mean TLX difference = 3.23 points; see Fig. 8). Across the input mode conditions self-reported workload was similar across the conditions (tiller = 8.02 (SD: 5.15); side stick = 8.25 (SD: 5.30), active stick = 8.22 (SD: 5.13). The full set of NASA-TLX means and standard deviations are included within Table 2. These observations were supported by the ANOVA results that demonstrated a large significant *HUD* main effect ($F(1, 11) = 35.075$, $p < 0.001$, $\eta_P{}^2 = 0.73$). The input mode for nose wheel steering also had an impact on pilots' self-reported workload, as reflected in the presence of a significant *Input* by *Dimension* interaction ($F(10, 130) = 2.08$, $p < 0.05$). There was a medium effect size for this interaction ($\eta_P{}^2 = 0.13$). Pairwise comparisons showed that the interaction was due to active stick trials producing lower TLX scores compared to the tiller conditions (Mean difference: 2.03; CI: −0.06 – 4.09) within the Effort dimension of the TLX. Though this comparison was only nearly significant ($p = 0.06$), thus the *Input* by *Dimension* interaction should be interpreted with caution.

3.5 Usability debrief questionnaire

Figure 9 displays the mean overall effectiveness ratings that participants gave to the three possible guidance cue modality combinations. Both unimodal cue presentations, HUD visual or active stick haptic guidance, were deemed to be of benefit in taxiing in low-visibility conditions. Participants strongly believed that the visual cues were effective (Mean/SD = +2.38/0.74). The haptic cues were believed to be less so (Mean/SD = +1.00/1.85), but notably the between participant variance here was relatively wide. Ratings of the multimodal cues were judged to be equally effective to cues presented via the visual channel alone (Mean/SD = +2.31/0.75). No significant difference was observed between these three ratings (Friedman's test: $X^2{}_F(2) = 4.87$, $p = 0.08$).

Participants' open answer responses on the questionnaire helped to describe the reported effectiveness of the HUD symbology. For example, one participant notes: "*When it is there it is 'Golden guidance' and helped me taxi more effectively and more safely.*" or "*It definitely supports my performance against not having a head-up display.*" As per the above usability scores, comments on the benefits of the active stick haptic feedback were more varied. One participant appraised the benefit of the haptic cues as "*Not that much I would say*" whilst another mentioned "*It definitely provided benefit over not having it since it provided feedback. Less effective compared to the HUD Symbology, since the visual stimulation is*

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

Table 1. *Mean taxi ground speed (GS) grouped by presence of HUD, nose wheel steering input mode, and taxiway route segment. Ground speed (GS) standard deviations shown in parentheses*

Input condition	Tiller				Side stick				Active stick					
HUD condition	Off		On		Off		On		Off		On		TOTAL	
Straight	9.21	(1.89)	11.84	(3.21)	11.60	(3.00)	11.09	(2.04)	10.85	(2.93)	11.92	(2.85)	11.08	(2.65)
Turn90	8.63	(2.61)	8.76	(2.87)	8.37	(3.27)	8.22	(2.85)	7.59	(2.21)	8.49	(2.95)	8.34	(2.79)
Turn120	7.98	(1.83)	8.17	(3.90)	7.31	(2.75)	6.99	(3.12)	7.30	(2.58)	6.58	(1.48)	7.39	(2.61)
SBend	6.61	(1.18)	8.26	(3.40)	8.36	(3.62)	7.14	(2.51)	6.55	(2.43)	6.41	(1.51)	7.22	(2.44)
TOTAL	8.11	(1.88)	9.26	(3.34)	8.91	(3.16)	8.36	(2.63)	8.07	(2.54)	8.35	(2.20)		

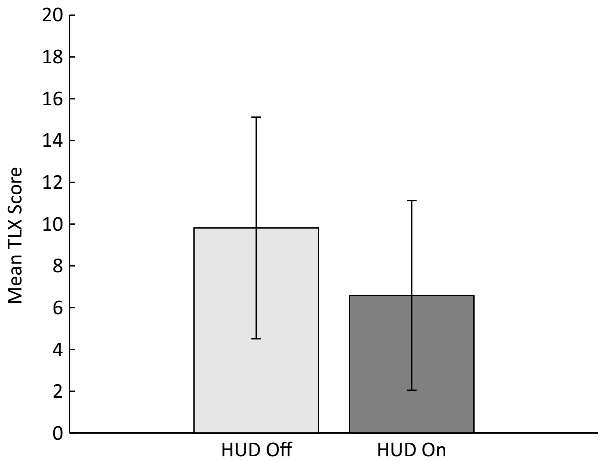

Figure 8. *Mean NASA Task Load Index scores for HUD conditions. Standard deviations represented as errors bars.*

significantly more effective." A number of participants mentioned that the vibrotactile distance to turn cue (the rumble strip) was the only haptic cue they attended to: *"I didn't notice any of these nudges apart from when I was off the runway centre line at the end."* and "*The only haptic cue you are really paying attention to is the rumble strips to notify me when a turn was coming up*." A possible reason for this is that the proprioceptive cues were conditional cues based on ground speed and centreline deviation, so it is possible that some participants may not have encountered them. For instance: "*I don't feel like you gave this cue since I was taxiing so slowly in this task*." The multimodal nature in which haptic and visual cues were utilised in the current study were exemplified by: "*These were really great! Particularly without the HUD [the haptics were] really useful to have. But when I had it with the HUD, I used it to cross check the HUD symbology*." Likewise, the benefit that the haptic cues provided as a situational awareness enhancing redundant cue was highlighted by a comment: "*In this way the haptic cue served as an attention getter. I certainly didn't feel I needed the haptics in the middle of a turn, but I did when I've been taxiing in a straight line to help get my attention back on to the task*." Therefore, the utility of the haptic cues changed according to the presence of the visual HUD cues.

Finally, there were a large number of remarks from participants on how quickly they found they could transition from the tiller to the side stick as means of nose wheel steering during taxiing. For example, one participant compared the transition to their experience moving from a yoke to a sidestick during their Boeing to an Airbus type transition: *"If I had to transition from a tiller to a stick, it wouldn't be a big leap. When I had to transition from a Boeing to Airbus everyone made a deal about going to a stick, but it only took me about 30 seconds to get used to that. And I can imagine that this would be just the same as that."* Similar comments included: "*Actually, I thought using a stick would be worse than it was. The fact is - it didn't take me long to get used to using it."*

4.0 Discussion

In the current study the benefits of a user-centred design based multimodal taxi symbology set were evaluated in a human-in-the-loop taxi simulator trial. The study presents the first examination of how a multimodal display, combining visual and haptic cues, could support the safety and efficiency in which pilots are able to conduct a taxi navigation task in low-visibility conditions. The study is also unique in that it explores the feasibility of pilots transitioning from tiller to side stick as the mode of nose wheel steering. This concept would facilitate the future application of task-relevant haptic information within the taxiing task, given the greater versatility of potential haptic cues that can be delivered via the side

Table 2. *Mean inter-item NASA-TLX score grouped by presence of the HUD and nose wheel steering input mode. NASA-TLX score standard deviations shown in parentheses*

Input condition	Tiller				Side stick				Active stick					
HUD condition	Off		On		Off		On		Off		On		TOTAL	
1 - Mental	12.13	(5.65)	10.13	(4.36)	14.64	(3.69)	9.17	(4.76)	13.13	(4.28)	9.50	(4.42)	11.45	(4.53)
2 – Physical	7.13	(4.16)	4.83	(4.24)	7.32	(4.53)	3.79	(3.58)	7.38	(5.42)	3.75	(3.55)	5.70	(4.24)
3 – Temporal	7.79	(4.94)	4.42	(3.50)	7.82	(5.38)	4.96	(3.95)	6.08	(4.02)	4.25	(3.17)	5.89	(4.16)
4 – Performance	8.17	(4.91)	6.71	(5.15)	7.55	(4.69)	5.79	(3.10)	8.33	(4.34)	4.96	(2.91)	6.92	(4.18)
5 – Effort	12.46	(4.98)	8.54	(4.02)	14.59	(3.31)	10.13	(3.89)	15.13	(2.99)	9.75	(4.14)	11.77	(3.89)
6 - Frustration	8.50	(5.43)	5.46	(4.50)	8.30	(4.66)	6.08	(5.78)	10.27	(4.69)	6.38	(3.73)	7.50	(4.80)
TOTAL	9.36	(5.01)	6.68	(4.30)	10.03	(4.38)	6.65	(4.17)	10.05	(4.29)	6.43	(3.65)		

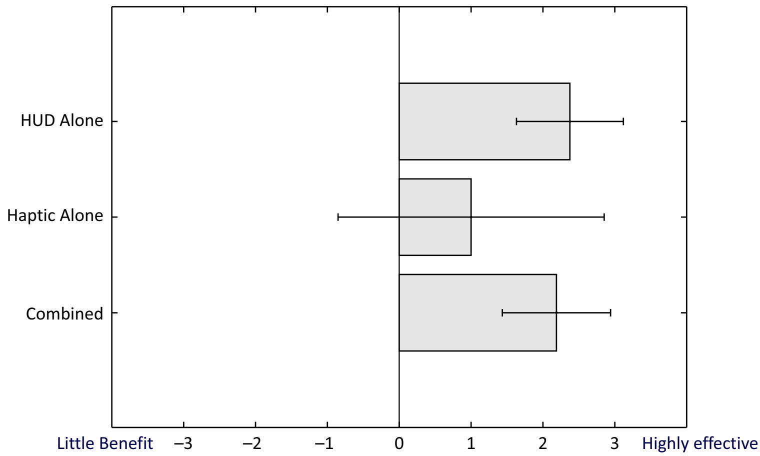

Figure 9. *Mean overall HUD taxi, haptic and combined guidance effectiveness. Rating standard deviations represented as error bars.*

stick compared to the tiller. The primary findings of the study are that pilot participants perceived the HUD interface to be highly intuitive and that it would have substantial benefit in improving the safety and efficiency of real-world taxiing operations. Participant feedback on the unimodal presentation of haptic cues was also positive, but less so relative to task-related visual information presented via the HUD. Furthermore, participant feedback on the utilisation of the side stick for nose wheel steering revealed that, whilst there was a preference for the inherently more familiar tiller, the perceived transition time from tiller to side stick was brief. Hence, this lends weight to the feasibility of the side stick nose wheel steering concept. In conjunction with the positive participant feedback, the objective evaluation of the HUD guidance interface was supported by participants' views on the effectiveness of the symbology. Overall, the findings of the study underline the benefits of adopting a user-centred design from early on in the design process of complex aviation systems/displays. Specifically, incorporating the end user in the design process results in the development of systems/displays that are received with high acceptance and perceived benefit by the end user.

Despite mixed feedback regarding the haptic cues, many participants praised the safety benefits they offered when presented as redundant cues alongside the visual HUD cues within the multimodal concept. Namely, the effect of cue redundancy was reported to increase pilot's trust in the information they were receiving via the HUD. Similar relationships between the presentation of multimodal information and trust have been reported recently in the domain of automation design [25]; there is a relationship between perceived trust and the clarity of a user's mental model, and that this mental model is clearer in multimodal compared to unimodal conditions. Another possibility is that if the interpretation of incoming multimodal information is congruent it will support the current situational assessment (positive weighting). As a result the summative content of the incoming information becomes more trusted, and the decision to act upon the information is both more likely and less cognitive resource intensive. This rationale is supported by multiple decision-making models [26–28]. Further safety benefits of the multimodal interface were the diminished opportunity for attentional capture presented by the conformal HUD symbology. For example, one participant remarked that the haptics acted as an effective "attention getter" during long straights to notify pilots of upcoming turns when arousal might have declined. Implementation of redundant cues from basic and applied research have reported similar effects [29–35]. Here, redundant cues are able to bestow a pre-attentive advantage to neglected information, enabling the neglected information to "pop-out" of the display. It is possible that future multimodal studies involving the detection of non-nominal events, such as a taxi-way incursions, will benefit from examining the interaction between visual and redundantly presented haptic cues.

The objective taxiing performance results revealed several important findings. Firstly, when taxiing with the HUD visual guidance in CAT-III visibility conditions participants made significantly fewer major taxiway deviations, either going off-route or departing the taxiway tarmac. In parallel, the second performance enhancement reported in the current study was that the presentation of the HUD guidance reduced main wheel centre line deviation during turns. These outcomes corroborate the findings of Arthur *et al.* [36] who found that the frequency of major navigation errors was higher when navigating with traditional paper charts compared to either a HUD or HMD. However, their finding was not significant, likely due to limited statistical power. Likewise, Arthur *et al.* reported that RMSE centreline deviations were greater whilst navigating with paper charts. In contrast, participants taxied slightly, albeit significantly, faster (1.2Kn faster) within the study by Arthur *et al.* No difference in taxiing speed were observed in the current study which is likely due to the more restricted visibility conditions employed in the current study (CAT III in the current study versus 700 RVR in Arthur et al). Thus, it is possible that the effect of the HUD guidance on taxiing speed might be exclusive to taxiing in better visibility conditions, such as CAVOK.

The clear performance benefits granted by the HUD navigational symbology were accompanied by a reduction in self-reported workload measured with the NASA TLX. This is promising in the sense that the HUD navigational symbology provides improved aircraft ground manoeuvring performance at a reduced workload 'cost'.

Evidence that side stick haptics (i.e. the active side stick) objectively benefited taxi performance, either as a unimodal or multimodal cue, was not found in the current study. However, several of the haptic related performance differences were associated with medium effect sizes ($\eta_P{}^2 > 0.06$), hence the comparisons that included haptics would have likely achieved significance with a larger sample size. Furthermore, the lack of a significant effect of haptics upon performance does conflict with a number of studies that have demonstrated the benefit of haptic cues in minimising spatial deviations during path following tasks [19–21, 37, 38]. Though the current findings do align with several studies that have found that, whilst multiple haptic concepts show great potential, many lack clear conclusive evidence [39, 40]. There are several potential reasons for the conflicting findings in the literature. Firstly, a meta-analysis on vibrotactile research [16] showed that the performance benefits of haptic cues is moderated by cue complexity. For example, haptic cues have been shown to be effective when used as alerting cues but not when used as directional cues. Likewise, the effectiveness of haptic cues will be highly dependent on the intuitiveness of the information contained. Studies where an effect of haptics been absent often communicate task-related information with discrete arbitrary signals that bear minimal resemblance to users' existing mental models. For instance, a recent study by Van Baelen *et al.* [41] in 2021 found no performance benefit associated with the representation of departing a safe flight envelope through the presentation of a discrete saw-tooth vibrotactile signal. In the current study, this observation was echoed within participant comments on the lack of intuitiveness of the lateral proprioceptive cues relative to the distance to turn vibrotactile cue. Specifically, the intuitiveness of the distance to turn vibrotactile cue was bolstered by accommodating participant's existing mental model of the haptic sensation experienced whilst driving over rumble strips on approach to a round-about in an automobile. In addition, comments by a few participants revealed that there was some confusion between two lateral proprioceptive cues that were implemented – a variable lateral force (linked to speed) and a lateral nudge to protect against nose wheel scrubbing and centreline deviation, respectively. Furthermore, unlike the other haptic cues, the information presented through the variable lateral force cue was not replicated visually on the HUD, undermining it's benefit as a redundant cue during multimodal trial conditions. Indeed, one participant indirectly suggested such an inclusion by showing speed limits on the ground speed dial. It is possible adding this information visually could alleviate some of the confusion participants experienced with the haptic cue. Finally, in the current study the concept of nose wheel scrubbing was difficult to communicate to the participants due to the fixed based simulator. Motion-base simulations are far more successful at presenting this information to participants. Hence, participants may require a demonstration of the cues within a motion-based simulator in conjunction with an extended period of training time.

Another potential explanation for the conflicting haptic performance findings in the literature could be due to task-difficulty differences in the scenarios employed across studies. Whilst the CAT-III visibility conditions exacerbated the navigational complexity of the current taxi task, a scenario also comprising other non-nominal events (e.g. taxiway/runway incursion of other aircraft) could potentially reveal objective navigational benefits. This proposal is supported by Wickens' multiple resource theory (MRT) [12, 13]. In particular, Wickens posits several principles where the presence of multimodal cue redundancy will be most advantageous: (1) when overall task load is high, there will be greater offloading to ease the high workload; (2) cue overlap on subtasks (e.g. visual scanning and haptic direction cues effectively combine to support a common task, such as target detection or navigation). Both principles are of particular relevance to navigation tasks conducted in low-visibility conditions. Future research examining higher task load taxiing scenarios would be desirable.

5.0 Conclusion

In conclusion, the findings of the study demonstrate how a multimodal display, combining visual and haptic cues, could support the safety and efficiency in which pilots are able to conduct a taxi navigation task in low-visibility conditions. Furthermore, the current findings exemplify the benefits of adopting a user-centred design from early in the design process of complex systems and displays, whereby the approach resulted in a novel multimodal display with a demonstratable performance benefit and accompanying high end user acceptance. Future multimodal research should examine the implementation of haptic cues within a design grounded in UCD to enhance cue intuitiveness, thus optimising pilots' ability to accommodate their meanings more effectively within existing mental model frameworks.

Acknowledgements. This paper is based on work performed in the Open Flight Deck project, which has received funding from the ATI Programme, a joint government and industry investment to maintain and grow the UK's competitive position in civil aerospace design and manufacture. The programme, delivered through a partnership between the Aerospace Technology Institute (ATI), Department for Business, Energy & Industrial Strategy (BEIS) and Innovate UK, addresses technology, capability and supply chain challenges.

References

[1] Blundell, J., Huddlestone, J., Scott, S. and Sears, R. Flight deck optimization for a future SESAR/NextGen operating environment, In *AHFE 2019*, 2019, p 786. doi: 10.1007/978-3-319-93885-1_13

[2] Arthur, J., Prinzel, L.J., Bailey, R.E., Shelton K.J. and Nicholas, S.N. Evaluation of head-worn and head-up displays for use in enhanced flight vision system operations, *Opt Eng*, 2018, **58**, (5), doi: 10.1117/1.OE.58.5.051802

[3] Arthur, J., et al. Review of head-worn displays for the Next Generation Air Transportation System, *Opt Eng,* 2017, **56**, (5), doi: 10.1117/1.oe.56.5.051405

[4] Borst C., Mulder M. and Van Paassen, M.M. Design and simulator evaluation of an ecological synthetic vision display, *J Guid Cont Dynam*, 2010, **33**, (5), pp 1577–1591. doi: 10.2514/1.47832

[5] Mulder, J.A., Van Paassen, M.M. and Mulder, M. Perspective guidance displays show the way ahead, *J Aerosp Comput Inf Commun*, 2004, pp 428–431. doi: 10.2514/1.14069

[6] Stanton, N.A., Plant, K.L., Roberts, A.P. and Allison, C.K. Use of Highways in the sky and a virtual pad for landing head up display symbology to enable improved helicopter pilots situation awareness and workload in degraded visual conditions, *Ergonomics,* Feb 2019, **62**, (2), pp 255–267. doi: 10.1080/00140139.2017.1414301

[7] Stanton, N.A., Plant, K.L., Roberts, A.P., Allison, C.K. and Harvey, C. The virtual landing pad: facilitating rotary-wing landing operations in degraded visual environments, *Cognit Technol Work,* 2018, **20**, (2), pp 219–232. doi: 10.1007/s10111-018-0467-1

[8] Arthur, J. et al. Design and testing of an unlimited field-of-regard synthetic vision head-worn display for commercial aircraft surface operations, In *Proceedings of SPIE - The International Society for Optical Engineering*, 2007. doi: 10.1117/12.719695

[9] R. Bailey, Arthur, J., Prinzel III, L. and Kramer, L. Evaluation of head-worn display concepts for commercial aircraft taxi operations, In *Head- and Helmet-Mounted Displays XII: Design and Applications*, vol. 6557, (65570Y), 2007. doi: 10.1117/12.717221

[10] Arthur, J., et al. Performance comparison between a head-worn display system and a head-up display for low visibility commercial operations, In *Proceedings of SPIE*, p 9086, 2014.

[11] Ververs, P.M. and Wickens, C.D. The effect of clutter and lowlighting symbology on pilot performance with head-up displays, *Proc Human Factors Ergonom Soc Ann Meet*, 1996, **40**, (2), pp 62–66. doi: 10.1177/154193129604000212

[12] Wickens, C.D. Multiple resources and mental workload, *Hum Factors*, 2008, **50**, (3), pp 449–455. doi: 10.1518/001872008X288394

[13] Wickens, C.D. Multiple resources and performance prediction, *Theor Issues Ergon Sci* 2002, **3**, (2), pp 159–177. doi: 10.1080/14639220210123806

[14] Bolopion, A. and Régnier, S. A review of haptic feedback teleoperation systems for micromanipulation and microassembly, *IEE Trans Autom Sci Eng,* July 2013, **10**, (3), pp 496–502. doi: 10.1109/TASE.2013.2245122

[15] Jones L.A. and Sarter, N.B. Tactile displays: Guidance for their design and application, *Hum Factors,* Feb 2008, **50**, (1), pp 90–111. doi: 10.1518/001872008X250638

[16] Prewett, M.S., Elliott, L.R., Walvoord, A.G. and Coovert, M.D. A meta-analysis of vibrotactile and visual information displays for improving task performance, *IEE Trans Syst Man Cybern Part C Appl Rev,* Jan 2012, **42**, (1), pp 123–132. doi: 10.1109/TSMCC.2010.2103057

[17] Overtoom, E.M., Horeman, T., Jansen, F.W, Dankelman, J. and Schreuder, H.W.R. Haptic feedback, force feedback, and force-sensing in simulation training for laparoscopy: A systematic overview, *J Surg Educ,* Jan 2019, **76**, (1), pp 242–261. doi: 10.1016/j.jsurg.2018.06.008

[18] Ellerbroek, J., Martin, M.J.M.R, Lombaerts, T., van Paassen, M.M. and Mulder M. Design and evaluation of a Flight Envelope Protection haptic feedback system, *IFAC-PapersOnLine*, 2016, **49**, (19), pp 171–176. doi: 10.1016/j.ifacol.2016.10.481

[19] Lam, T.M., Mulder, M. and Van Paassen, M.M. Haptic interface for UAV collision avoidance, *Int J Aviat Psychol*, 2007, **17**, (2), pp 167–195. doi: 10.1080/10508410701328649

[20] Ellerbroek, J., Brantegem, K., Van Paassen, M.M., De Gelder, N. and Mulder, M. Experimental evaluation of a coplanar airborne separation display, *IEEE Trans Human-Mach Syst*, May 2013, **43**, (3), pp 290–301. doi: 10.1109/TSMC.2013.2238925

[21] Mulder, M., Abbink, D.A. and Boer, E.R. Sharing control with haptics: seamless driver support from manual to automatic control, *Human Factors*, Oct 2012, **54**, (5), pp 786–798. doi: 10.1177/0018720812443984

[22] Harris, D. *Human Performance on the Flight Deck*, Ashgate, UK, 2011.

[23] Hart, S.G. and Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research Sandra, *Adv Psychol*, 1988, **52**, pp 139–183.

[24] Hart, S.G. Nasa-Task Load Index (NASA-TLX); 20 years later, *Proc Human Factors Ergonom Soc Ann Meet*, 2006, **50**, (9), pp 904–908.

[25] Hastie, H., Liu, X. and Patron, P. Trust triggers for multimodal command and control interfaces, In *The 19th ACM International Conference on Multimodal Interaction*, Nov. 2017, pp 261–268. doi: 10.1145/3136755.3136764

[26] Rasmussen, J. Signs and symbols, and other distinctions in human performance models, *IEEE Trans Syst Man Cybern*, 1983, **13**, (3), pp 257–266.

[27] Peterson, D.S., King, L.A., Cohen, R.G. and Horak, F.B. Cognitive contributions to freezing of gait in parkinson disease: Implications for physical rehabilitation, *Physical Therapy,* 2016, **96**, (5), pp 659–670. doi: 10.2522/ptj.20140603

[28] Klein, G.A. A recognition-primed decision (RPD) model of rapid decision making, In *Decision Making in Action: Models and Methods,* Jan 1993, pp 138–147.

[29] Blundell, J., Frisson, S., Chakrapani, A., Kearney, S., Vijay, S., MacDonald, A., Gissen, P., Hendriksz, C. and Olson, A. Markers of cognitive function in individuals with metabolic disease: Morquio syndrome and tyrosinemia type III, *Cogn Neuropsychol*, May 2018, **35**, (3–4), pp 120–147. doi: 10.1080/02643294.2018.1443913

[30] Blundell, J., Scott, S., Harris, D., Huddlestone, J. and Richards, D. Workload benefits of colour coded head-up flight symbology during high workload flight. *Displays*, Dec 2020, **65**. doi: 10.1016/j.displa.2020.101973

[31] Treisman, A.M. and Gelade, G. A feature-integration theory of attention, *Cogn Psychol*, 1980, **12**, pp 97–136. doi: 10.1016/0010-0285(80)90005-5

[32] Leonards, U., Sunaert, S., Van Hecke, P. and Orban, G.A. Attention mechanisms in visual search – An fMRI study, *J Cogn Neurosci*, Jan 2000, **12**, (2), pp 61–75. doi: 10.1162/089892900564073

[33] Wolfe, J.M. *Seeing*, Academic Press, San Diego, CA, USA, 2000.

[34] Nakayama K. and Silverman, G. Serial and parallel processing of visual feature conjunctions, *Nature*, 1986, **320**, (20).

[35] Blundell, J., Scott, S., Harris, D., Huddlestone, J. and Richards, D. With flying colours: Pilot performance with colour-coded head-up flight symbology, *Displays*, *Jan 2020, 61.* doi: 10.1016/j.displa.2019.101932

[36] Arthur, J., et al. Synthetic vision enhanced surface operations with head-worn display for commercial aircraft, *Int J Aviat Psychol*, Mar 2009, **19**, (2), pp 158–181. doi: 10.1080/10508410902766507

[37] Gilson, R.D., Redden, E.S. and Elliott, L.R. *Remote Tactile Displays for Future Soldiers*, Defence Technical Information Centre, 2007.

[38] Van Erp, J.B. Tactile Displays for Navigation and Orientation: Perception and Behaviour, NATO, RTO Technical Report, TR-HFM-122, 2007.

[39] Lombaerts, T., Looye, G., Ellerbroek, J. and Martin, M.R. Design and piloted simulator evaluation of adaptive safe flight envelope protection algorithm, *J Guid Control, Dynam*, May 2017, **40**, (8), pp 1902–1924. doi: 10.2514/1.G002525

[40] Stepanyan, V., et al. Loss-of-control mitigation via predictive cuing, *J Guid Control Dynam*, Aug 2017, **40**, (4), pp 831–846. doi: 10.2514/1.G001731

[41] van Baelen, D., van Paassen, M.M., Ellerbroek, J., Abbink, D.A. and Mulder, M. Flying by feeling: Communicating flight envelope protection through Haptic feedback, *Int J Human-Comput Interact*, 2021, **37**, (7), pp 655–665. doi: 10.1080/10447318.2021.1890489

Appendix

Appendix A. EDDM taxi route. Example of one gate-runway routes from gate 310 to runway B2.

Appendix B. Usability Questionnaire

Date:______________

Participant ID:______________

You have just used the head-up-display (HUD) along with active stick haptic feedback to complete taxiing phase in different modes of operation. While we have collected data during your performance on the task, we would now like to get your views on the various technology features that you have just experienced.

We would like to understand to what extent the candidates believe the new HUD and stick capabilities would enable them to conduct their duties in a more effective manner:

(a) HUD taxi symbology alone

Little Benefit						**Highly Effective**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(b) Stick taxi steering with tactile cueing alone

Little Benefit						**Highly Effective**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(c) Combined taxi guidance symbology with stick taxi steering and tactile cueing

Little Benefit						**Highly Effective**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

Further comments:

HUD SYMBOLOGY

Considering the visual properties and presentation of information of the HUD taxi symbology set as a whole.

(a) Do you feel the current HUD taxi symbology would offer *performance benefits* over your present daily practice (e.g. controlling speed of aircraft)?

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(b) Do you feel the current HUD taxi symbology would offer *safety benefits* over your present daily practice?

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

(c) Did you feel there are any taxi guidance HUD features that were particularly distracting and counter-intuitive to completing the task?

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(d) Do you feel you were able to find the information you required, when you needed it?

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(e) There were features of the HUD symbology that I completely ignored.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(f) The HUD taxi information architecture made sense to me.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

Further comments:

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

TAXI STEERING VIA SIDE STICK

(a) Steering with the side stick was preferable to steering with the tiller.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(b) It did not take me long to become accustom to steering the nose wheel with the side stick.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

(c) The sensitivity of the required side stick deflections to steer the nose wheel was?

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

Further comments:

OVERALL TASK AND HUD/ACTIVE STICK INTEGRATION

In this final section we'd like you to consider the overall difficulty of the task as well as the collective effectiveness of the HUD and active stick cues in supporting your task performance.

Did you find any aspect of the taxi task more difficult than other tasks? If so, which aspects:

(a) The concurrent presentation of HUD and active stick guidance information supported me in the task more so than when the HUD and active stick information was presented in isolation.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

Further comments:

(b) The concurrent presentation of HUD and active stick guidance confused me at certain points of the task.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

Further comments:

(c) There are HUD and active stick features I would have liked that are not presently here.

Strongly Disagree						**Strongly Agree**
1	**2**	**3**	**4**	**5**	**6**	**7**
O	**O**	**O**	**O**	**O**	**O**	**O**

Further comments:

Appendix C. Demographic Analysis

Pilot Aircraft Type Experience

The influence of participants' primary aircraft type experience, Airbus versus Boeing, on the performance outcomes, main gear deviation (*MG*) and ground speed (*GS*), were analysed using two-way split-plot ANOVAs. Where *Aircraft Type* (2 levels: Airbus/Boeing) and *Route* (4 levels) were treated as between- and within-subject variables, respectively. Performance differences across *Input* and *HUD* conditions were averaged across route type.

For mean MG deviation there was no main effect of *Aircraft Type* ($F(1, 10) = 1.00, p = 0.34$), nor was there an interaction between *Aircraft Type* and *Route* ($F(3, 30) = 0.13$, $p = 0.94$). Likewise, for mean *GS* there was no main effect of *Aircraft Type* (F (1, 10) = 0.24, p = 0.64), nor was there an interaction between *Aircraft Type* and *Route* (F (3, 30) = 1.23, p = 0.07).

The effect of pilot aircraft type experience on pilot subjective workload, measured on the NASA-TLX, was also investigated. TLX average workload scores between Airbus and Boeing pilots were compared using an independent-samples t-test. Results found no significant difference ($t(10) = 0.47$, $p = 0.65$).

https://doi.org/10.1017/aer.2022.81 Published online by Cambridge University Press

Pilot Years Flight Experience

The second series of pilot demographic analyses was conducted to examine the relationship between pilot flight experience (years) with task performance and workload. Correlations between flight experience (years) and main gear deviation (*MG*) and ground speed (*GS*) on the 4 different route segments were not significant (see Table A1).

Table A1. *Pilot flight experience (years) correlations with mean main gear deviation (MG) and with taxi ground speed (GS)*

	Mean main gear deviation (*MG*)		Mean ground speed (*GS*)	
	r	*p*	*r*	*p*
Straight	0.23	0.48	−0.24	0.50
Turn90	0.19	0.57	−0.06	0.87
Turn120	−0.19	0.56	0.17	0.60
SBend	0.16	0.63	−0.30	0.35

The relationship between pilot flight experience and subjective workload was also investigated using a correlation analysis. The results were not significant ($r(10) = 0.03$, $p = 0.92$)

Cite this article: Blundell J., Collins C., Sears R., Plioutsias T., Huddlestone J., Harris D., Harrison J., Kershaw A., Harrison P. and Lamb P. Low-visibility commercial ground operations: An objective and subjective evaluation of a multimodal display. *The Aeronautical Journal*, https://doi.org/10.1017/aer.2022.81