Neuroergonomics and Cognitive Engineering, Vol. 102, 2023, 76–82
https://doi.org/10.54941/ahfe1003007

Multimodal Learnability Assessment of a Touch-Based Large Area Display With Eye Tracking and Optical Brain Imaging

Ezgi Çigdem Şahin[1], Dilan Diğba Özmenoğlu[1], Ahmet Paker[1], Yasin Kaygusuz[1], Hakan Aydemir[1], Cengiz Acartürk[2], and Murat Perit Çakır[3,4]

[1]Turkish Aerospace Industries, Ankara, Türkiye
[2]Department of Cognitive Science, Jagiellonian University, Krakow, Poland
[3]Department of Cognitive Science, Informatics Institute, Middle East Technical University, Ankara, Türkiye
[4]METU TAF Modelling & Simulation R&D Center, Middle East Technical University, Ankara, Türkiye

ABSTRACT

Multifunctional Large-Area Displays (LAD) have become an integral part of modern airplane cockpits, offering pilots flexible access to flight controls and mission-critical information. The modern glass cockpit paradigm is expected to reduce the workload of pilots by reducing the complexity and clutter of traditional cockpit layouts and improving their situational awareness by providing flexible access to rich information through the interface. On the other hand, the new paradigm has led to interaction design challenges in utilizing the affordances of the novel interface components, presenting barriers against effective use of the interface by the pilots. In the present study, we report empirical investigations of such a LAD interface and its learnability via behavioural performance, established upon the neuroergonomics approaches, such as eye tracking and optical brain imaging measurements. Two test pilots who had prior experience in traditional cockpit layouts were recruited to perform flight tasks in a flight simulator incorporating a LAD touch screen to interact with the basic flight instruments of a trainer jet. The Pupil Labs Invisible mobile eye tracker and the fNIR Imager 1100 system were used to monitor the pilots' eye movements and prefrontal oxygenation changes during the flights. Two weeks later, the tasks were replicated for the learnability assessment. The results indicated that pilots' task completion times decreased for the majority of the tasks, accompanied by an increase in eye fixation durations, a decrease in the number of fixations, and a decrease in right dorsolateral and right superior prefrontal cortex activation. Overall, these results suggest supporting evidence for the learnability of the new interface paradigm through task measurements in multiple scenarios.

Keywords: Learnability assessment, Large area display, Aviation, Eye tracking, fNIRS

INTRODUCTION

Multifunctional Large-Area Displays (LAD) have become an integral part of modern airplane cockpits, offering pilots flexible access to flight controls and mission-critical information. The modern glass cockpit paradigm is expected to reduce the workload of pilots by reducing the complexity and clutter of traditional cockpit layouts and improving their situational awareness by providing flexible access to rich information through the interface. On the other hand, the new paradigm has led to interaction design challenges in utilizing the affordances of the novel interface components, presenting barriers against effective use of the interface by the pilots.

Aircraft control is a demanding task. Therefore, cockpit interface design and usability testing play a crucial role. An incompatibility in providing the relevant functionality may cause the pilot to fail the task in an emergency or even in common flight segments. A known example of such an incident is the TransAsia Airways Flight accident, in which the pilot accidentally shut down the functioning engine instead of the malfunctioning engine. This resulted in the loss of many passengers on board (Taiwan Aviation Safety Council, 2015). Recently, the main challenge is that the automatization of cockpit interfaces and the availability of functions to the pilot may ease the mental workload and task difficulty. Nevertheless, a cockpit interface may also introduce issues in user interaction due to the sensory capacities of humans, such as the limitations in controlling touch-based interfaces in recent Large-Area Displays (LADs).

In the present study, we analyse perceptual and cognitive aspects of the interaction between pilots and various cockpit interface components from a user-centric perspective, with a special emphasis on learnability. Alongside efficiency and effectiveness, learnability has been a major constituent of usability analysis and human factors for the past several decades (e.g., Booth, 1989; Shackel, 1991). More recent methodologies, such as physiological sensors and brain imaging, allow an analysis of vague concepts, such as mental workload and situational awareness in the field, beyond laboratory settings (Wickens, 2008).

In the user-centric design approach, the main goal is to close the gap between the assumptions made by the designer on behalf of the user and the expectations of the user from the system (Saeed et al., 2016). The priority is given to the individual needs of the users and their expectations at the onset of the design rather than augmenting user opinions by the end of the design.

In this study, we present a neuroergonomics perspective on the learnability of a modern cockpit interface through the analysis of behavioural performance, eye movement recordings, and fNIRS brain imaging data. We report empirical investigations of a LAD interface and its learnability as presented below.

MATERIALS & METHODS

Two test pilots, who had experience in traditional cockpit layouts but had no prior experience with the LAD display of the HURJET platform were recruited for this study. The pilots were first given standardized training to familiarize themselves with the basic information layout and features of the LAD interface (Figure 1). Following the initial training, both pilots flew a

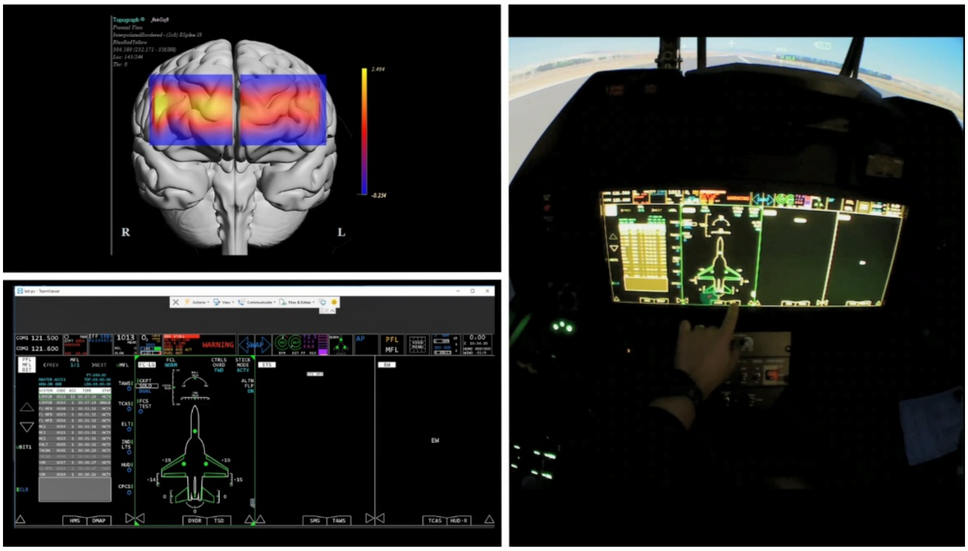

Figure 1: Screenshot from the LAD layout assignment task before the taxi phase. Video streams display the prefrontal oxygenation changes (top-left), the LAD screen feed (bottom-left), and the eye-tracking world camera view overlaid with a gaze indicator (right).

scenario in the HURJET flight simulator that included various tasks (e.g., functional layout assignment, setting communication parameters, monitoring barometric settings and fuel levels, etc.) over the LAD interface for the before take-off, taxi, take-off, climb, level flight, approach, and landing phases. The flight took about 1.5 hours. Two weeks later, the tasks were replicated for the learnability assessment, where the same pilots went through the same set of scenarios in the flight simulator.

During the trials, the Pupil Labs Invisible mobile eye tracker (120 Hz, binocular by Pupil Labs GmbH, Germany) and the fNIR Imager 1100 (fNIR Devices LLC, Potomac, MD, USA) system were used to monitor the pilots' eye movements and prefrontal oxygenation changes, respectively. The mobile eye tracker had two infrared cameras facing towards both eyes for binocular recording, as well as an integrated world camera capturing the field of view from the participant's perspective. The system has a sampling rate of 120 Hz (see Tonsen, Baumann & Tierkes, 2020 for a high-level description and performance evaluation). The Pupil Invisible Companion mobile application, accompanied by the mobile eye tracker, was used to calibrate and record eye movements.

The fNIRS system is composed of a flexible headband that hosts four LED infrared light sources and 10 photodetectors, a control box for hardware management, and a computer that runs the COBI Studio software (Ayaz et al., 2011) for data acquisition. fNIRS 1100 is a continuous wave NIRS system that uses wavelengths of 730 nm and 850 nm for optical imaging of the prefrontal cortex. The sensor has a source-detector separation of 2.5 cm, which allows for approximately 1.25 cm penetration depth. This system can monitor changes in relative concentrations of oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) at a temporal resolution of 2 Hz from 16 locations over the prefrontal cortex (Ayaz et al., 2013). Co-registration studies that mapped the measurement optodes on a standardized brain template associated these locations with Brodmann areas 9, 10, 44, and 45 (Ayaz et al., 2006; Chen et al., 2017).

ANALYSIS & RESULTS

The scenario tasks were grouped into three parts, and their analyses consisted of the analysis of behavioural performance (task completion times), eye movement parameters (mean fixation duration and total number of fixations for all parts), and fNIRS parameters. For this, the recordings (eye movements, speech instances, fNIRS) were annotated and time-synchronized manually at a higher level (i.e., the level of fixations and speech instances). Critical regions of interest (RoIs) were identified to narrow down the scope of the analysis to a manageable size for a comparative analysis between the sessions.

Behavioural Performance Analysis Results

Behavioural performance analysis consisted of the analysis of task completion times. The results revealed statistically significant differences ($p < 0.05$) between the two sessions for the majority of the subtasks. In the second session, the pilots completed the tasks in shorter times. For instance, for the portal setting task where the pilots were asked to assign specific functions to the four sections on the LAD screen, the pilots took on average 124.8 seconds to complete the scenario, whereas, in their second flight, it took them on average 97.5 seconds to finish the same task.

Eye Movement Analysis Results

For the analysis of eye movements, annotated fixations were analysed in terms of the mean fixation durations and the total number of fixations on predefined areas of interest (AoIs). In the majority of the subtasks, statistically significant differences ($p < 0.05$) were obtained between the two sessions. As for the relationship between mean fixation durations (i.e., the duration of single fixations) and the number of fixations (on a predefined AoI), the general tendency was that the pilots had longer mean fixation durations accompanied by less number of fixations on the target (correct) AoIs. For instance, in one of the subtasks and one of the pilots, the mean fixation durations increased from 570 ms to 709 ms (ANOVA test, $F=4.221$, $p<0.005$, $\eta^2=0.078$), whereas the total number of fixations to complete the subtask reduced from 59 to 28. The results for most of the subtasks were in line with this finding.

We also analysed gaze locations in terms of their congruency with the target AoIs. Each subtask required gaze allocation to a specific part of the display (a predefined, target AoI in the analysis). The results revealed that the overall accuracy of the pilots was overall above 90%, mostly approximating full accuracy (viz. a ceiling effect) in both sections. Therefore, an analysis of accuracy, in terms of gaze allocations to target AoIs did not return comparable results.

fNIRS Analysis Results

The fNIRSoft v4.11 software (Ayaz, 2010) was used for the inspection and processing of raw infrared signals. Raw fNIRS data were visually inspected for cases, including excessive noise and motion artifacts. Saturated optodes exceeding the signal-to-noise ratio limit of the sensors were excluded from the analysis. The data obtained from optodes 8 and 10 were excluded due to poor sensor-skin contact. The Sliding Windows Motion Artifact (SMAR) filter (Ayaz et al., 2010) was used to attenuate signal changes due to head movements. Next, raw fNIRS data were low-pass filtered with a linear phase filter with order 20 and a cut-off frequency of 0.1 Hz to reduce the effects of high-frequency noise due to cardiac cycles. The filtered fNIRS data were then converted into oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) concentration changes by using the Modified Beer-Lambert Law. The HbO values were used as the main dependent variable to explore prefrontal oxygenation changes due to learning. The HbO measures were baseline-corrected with respect to the beginning of each scenario block.

The HbO responses observed during the first and second flights were contrasted for both pilots. Both pilots exhibited improved performance in finding the corresponding buttons on the LAD screen to make the required functional assignments to each LAD section. Overall, we observed a decrease in HbO response during the second flight, especially at optodes 11–16 in the right prefrontal cortex, accompanied by a decrease in response times. Figure 2 shows the mean HbO levels observed during the portal setting task, where the pilots configured what information would be displayed at each display area. To conduct a statistical test, we divided the scenario segment into 10-second-long blocks and then compared the mean HbO values in the overlapping first 7 blocks. We found a significant decrease in HbO responses observed during flight 2 as compared to flight 1 for optode 11 ($t(6)=3.99$, $p<0.01$, Cohen $d=1.51$), optode 12 ($t(6)=4.44$, $p<0.01$, Cohen $d=1.68$), optode 13 ($t(6)=3.50$, $p<0.05$, Cohen $d=1.33$), optode 14 ($t(6)=3.05$, $p<0.05$, Cohen

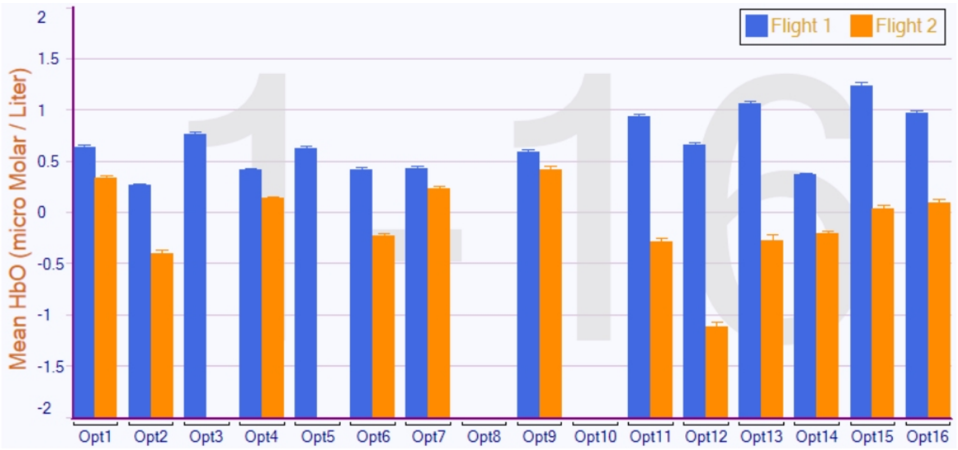

Figure 2: Contrast of the first and second flights in terms of mean prefrontal HbO response levels observed during the portal configuration task.

d=1.15), optode 15 ($t(6)=3.24$, $p<0.05$, $p<0.05$, Cohen $d=1.22$), and optode 16 ($t(6)=3.64$ $p<0.05$, Cohen $d=1.38$). A similar pattern of deactivation was observed in the remaining scenarios as well.

DISCUSSION & CONCLUSION

Overall, our results indicated that pilots' task completion times decreased for the majority of the tasks, accompanied by an increase in eye fixation durations, a decrease in the number of fixations, and a decrease in the right superior and dorsolateral prefrontal cortex activation. Despite the small sample size, these results suggest supporting evidence for the learnability of the new interface paradigm through task measurements in multiple scenarios. In particular, both pilots were able to remember the locations of relevant functions and the touch-based gestures to operate them. Although they completed only two sessions with a separation of 2 weeks, both pilots were able to improve their performance by remembering the relevant functional areas of the display and utilizing the affordances of the LAD screen for the majority of the test scenarios.

The improvement in task completion times, the decrease in the number of fixations per scenario, and the increase in fixation durations are altogether indicative of the pilots' reduced effort in allocating their visual attention while performing the same tasks for the second time. When viewed together with the decreased oxygenation trend in the right superior frontal and dorsolateral prefrontal regions that play a key role in the right frontoparietal network associated with deliberate guidance of visual attention (Duncan, 2013), one can argue that the pilots succeeded in developing a sensorimotor adaptation to the LAD mediated cockpit environment within a rather short amount of time. Such neuroergonomic evidence of sensorimotor adaptations can be a useful resource for the assessment of the learnability of a mission-critical human-machine interface.

REFERENCES

Ayaz, H., Izzetoglu, M., Platek, S. M., Bunce, S., Izzetoglu, K., Pourrezaei, K., et al. (2006). "Registering fNIR data to brain surfaceimage using MRI templates," in Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society (NewYork, NY: IEEE), 2671–2674.

Ayaz, H. (2010). Functional Near Infrared Spectroscopy Based Brain Computer Interface. Philadelphia, PA: Drexel University.

Ayaz, H., Izzetoglu, M., Shewokis, P. A., & Onaral, B. (2010). Sliding-window motion artifact rejection for functional near-infrared spectroscopy. In 2010 Annual International Conference of the IEEEEngineering in Medicine and Biology (pp. 6567–6570). IEEE.

Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., &Onaral, B. (2012). Optical brain monitoring for operator training andmental workload assessment. Neuroimage, 59(1), 36–47.

Ayaz, H., Onaral, B., Izzetoglu, K., Shewokis, P. A., McKendrick, R., and Parasuraman, R. (2013). Continuous monitoring of brain dynamics withfunctional near infrared spectroscopy as a tool for neuroergonomicresearch: empirical examples and a technological development. Frontiersin Human Neuroscience. 7:871.

Booher, H. R., Minninger, J. (2003) “Human systems integration in armysystems acquisition”, in: Handbook of human systems integration, Booher, Harold (Ed.). pp. 663–698

Blignaut, P. J., Beelders, T. R., & So, C. Y. (2008). The visual spanof chess players. In Proceedings of the 2008 symposium on Eye TrackingResearch & Applications (pp. 165–171). ACM.

Booth, P. (1989). An Introduction to human computer interaction. London: Lawrence Erlbaum Associates.

Chen, M., Blumen, H. M., Izzetoglu, M., and Holtzer, R. (2017). Spatialco-registration of functional near-infrared spectroscopy to brainMRI. Journal of Neuroimaging, 27, 453–460

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontallobe recruited by diverse cognitive demands. Trends in neurosciences, 23(10), 475–483.

Duncan, J. (2013). The structure of cognition: attentional episodes inmind and brain. Neuron, 80, 35–50.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertisedifferences in the comprehension of visualizations: A meta-analysis ofeye-tracking research in professional domains. Educational PsychologyReview, 23(4), 523–552.

Hill, N. M., and Schneider, W. (2006). Brain changes in the developmentof expertise: neuroanatomical and neurophysiological evidence aboutskill-based adaptations. In K. A. Ericsson, N. Charness, P. Feltovichand R. Hoffman (Eds) The Cambridge Handbook of Expertise and ExpertPerformance (pp. 653-682). New York, NY: Cambridge University Press.

Saeed, S., Bamarouf, Y. A., Ramayah, T., & Iqbal, S. Z. (Eds.). (2016). Design solutions for user-centric information systems. IGI Global.

Shackel, B. (1991). Usability - context, framework, definition, designand evaluation. In B. Shackel & S. Richardson (Eds.), Human factors forinformatrics usability (pp.21–37). New York, NY: Cambridge UniversityPress.

Thiebaut de Schotten, M., Dell’Acqua, F., Forkel, S., Simmons, A., Vergani, F., Murphy, D. G., & Catani, M. (2011). A lateralized brainnetwork for visuo-spatial attention. Nature Neuroscience, 14(10), 1245–1246.

Tonsen, M., Baumann, C. K., & Dierkes, K. (2020). A high-leveldescription and performance evaluation of pupil invisible. arXivpreprint arXiv:2009.00508.

Wickens, C. D. (2008). Multiple resources and mental workload. HumanFactors, 50(3), 449–455.