In: Harris, D., Li, WC. (eds) Engineering Psychology and Cognitive Ergonomics. Cognition and Design. HCII 2020. Lecture Notes in Computer Science, vol 12187
DOI: 10.1007/978-3-030-49183-3_24

Understanding human behaviour in flight operation using Eye-Tracking technology

Wojciech Tomasz Korek[1,2], Arthur Mendez[2], Hafiz Ul Asad[2], Wen-Chin Li[3], and Mudassir Lone[2]

[1] Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, Gliwice, PL
[2] Dynamics, Simulation and Control Group, Cranfield University, Cranfield, UK
`{w.t.korek,a.mendez,a.ul-asad,m.m.lone}@cranfield.ac.uk`
[3] Safety and Accident Investigation Centre, Cranfield University, Cranfield, UK
`wenchin.li@cranfield.ac.uk`

Abstract. A clear understanding of how the pilot processes the information in the cockpit while carrying out particular tasks is crucial for developing the Human-Machine Interface and inceptors that help reduce pilot workload. Eye-tracking data synchronised with aircraft dynamics data is used here to study the high-workload scenario of executing an offset landing in an engineering flight simulator. The study focused on identifying differences in behavioural patterns between line pilots and test pilots. Evidence for significant differences were found regarding the ability to multitask and monitor aircraft states. The research output will lead to reduction of the pilot's workload and, in further study, proposition of a new display setups and inceptors.

The final authenticated version is available online at https://doi.org/10.1007/978-3-030-49183-3_24.

Keywords: Human behaviour · Situation awareness · Flight simulation · Eye-tracking.

1 Introduction

The design of human machine interface in general and even more so in aerospace requires an in-depth understanding of the way a pilot utilises the information provided to him/her. Consequently, human behaviour has significant implications on aspects ranging from the design of individual displays and inceptors all the way to the design of flight control systems. Numerous studies exist that focus on developing an understanding of pilot's control behaviour, some of which are limited to purely mathematical representation of control action [12] while others focus on psychological aspects [5]. In [1], the author investigated the workload on an air traffic controller from the use of weather display using an eye tracker and concluded that the eye movement tracker can give a good measure of a controller workload when he was subjected to have a frequent visit of a weather display. The effect of anxiety on a pilot performance during an aircraft landing has been studied in [2] where they showed, using simulated landings, that the pilot anxiety increased when encountered with difficult situation, such as landing under stressful situation. Cognitive workload of an air traffic controller has been investigated in [3] using Functional Near-Infrared spectroscopy with the findings that accuracy and speed of a pilot decreased and blood oxygenation increased with the increase in task difficulty. Eye tracking has been also used in studying different aspects of human-machine interaction [15, 14]. In this paper, the authors attempt to address some of the limitations found in mathematical models of the human pilot by using eye tracking data to provide insight into the way the pilot uses the available information. Although numerous studies have been carried out using eye-tracking for various purposes in aviation, for example recognising scan patterns of a remote air traffic control by a single controller [10], flight deck design [11], augmented reality (AR) in Primary Flight Display (PFD) [9], air traffic controller's situation awareness [7] and understanding human behaviour during aircraft-pilot coupling events [8], there has been limited transfer of this knowledge into the mathematical formulation of pilot behaviour. In this study the authors synchronise temporal and spatial eye-tracking data with aircraft data to study scenarios where the pilot is given an urgent safety critical task. The insight from this study, where pilot's attention allocation is quantified and correlated with aircraft data, enables the reconsideration of critical cockpit information for specific tasks. The overall research aim is to first understand the control behaviour and information needs of the pilot, and in doing so reduce his/her workload by proposing new display setups and inceptors. The work discussed in this paper focuses on the pilot's visual cues (situation awareness) which consists of the eye-tracker data analysis and its correlation with the aircraft data. This will help understanding and defining the whole process of the pilot's decision making within the pilot-aircraft control loop. The diagram of the control system is shown in Figure 1. This is a part of the pilot-vehicle-system under manual control which is presented in more detail in [12]. The visual pilot cues that affect his/her decision making are directly influenced by pilot looking either on a PFD or outside the aircraft's window.

2 Task & Method

2.1 Experimental setup

Data for the test was gathered using flight simulator and off-the-shelf eye-tracker. The eye tracker and aircraft dynamics were synchronised to provide a more complete picture of how the pilot controls the aircraft in high stress cases. Statistical analysis of experimental data was carried out to see correlations between the stick deflection levels and eye gaze positions. The tests were conducted using an engineering flight simulator called the Future Systems Simulator. The rig is a flexible aircraft systems simulator platform for the development of intelligent, integrated aerospace technologies. It is a test-bed for a wide variety technologies impacting today's aircraft and future aircraft concepts. The tests were carried out using the mathematical model of a regular business jet.

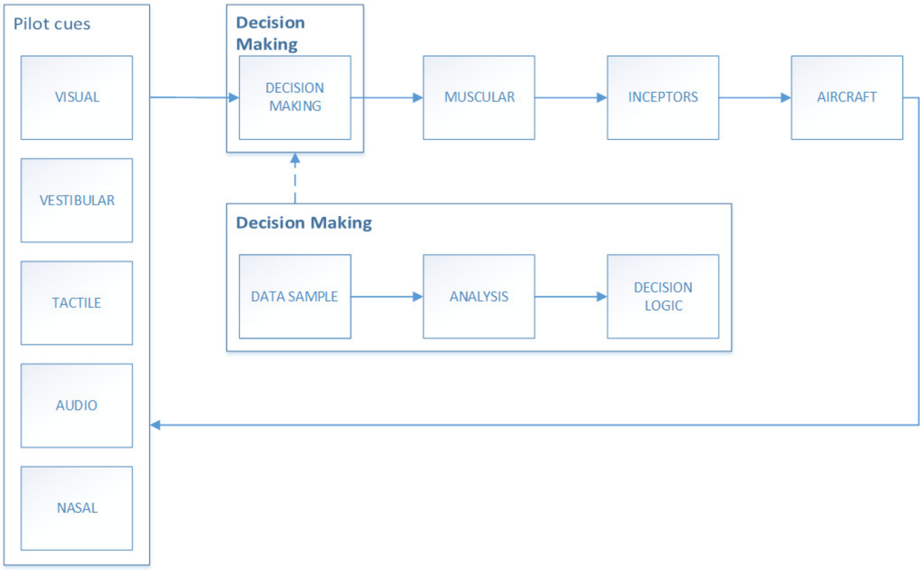

Fig. 1. Diagram showing the pilot-aircraft control loop.

The Pupil Labs' Pupil Core has been used for tracking eye movements. It includes a pair of glasses with a camera pointing towards where the wearer is looking and a second camera tracking the eyeball movement. Figure 2 presents the calibration process and the general eye-tracker setup. During the calibration the world frame, which is the wearer field of view, is defined. The red lines indicate the point of view of the world camera, while blue lines show the eye-camera video transmission. After calibration, a dedicated software is able to visualise the point where the test subject is looking at. Recording allows the software to create the normalised fixation positions in the world frame [13]. Fixations are a group of gaze positions focused on one point at a given time frame [6]. The Pupil Labs' Pupil Capture application allows the user to specify Areas Of Interest (AOIs). By attaching the AprilTags on the viewing range of the front-facing eye-tracker camera and defining the surface between them in the software two AOIs could be specified: PFD (marked by four AprilTags in Figure 3) and the outside window view. These are two aspects that affect pilot's decision making [12]. With eye-tracking technology some limitations still exist regarding the collection and analysis of eye-tracking data. This is discussed in the results & analysis part.

2.2 Scenario

The test case used in this study is the offset-landing task as shown in Figure 4. The pilot/test subject is given control of the aircraft in the final stages of the landing phase and he/she is required to correct for the lateral offset from

Fig. 2. Eye-tracker setup.

Fig. 3. Pilot's perspective and example fixation (here: the pilot is looking on the outside of the window).

the runway centreline and the glide path angle change that results from the manoeuvre. Hence, the pilot is required to make inputs in pitch, roll, and throttle to attain the desired trajectory of the aircraft. This task is classified as a high workload scenario due to the need for manoeuvring the aircraft close to the ground. Furthermore, the stress level of the task can be increased by reducing the distance from the runway, at which the pilot is allowed to execute the S-shaped manoeuvre. This task is commonly used for the handling qualities assessment of flight control systems, where the feedback from the pilot is captured on a Cooper-Harper rating scale (CHR) [4]. It should be noted that for the tests conducted in this study, the Auto-throttle functionality was enabled so as to eliminate the need for the pilot to focus on airspeed management. Hence, during the task the pilots need only focus on lateral offset distance correction and glide path angle capture and maintenance. The aircraft was also configured for landing with flaps and landing gear deployed, which eliminated the need for a pilot to get accustomed to the touch-screen based flaps and landing gear levers and any influence this might have on their behaviour. The pilots were also given a flight time in the simulator beforehand to familiarise themselves with the simulator's PFD and the sidestick sensitivity.

The experimental parameters for this offset-landing task are δ - the distance from the runway at which the pilot is allowed to execute the sidestep manoeuvre and Δ - the lateral offset distance from the runway centreline to the position of the aircraft before the manoeuvre. The initial altitude of the aircraft was determined using the longitudinal distance (δ) to the runway and a desirable glide slope angle of 3°. Here, in these tests the offset distance from the runway centreline was fixed at 416 metres which corresponds to the lateral distance between two parallel runways at the test airport. Hence, the pilot was in short conducting a switch from the right-hand side runway to the left-hand side runway. The two values of δ chosen for this study were 4 and 2 nautical miles (nm). These distances were chosen to reflect a moderate workload and a high workload scenario respectively. These distances were established through an initial handling qualities assessment of this offset landing test with varying values of δ and then capturing the feedback from the pilots on the CHR scale [4]. From these tests it was revealed that for $\delta = 4$ nm, the rating from the pilot was a "3" on the CHR scale which corresponds to the desired performance being achieved with minimal pilot compensation. For the longitudinal distance, $\delta = 2$ nm, the rating on the CHR scale was a "5" which corresponds to considerable pilot compensation being required for achieving adequate performance. Hence, $\delta = 4$ and $\delta = 2$ nm were chosen to signify a moderate and high workload scenarios respectively along with a straight-ahead landing without the S-shaped manoeuvre as the baseline.

2.3 Test subjects

For the purpose of this test two pilots were asked to carry out the baseline and two offset landings. One of them was a line pilot who flies business jets on a regular basis with 4000 hours of total flying time. The second pilot was a test pilot with significantly more experience on different types of aircraft and flight

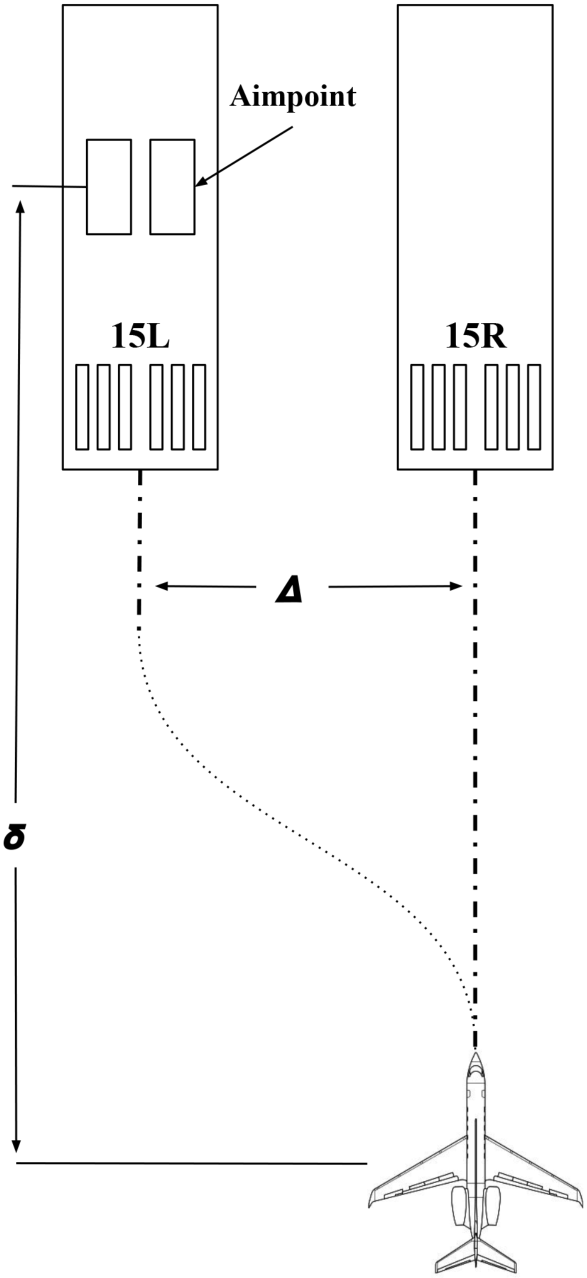

Fig. 4. Plan view of the offset-landing task.

simulators: 3000 total flying hours on over 50 different aircraft types. Because of this he was able to notice more aspects of a flight in a new environment which made his flight pattern noticeably different.

2.4 Data curation

Two sets of data were gathered during each test: one from the eye-tracker and the other from the flight simulator. The first one, concerning the eye gaze fixations, consists of timestamp, fixations' x and y position and fixation duration. The latter includes flight dynamics data, from which the simulation time, glide angle (γ), longitudinal and lateral stick deflection (pitch and roll input) and aircraft's position and altitude are used for analysis. Two ways of data synchronisation were taken into account: distance from the airport and fixed time from the touchdown moment. The pilot's behaviour changes depend on whether the aircraft is on the ground or in the air. Because of that and since the time it takes to get from the starting point (which was identical for all the trials) to the landing point is slightly different for each flight, the second method was chosen. Thus, to synchronise the data, all recordings have been cut to end at the touchdown time t_{TD} and begin 180 seconds before it. This allowed for the correlation of the simulation inputs and outputs with the eye-tracker data. This time duration included the offset manoeuvre for each test/scenario. The data has been processed using MATLAB® software.

3 Results

For the initial trials the authors had a test pilot as a subject. Preliminary results of these trials have shown that during the offset landing the pilot was looking outside the cockpit for 82% of the simulation time, making it the most critical element of the pilot's field of view. In the main tests the authors decided to focus on the correlation between eye-tracking data and simulation data. Moreover, some differences of the line and test pilots' behaviour were noticed.

Figures 5 and 6 show the spatial distribution of fixation positions for line pilot and test pilot respectively. It is evident that both pilots were looking outside the aircraft's window for the majority of time for all three landings. However, the major difference is the focus of the fixations: the test pilot has a dispersed scanning pattern. He is moving his gaze more often and monitoring more elements during the flight. Figures 7 and 8 present the temporal distribution of fixation points' position in Y axis, as this was the variable used to specify the outside window/PFD threshold.

The difference in Window/PFD threshold occurs due to the eye-tracker calibration being specific to each pilot. It was evaluated by assessing the video recording and scatter plot showing the fixations' positions. Due to the limitation in the eye-tracker accuracy, both in terms of calibration and resolution, it was not possible to distinguish specific PFD elements that pilot was looking

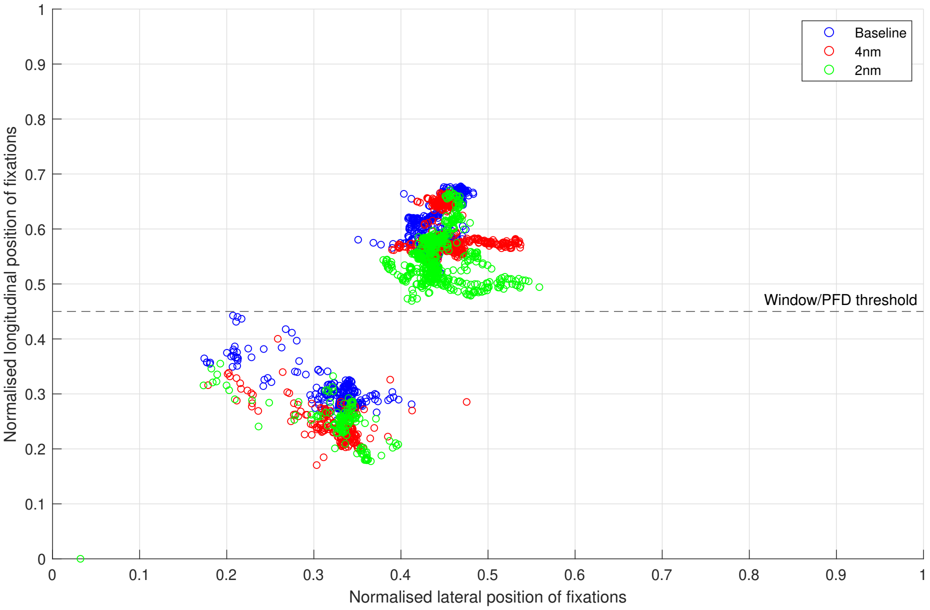

Fig. 5. Spatial distribution of fixation points for the line pilot.

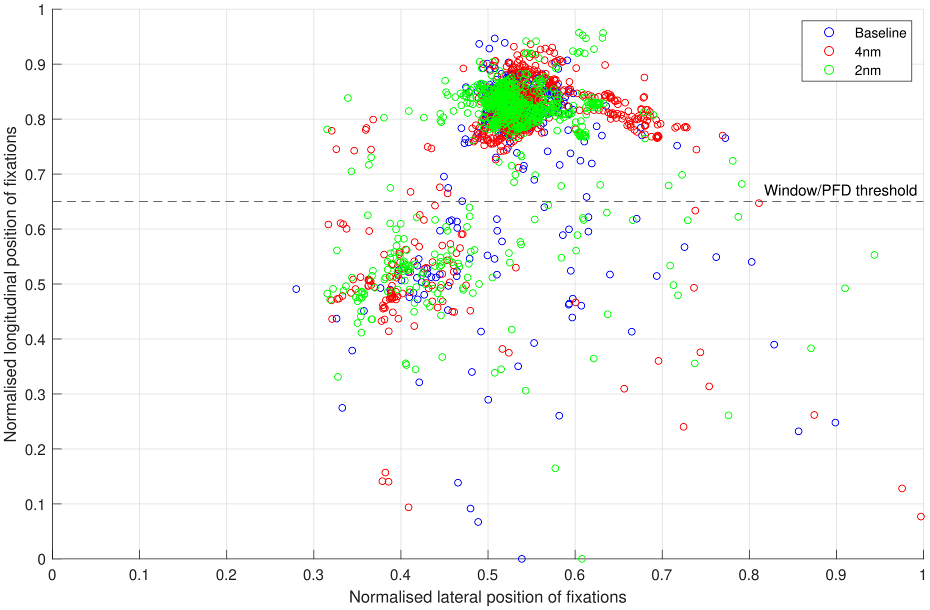

Fig. 6. Spatial distribution of fixation points for the test pilot.

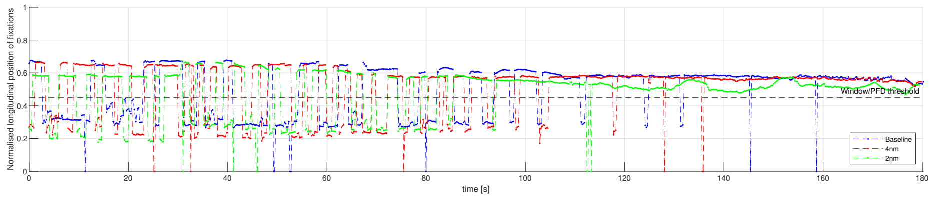

Fig. 7. Temporal distribution of fixation points' position on Y axis for the line pilot.

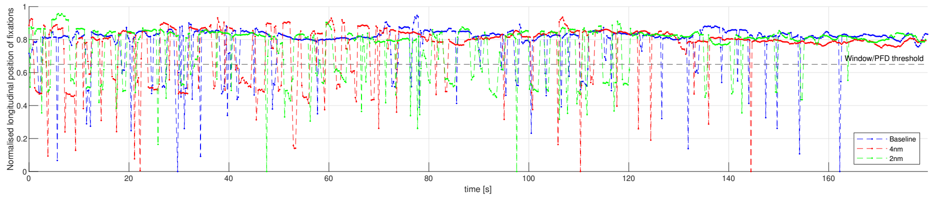

Fig. 8. Temporal distribution of fixation points' position on Y axis for the test pilot.

at, such as airspeed or altitude indicators. However, the gathered data provides insight into the main differences between the test subjects.

Next difference between the pilots was that, with each landing being more stressful, behaviour of each pilot was changing differently - the line pilot preferred to look more outside the window, while the test pilot was focusing more on the PFD, although in total it was still less than the line pilot, as shown in Figure 9.

Figures 10, 11 and 12 present the flight dynamics data from the simulator combined with pilot's AOI fixation gaze. The flight dynamic parameters consists of the aircraft's glide path angle γ (which is the landing approach angle with a $-3°$ target), longitudinal and lateral stick deflection (pitch and roll input, ranging from -1 to 1 for each axis), and Δ - the aircraft's offset distance from the runway's centreline. A fixation location above the threshold corresponds to the pilot's attention being allocated outside the cockpit and a value below the threshold corresponds to their attention fixated on the PFD.

For the baseline test (Figure 10), it can be seen that the line pilot's pitch inputs have much higher gain/amplitude compared to the test pilot. As a consequence, the test pilot was closer to the target of $-3°$ throughout the landing unlike the line pilot who's γ had more fluctuations before the flare. Figure 10 also shows the subplot with roll inputs for both pilots. From this it can be seen that the test pilot makes more small amplitude-refining inputs compared to the line pilot and is also able to maintain a smaller offset error from the runway throughout. From the subplot of the fixation location, the test pilot makes a slightly more frequent switches between allocating attention to the PFD and outside the cockpit compared to the test pilot.

Figure 11 shows the different parameters for the 4 nm offset test. The similar observation seen in pitch inputs of baseline test can be seen here as well, i.e. the line pilot has larger amplitude inputs. In this test the roll inputs of the

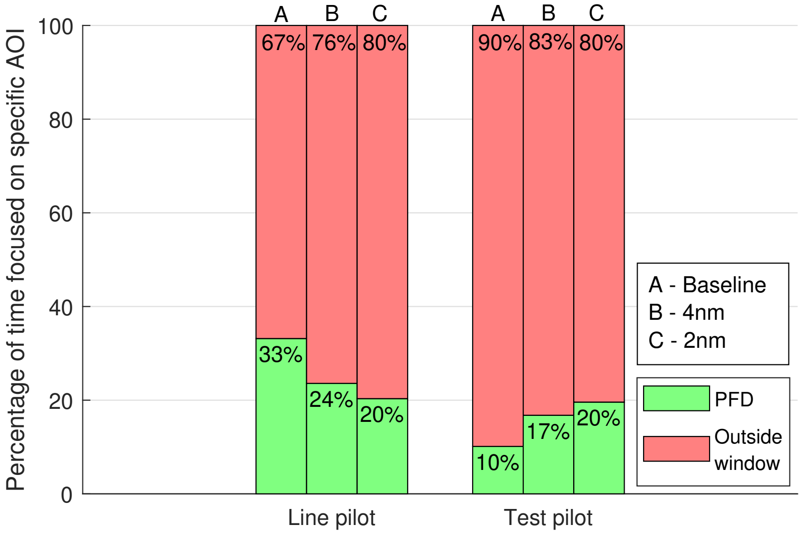

Fig. 9. Percentage arrangement of fixation focus for each pilot during each test flight.

test pilot are much larger in amplitude compared to the line pilot. Hence, the test pilot executes the necessary manoeuvre more aggressively compared to the line pilot. The test pilot in this test, similar to the baseline test, also switches focus between PFD and outside cockpit more often than the line pilot. Figure 12 shows the same data but for the 2 nm offset test. Here pitch inputs of larger magnitude of the line pilot are observed again like in the previous tests. However, unlike the previous tests, the roll inputs of the line pilot become higher in the amplitude and the offset distance error trend for both pilots is very similar. From the fixation location subplot of Figure 12, it can be seen that the line pilot keeps his attention fixed to the outside of the cockpit in the last minute before touchdown. As a consequence of this, his accuracy in maintaining the glide path angle is worse compared to the test pilot as observed in the γ subplot.

Table 1 shows the root mean squares (RMS) values of pitch and roll inputs made by the pilots for the different tests. The RMS value was calculated to gain an insight into how active the pilots were at the controls for each of tests. The test pilot's activity at the controls was lower compared to the line pilot for all three tests in both pitch and roll commands. It can also be observed that as the test becomes more stressful both pilots increase their activity at the controls.

This high gain nature of the line pilot is also evident in Figure 13 which categorises the stick movements on whether the pilot was looking at the PFD or outside the window while the input was made. Figure 14 shows the same plot for the test pilot. These Figures were generated for the most stressful of the three tests, i.e. test C (2 nm offset). When the pilot is focused on the outside view while making the input it can be inferred that the pilot is looking at the

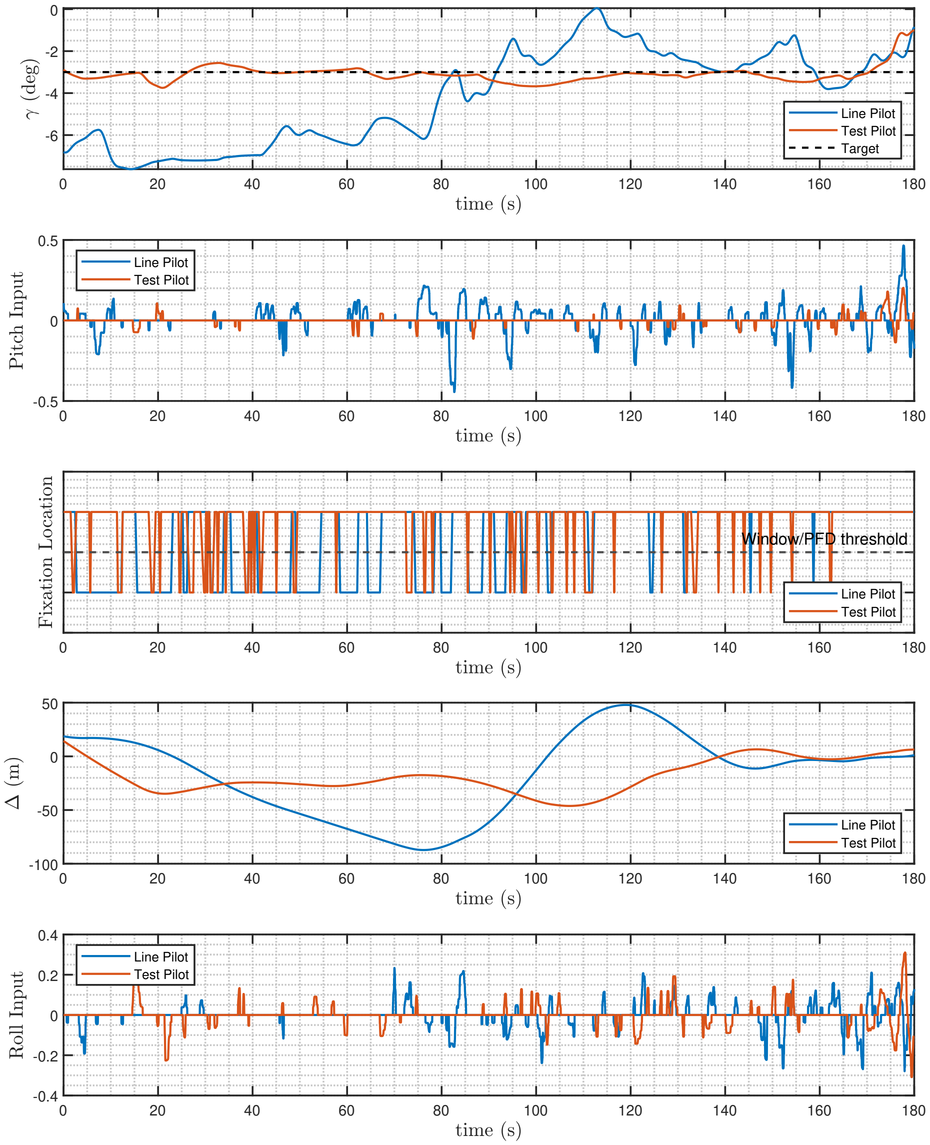

Fig. 10. Longitudinal normalised eye gaze position synchronised with simulator data for both pilots during the baseline landing.

	Pitch Input RMS			**Roll Input RMS**		
	Baseline	**4nm**	**2nm**	**Baseline**	**4nm**	**2nm**
Line Pilot	0.09	0.09	0.19	0.08	0.12	0.17
Test Pilot	0.04	0.04	0.06	0.05	0.09	0.1

Table 1. RMS values of pitch and roll inputs during each test for both types of pilots.

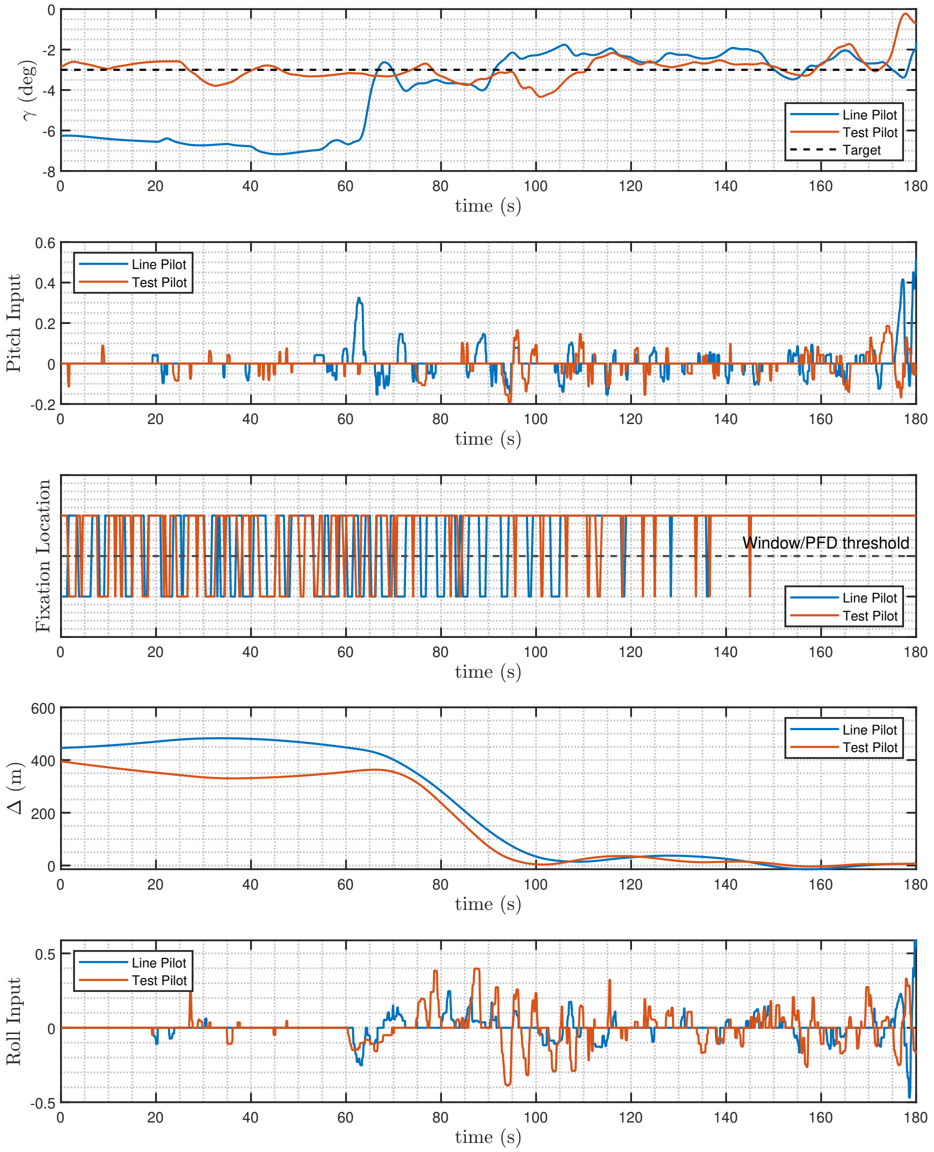

Fig. 11. Longitudinal normalised eye gaze position synchronised with simulator data for both pilots during the 4-nm offset landing.

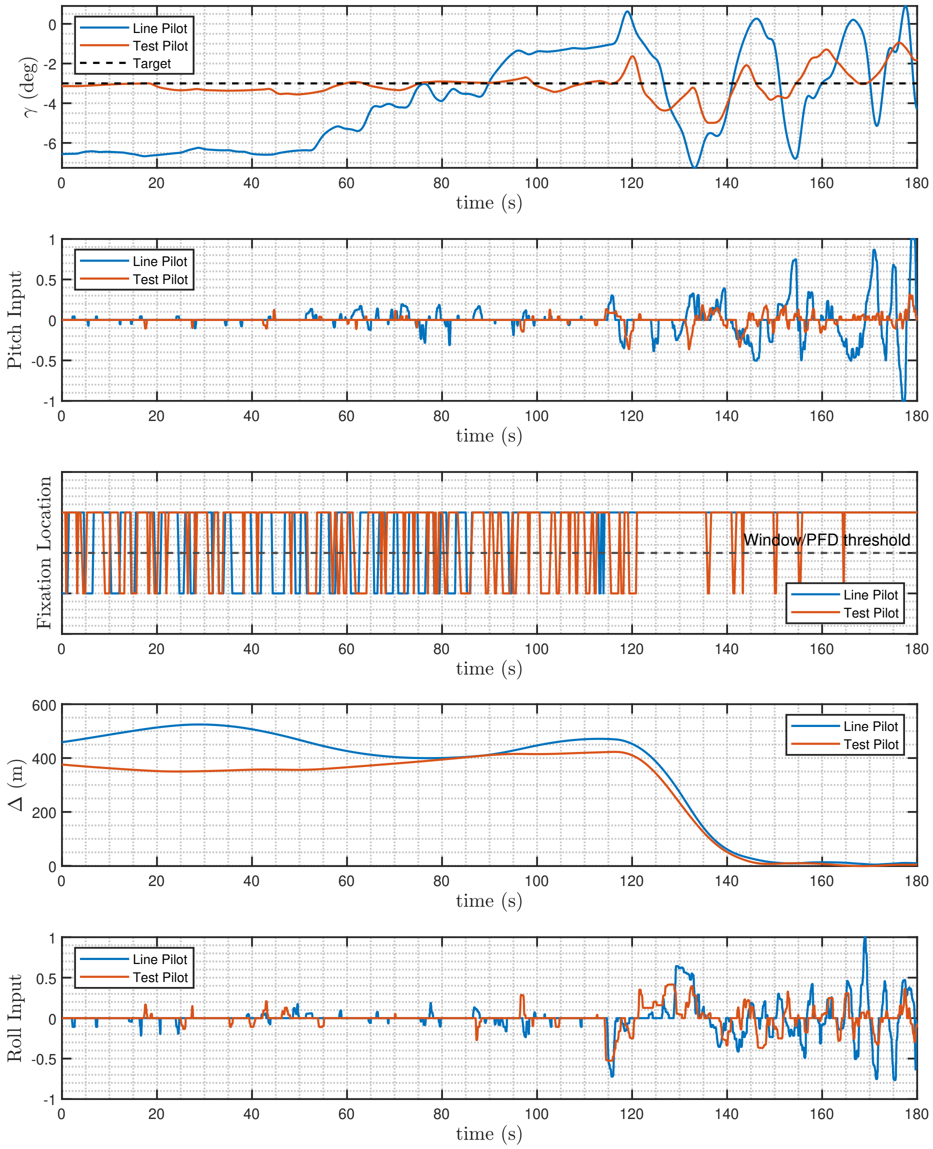

Fig. 12. Longitudinal normalised eye gaze position synchronised with simulator data for both pilots during the 2-nm offset landing.

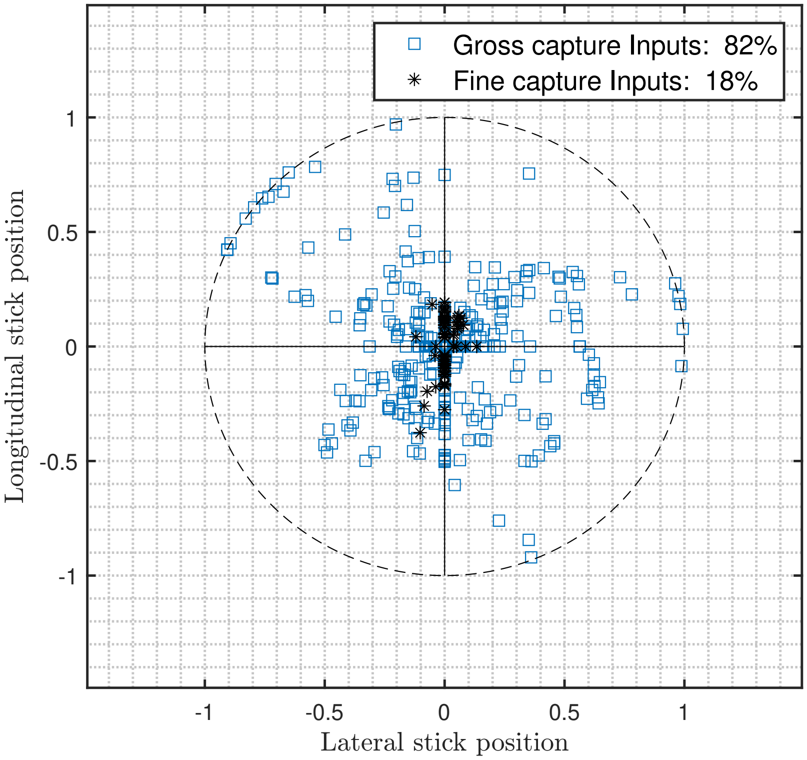

Fig. 13. Stick movements categorised according to line pilot's fixation position for the 2nm test.

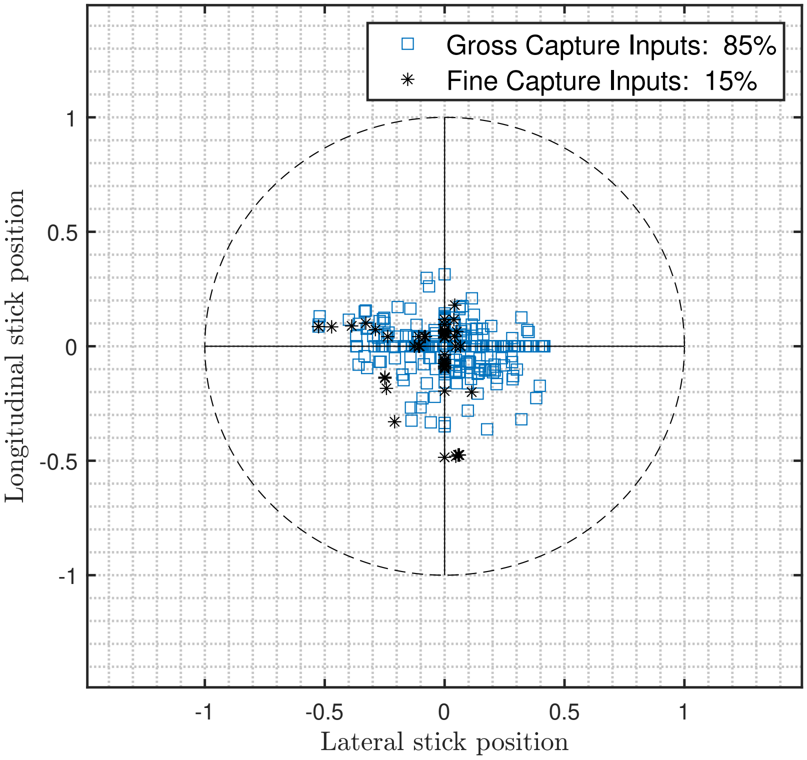

Fig. 14. Stick movements categorised according to test pilot's fixation position for the 2nm test.

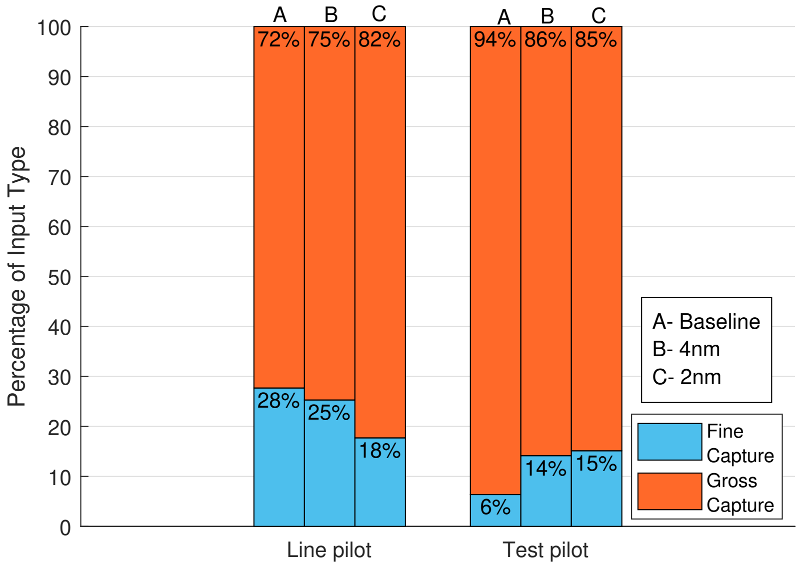

Fig. 15. Comparison of input type distribution for line pilot and test pilot for all three tests.

runway and aiming at touchdown point. Hence, the inputs made by the pilots in these instances are for gross capture. In contrast, when the pilots are looking at the PFD, they are trying to control and maintain their glide slope angle to an ideal value. Hence, the inputs made in these instances are for the fine/precise capture of the glide slope angle. From both these Figures it can be seen that the majority of the inputs made by both the pilots are gross capture inputs. Furthermore, the majority of their fine capture inputs are purely longitudinal. This is expected as the flight path vector and vertical speed readings on the PFD are the only cues that determine the aicraft's glide slope angle. From Figure 14 it is also clear that the test pilot has executed high gain roll inputs while his gaze was fixed on the PFD. As the PFD doesn't give information to the pilot about the offset distance, it can be inferred that the test pilot in these instances was multi-tasking by correcting the offset through a roll input while monitoring the glide slope angle at the same time. Hence, the Figures suggest that the test pilot might be more comfortable with multi-tasking compared to the line pilot.

Figure 15 shows the input type distribution for both the pilots and all three of the tests conducted. It is interesting to note that as the landing scenario becomes more stressful, the line pilot tends to increase their gross capture inputs compared to their baseline landing, while the test pilot's fine capture inputs increase in the same situation.

4 Conclusions & further work

In this paper the authors present a detailed analysis of the data collected over the entire test program. The analysis focuses on developing an understanding the pilot attention allocation in high workload scenarios and therefore, provides an understanding of the pilot's adopted control strategy. The Eye-tracking fixations' data gathered was correlated with the flight simulator output. The differences and similarities between the line pilot and test pilot were observed. The further analysis may use the heatmaps within AOIs to distinguish specific elements that the pilot focuses on (airspeed, altitude etc.) and develop an information usage profile. This in turn will allow the proposition of new display setups and inceptors to reduce workload. In the future more pilots may be involved as a test subjects to give more significance (and possibly variance dependable on a pilot's previous experience) to the gathered data. Moreover, different scenarios, such as air-to-air refuelling or path following will be considered along.

Acknowledgements. Co-financed by the European Union through the European Social Fund (grant POWR.03.02.00-00-I029).

References

1. Ahlstrom, U., Friedman-Berg, F.J.: Using eye movement activity as a correlate of cognitive workload. International Journal of Industrial Ergonomics **36**(7), 623–636 (jul 2006). https://doi.org/10.1016/j.ergon.2006.04.002, https://linkinghub.elsevier.com/retrieve/pii/S0169814106000771
2. Allsop, J., Gray, R.: Flying under pressure: Effects of anxiety on attention and gaze behavior in aviation. Journal of Applied Research in Memory and Cognition **3**(2), 63–71 (jun 2014). https://doi.org/10.1016/j.jarmac.2014.04.010, https://linkinghub.elsevier.com/retrieve/pii/S2211368114000333
3. Ayaz, H., Willems, B., Bunce, S., Shewokis, P., Izzetoglu, K., Hah, S., Deshmukh, A., Onaral, B.: Cognitive Workload Assessment of Air Traffic Controllers Using Optical Brain Imaging Sensors. Advances in understanding human performance: Neuroergonomics, human factors design, and special populations pp. 21–31 (jun 2010). https://doi.org/10.1201/EBK1439835012-4, http://www.crcnetbase.com/doi/abs/10.1201/EBK1439835012-4
4. Cooper, G.E., Harper, R. P., J.: The use of pilot rating in the evaluation of aircraft handling qualities. Tech. rep., NASA Ames Research Center, Washington (1969). https://doi.org/NASA-TN-D-5153
5. Demerouti, E., Veldhuis, W., Coombes, C., Hunter, R.: Burnout among pilots: psychosocial factors related to happiness and performance at simulator training. Ergonomics **62**(2), 233–245 (feb 2019). https://doi.org/10.1080/00140139.2018.1464667, https://www.tandfonline.com/doi/full/10.1080/00140139.2018.1464667
6. IMotions: 10 Most Used Eye Tracking Metrics and Terms (2019), https://imotions.com/blog/7-terms-metrics-eye-tracking/
7. Kearney, P., Li, W.C., Yu, C.S., Braithwaite, G.: The impact of alerting designs on air traffic controller's eye movement patterns and situation awareness. Ergonomics **62**(2), 305–318 (feb 2019). https://doi.org/10.1080/00140139.2018.1493151

8. Künzel, D.: Flight simulator assessment of pilot behaviour during aircraft-pilot-coupling events. Diploma thesis, Munich University of Applied Sciences (2016)
9. Li, W.C., Horn, A., Sun, Z., Zhang, J., Braithwaite, G.: Augmented visualization cues on primary flight display facilitating pilot's monitoring performance. International Journal of Human-Computer Studies **135** (mar 2020). https://doi.org/10.1016/j.ijhcs.2019.102377
10. Li, W.C., Kearney, P., Braithwaite, G., Lin, J.J.: How much is too much on monitoring tasks? Visual scan patterns of single air traffic controller performing multiple remote tower operations. International Journal of Industrial Ergonomics **67**, 135–144 (sep 2018). https://doi.org/10.1016/j.ergon.2018.05.005
11. Li, W.C., Zhang, J., Le Minh, T., Cao, J., Wang, L.: 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 (jul 2019). https://doi.org/10.1016/j.ergon.2019.04.003
12. Lone, M.: Pilot modelling for airframe loads analysis. Phd thesis, Cranfield University (2013)
13. Pupil Labs: Pupil Labs — Core. (2019), https://pupil-labs.com/products/core/
14. Saravanakumar, S., Selvaraju, N.: Eye Tracking and Blink Detection for Human Computer Interface. International Journal of Computer Applications **2**(2), 7–9 (2010). https://doi.org/10.5120/634-873
15. Shimizu, S., Tanzawa, Y., Hashizume, T.: Classification of gaze preference decision for human-machine interaction using eye tracking device. International Journal of Mechatronics and Automation **2**(2), 75 (2012). https://doi.org/10.1504/IJMA.2012.048183, http://www.inderscience.com/link.php?id=48183

2020-07-10

Understanding human behaviour in flight operation using eye-tracking technology

Korek, Wojciech Tomasz

Springer

Korek WT, Mendez A, Asad HU, et al., (2020) Understanding human behaviour in flight operation using eye-tracking technology. In: 22nd International Conference on Human-Computer Interaction (HCI International) 2020, 19-24 July 2020, Copenhagen, Denmark, pp. 304-320

https://doi.org/10.1007/978-3-030-49183-3_24