Monitoring eye movements in real flight conditions for flight training purpose

Frédéric Dehais[✉][*], Stéphane Juaneda[*], and Vsevolod Peysakhovich[*]

[*]ISAE-SUPAERO, Université de Toulouse, France

Eye-tracking is a relevant technique to investigate pilots' cognitive performance. It offers promising prospects to enhance pilots' training and be used for on-line monitoring for smart cockpits purposes. Most of the studies are conducted in flight simulators which may limit the interpretation of the eye-tracking data. In this study, we investigate the possibility to measure eye movement in real flight. We conducted an experiment with 7 pilots performing two traffic patterns and basic flight maneuvers in a real light aircraft. We analyzed the distribution of attention over the main areas of interest (flight instruments) during the different flight phases. These data were confronted with operational procedures and discussed with flight instructors in terms of flight safety and recommendations for training. Also, a classifier trained on the first traffic pattern could predict the three phases (take-off, downwind, and landing) of the second one with a mean accuracy of 70%.

Keywords & Phrases: Eye tracking, training, monitoring, real flight condition, flight safety

1 INTRODUCTION

Flying is a challenging activity that takes place in a complex, uncertain and dynamic environment. It is of great importance for pilots to constantly monitor the flight deck as well as the out-of-the window to update situation awareness and detect potential threats (eg. failure, incoming traffic). Ab-initio pilots are trained to scan the main panel instruments while glancing at the outside world. It is now well admitted that several factors such as stress, workload, time-on-task, and fatigue can impair monitoring abilities to an extent that critical events (eg. flightpath deviation, alarms) remained unnoticed [1–4]. Several studies have pointed out that experts exhibited higher scanning proficiency than novices (see [5–7] for comprehensive reviews) thus being less affected by the deleterious effects of stressors [8, 9]. These findings suggest that monitoring is a key technical skill that needs to be trained to improve the efficiency and safety of aircraft operations. However, little is known on how pilots scan the cockpit especially from an instructor's point of view when his/her latter gives flight lessons to their trainees.

To that end, eye tracking offers fruitful perspectives for training [10, 11] since it provides both behavioral [9] and physiological outputs [12] to assess the trainee's performance and mental effort. For instance, the video recording of the session can be replayed during the debriefing by the flight instructor to point out the poor distribution of attention and make recommendations to their students. More sophisticated analyses can also be performed using metrics such as saccadic activity, fixations and dwell time, gaze entropy, blink rates and variation of pupil diameter (see [13] for a review). This approach allows computing objective indicators of visual performance leading to disclose statistical evidence of efficient or non-efficient ocular behaviors if conducted over large samples of pilots. Eventually, machine learning techniques could be applied to automatize the detection of abnormal ocular events with regard to a reference model to assist the flight instructor during the debriefing. In the long run, this approach would help flying clubs, airlines or aviation authority administrations to improve training and flight procedures.

[✉]Contact: frederic.dehais@isae-supaero.fr

Article №6 · Pages 54–60 · https://doi.org/10.3929/ethz-b-000407652

Whereas eye tracking has been extensively used for aviation psychology research since the pioneering work of Fitts [14], there is still a long way before such an approach could be implemented and systematized for training purpose [7, 11]. One reason is that experiments generally involved small cohorts of pilots and there is an obvious lack of data to constitute reference models of nominal ocular behavior. Another issue is that most of these experiments have been conducted in simulated conditions thus limiting the interpretability of the data. To date, and to the best of the author's knowledge, few studies have collected eye movements under actual flight conditions [15–18].

The objective of the present study was to evaluate the potential of eye-tracking in real flight conditions for training purposes. Seven pilots were asked to perform a flight scenario involving basic flight maneuvers and two traffic patterns. Their ocular behaviors were recorded using a head-mounted eye-tracking device. We first performed descriptive analysis related to the dwell times over the different flight instruments with regards to the different flight phases. Three flight instructors were then asked to subjectively assess their distribution of attention if they had to fly similar scenarios. Participant's eye-tracking data were confronted with operational procedures and compared with the flight instructors' subjective data. We eventually used classification techniques over the eye-tracking data collected during the first traffic pattern to predict the different flight phases of the second traffic pattern. The motivation was to show that such techniques could be used to classify pilots' ocular performance.

2 MATERIAL AND METHOD

2.1 Participants

Seven pilots were recruited among the students of the ISAE-SUPAERO engineering school to participate in the study (all males; 25-35 years old, with 80-250 flight hours experience). All had a normal or corrected-to-normal vision and no history of neurological or psychiatric disorders. The study was approved by the European Aviation Safety Agency (2403 2424 2487–EASA 0010011661) and all participants gave their informed written consent. Three experienced flight instructors participated in a subjective dwell time survey.

2.2 Aircraft

We used the ISAE-SUPAERO experimental Robin DR400 light aircraft. This four-seated aircraft is powered by a 180HP Lycoming engine and equipped with classical gauges, radio navigation equipment, and actuators such as rudder, stick, thrust, and switches to control the flight.

2.3 Experimental scenario and protocol

The flight scenario consisted of performing basic maneuvers such a steep turn, climb and descent and two consecutive standard traffic patterns at Lasbordes airfield (Toulouse). The scenario lasted around 40 minutes. The steep turn was a left 360-degree turn with a 45-degree bank angle while maintaining altitude, speed, and bank. The descent was initiated at 2500 feet to reach an altitude of 1500 feet and consisted of maintaining a constant negative vertical speed (500 feet per min) and a constant speed of 150 km per hour. The climb was following the descent and consisted of adapting the pitch to maintain a constant speed of 150km per hour until an altitude of 2500 feet was reached. The pilots were then asked to achieve two consecutive traffic patterns. Each traffic pattern, according to the standards of visual flight rules (VFR), is divided into five flight phases—the upwind take-off leg, the crosswind leg (2000 ft), the downwind leg, the base leg and the final landing (see Fig. 1).

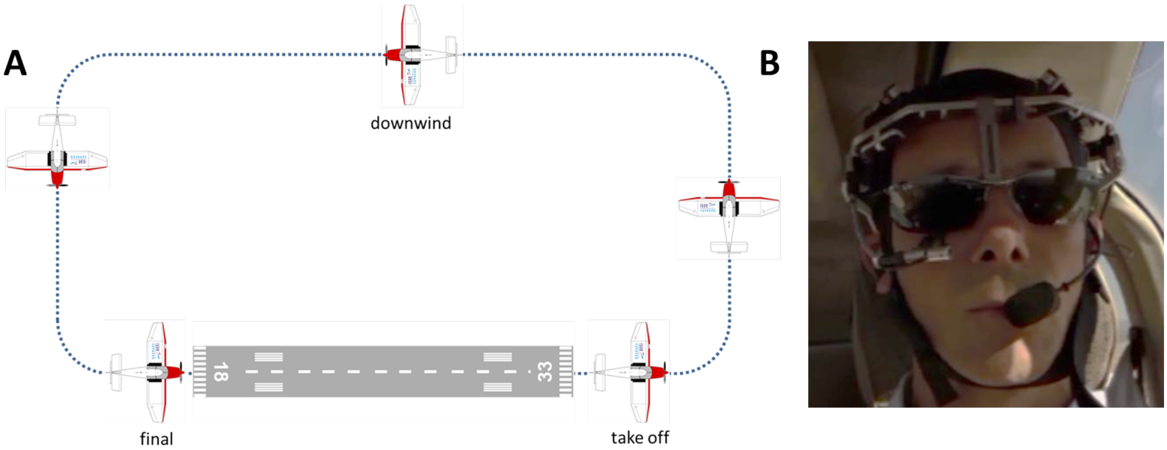

Figure 1: A) Traffic pattern at Lasbordes airfield. B) Pilot equipped with a Pertech eye tracker.

2.4 Measurements

Eye movements were recorded using a head-mounted Pertech eye-tracker (Pertech, Mulhouse, France). This device has a 0.25° accuracy and a 50 Hz sampling rate. The data from a field view camera and a monocular infrared-sensitive eye camera mounted to the head support were recorded by the computer unit installed in the aircraft baggage compartment. A calibration procedure was performed using five distant points on the instrument flight panel.

2.5 Analyses

2.5.1 Objective and subjective dwell times. The participants' fixations were mapped on an image of reference, chosen for each participant so that it contained 8 Areas Of Interest (AOI): SPEED, ATT (attitude indicator), ALT (altimeter), VS (Vertical Speed), HDG (heading), TS (turn and slip indicator), RPM (engine power) and OTW (out-of-the window). We then computed descriptive analyses by computing mean percentage dwell fixation time over the 8 AOIs during the three main phases of the two traffic patterns (take off, downwind, and final) and the three basic maneuvers (360-degree steep turn, climb and descent).

The real flight eye-tracking data were used to obtain objective dwell times. We also asked three experienced flight instructors to describe ideal division of attention among instruments during each maneuver. We compared these subjective dwell times with the objective dwell times of our seven participants collected with the eye tracker.

2.5.2 Flight phase classification. Similarly to Scannella et al. [15], we trained a classifier to discriminate the three different phases (take-off, downwind and final) using eye metrics. To meet this goal, we epoched our data in non-overlapping and continuous 5s time-window and computed, for each leg, the mean dwell time on the SPEED, OTW, ATT, ALT and RPM AOIs. The other three AOIs were excluded as they did not receive a sufficient amount of fixations. We then trained a support vector machine classifier on the first traffic pattern (5 fold, 10-time cross-validation procedure) to discriminate the three phases of the second traffic pattern. Please not that the number of trials in each class was not equal as the three flight phases had not the same duration. This could bias our classification results so we computed and reported the mean balanced accuracy that is the average of the proportion corrects of each class individually.

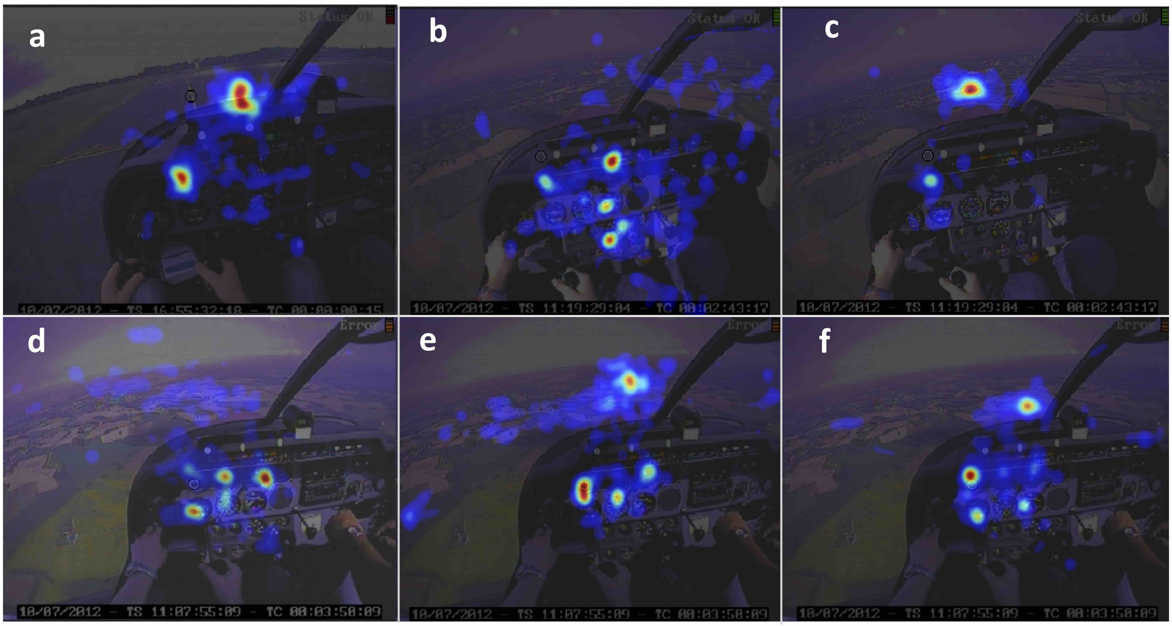

Figure 2: Heatmaps during a) take-off, b) downwind, c) final landing, d) 360-degree turn, e) descent and f) climb.

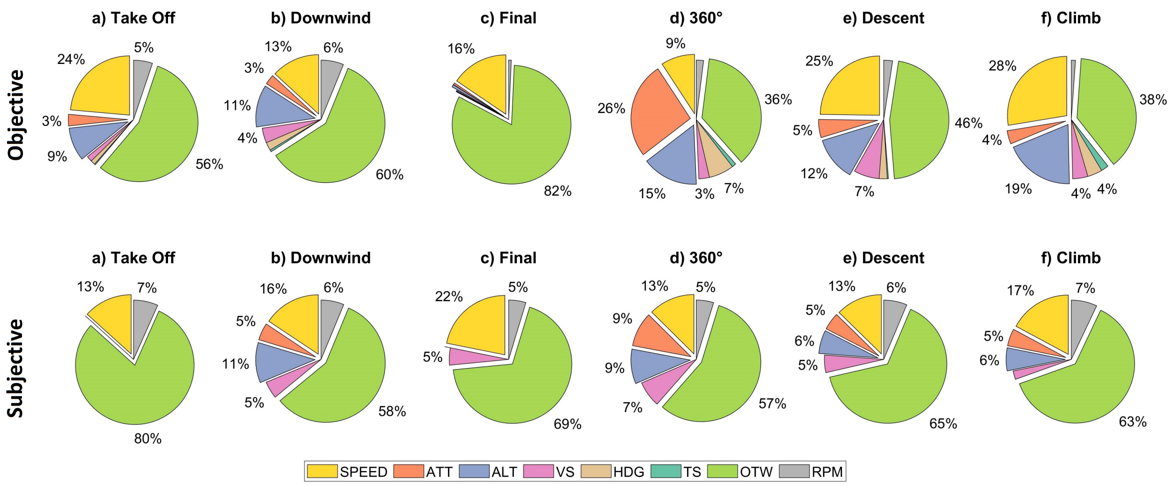

Figure 3: Objective distribution of dwell time (top row) and subjective distribution of attention (bottom row, as reported by three experienced instructors) over the 8 different AOIs during the different flight exercices. Labels for the AOIs with less than 3% of dwell time were ignored for the ease of reading.

3 RESULTS

3.1 Dwell time on the 8 AOIs across the different phases

Figure 2 represents typical heatmaps of one of the participant during the different flight phases.

Figure 3 shows the mean percentage of dwell time (top row) on the 8 different AOIs during the different flight phases of the traffic pattern, the 360 steep turn, the descent and climb.

3.2 Flight phases classification during the landing

The next table presents the mean balanced accuracy to classify the three flight phases of the second traffic pattern using eye tracking features collected during the first run.

Table 1: Mean balanced classification accuracy for the different features.

	Features used					
Leg	SPEED	EXT	ATT	ALT	RPM	All features
Take-off	61.0	53.0	50.0	50.0	50.0	49.3
Downwind	50.0	72.8	76.1	81.7	56.2	85.5
Landing	58.0	76.1	74.4	66.4	53.4	75.2
Mean accuracy	56.3	67.3	66.8	66.0	53.2	70.0

4 DISCUSSION

The objective of the present study was to investigate the potential of eye-tracking for training purposes. The originality of this work was that the data were collected in real flight conditions and discussed with flight instructors. We mainly focused our analysis on the qualitative measure of the mean dwell time over the different AOIs. One first observation is that the flight phases have an impact on the distribution of the pilot's attention over the different AOIs. For instance, pilots' attention is highly focused during the landing while it is more spread during the other phases (see Fig 2 and 3). More specifically, no AOIs received the same dwell time across the different phases as their relative relevancy highly depends on the maneuver to be performed. Despite these local differences, our findings indicated that the outside world corresponded to the most glanced AOI, whatever the flight maneuvers were. This result was expected as long our participants are only allowed to fly in Visual Flight Rules (VFR) conditions and are trained to spend most of their time "head up" to supervise the environment to anticipate potential threats (clouds, traffic, terrain). Another finding is that the speed indicator was the second most glanced AOI. This flight parameter is of paramount importance and has to be frequently monitored during the take-off, landing and climb to avoid stalls and during the downwind or the descent not to exceed speed limits. Inadequate monitoring of the speed is thought to be a precursor of poor performance [19] and fatal accidents [20]. Eventually, it is interesting to denote that a highly procedural flight phase such as the downwind leg is well reflected by the eye-tracking measures: the engine thrust power has to be set properly (6% of dwell time) to reduce the speed (13% of dwell time) so as to lower the flap, the altitude has to be carefully maintained (11 %) leading to slight attitude (3%) and vertical speed adjustment (4%) while spending much time on the outside world over (60%) to keep an eye on the runway. Taken together these findings indicated that eye-tracking appears to be a reliable sensor to be used in actual flight conditions as previously demonstrated by [17].

It is interesting to compare these numbers with a dwell time that the flight instructors would expect their trainees to have on the flight deck. There are some obvious discrepancies between our participants' scan patterns and the ideal ones defined by our three experts for each maneuver, to the exception of the downwind phase. The main difference was related to the time spent on the outside view that was much lower than thought by the instructors. Specifically, these latter were particularly surprised to find out that the participants spent most of their time head down (64%) when performing the 360° steep turn. During this latter maneuver, pilots are taught to take cues from the outside world (e.g. sighting the horizon line) to fly their aircraft as this steep maneuver can lead to a potential collision with traffic. However, the flight instructors confirmed that steep turn remains difficult to control from a psycho-motor and proprioception point of views leading to focus on the flight parameters to maintain constant the attitude indicator, the altitude and the speed of the aircraft while ending the turn at the correct heading as indicated by the dwell time over these AOIs. These findings intrigued the flight instructors who concluded that the systematic use of eye-tracking would benefit to improve training.

One last analysis was related to the use of classification techniques. In the best case, the balanced accuracy reached 70% to discriminate the three legs of the second run when combining all the AOIs collected during the first run. This result is more than promising considering that it is a three-class problem and that we randomly mixed the data from all our participants (i.e. intersubject classification). It shows, similarly to [15], that eye metrics are quite robust over time and do not differ across pilots. We used 5s-windows, which means that it might be technically possible to assess pilots' ocular behavior with regard to a specific flight phase in an online fashion. Such a goal would be reachable by collecting more data with expert pilots to build a reference model that could be learned by classification algorithms and used to classify trainees' performance.

This study has several limitations as our sample of pilots and experts were small data thus preventing us from building an "attentional" model of a good vs bad student pilot. We also only reported mean dwell time that does not account for the dynamic of attention and future study should focus on the temporal aspect and the analysis of scan patterns. The incorporation of such metrics could help to improve classification accuracy and using exponential moving averaged so as to meet aeronautical standards in terms or reliability. Eventually, one has to consider that measuring eye movements in real flight conditions remains challenging as the infrared radiation from the sun can degrade the tracking of the pupil. This particularly true in summer during warm weather. To that end, our experiments were collected during the morning and signal quality was higher during high-level cloudy conditions. One possibility would be to integrate solar glass canopy to shield against such contamination for future operation. Nonetheless, we believe that this first study offers relevant perspectives for the consideration of eye-tracking for flight safer and more efficient flying [11, 21].

ACKNOWLEDGEMENTS

This study was supported by the AXA Research Fund and ANITI ANR-19-PI3A-0004 .

REFERENCES

[1] Emilien Dubois, Colin Blättler, Cyril Camachon, and Christophe Hurter. Eye movements data processing for ab initio military pilot training. In *International Conference on Intelligent Decision Technologies*, pages 125–135. Springer, 2017.

[2] Christopher D Wickens and Amy L Alexander. Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology*, 19(2):182–199, 2009.

[3] Frédéric Dehais, Raphaëlle N Roy, and Sébastien Scannella. Inattentional deafness to auditory alarms: Inter-individual differences, electrophysiological signature and single trial classification. *Behavioural brain research*, 360:51–59, 2019.

[4] Frédéric Dehais, Helen M Hodgetts, Mickael Causse, Julia Behrend, Gautier Durantin, and Sébastien Tremblay. Momentary lapse of control: A cognitive continuum approach to understanding and mitigating perseveration in human error. *Neuroscience & Biobehavioral Reviews*, 2019.

[5] Christopher D Wickens and Frédéric Dehais. Expertise in aviation. 2019.

[6] Sylvia Peißl, Christopher D Wickens, and Rithi Baruah. Eye-tracking measures in aviation: A selective literature review. *The International Journal of Aerospace Psychology*, 28(3-4):98–112, 2018.

[7] Gal Ziv. Gaze behavior and visual attention: A review of eye tracking studies in aviation. *The International Journal of Aviation Psychology*, 26(3-4):75–104, 2016.

[8] JR Tole, AT Stephens, M Vivaudou, AR Ephrath, and LR Young. Visual scanning behavior and pilot workload. 1983.

[9] Andrew H Bellenkes, Christopher D Wickens, and Arthur F Kramer. Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. *Aviation, space, and environmental medicine*, 68(7):569–579, 1997.

[10] Chiara P Ryffel, Celine M Muehlethaler, Sandro M Huber, and Achim Elfering. Eye tracking as a debriefing tool in upset prevention and recovery training (uprt) for general aviation pilots. *Ergonomics*, 62(2):319–329, 2019.

[11] Vsevolod Peysakhovich, Olivier Lefrançois, Frédéric Dehais, and Mickaël Causse. The neuroergonomics of aircraft cockpits: the four stages of eye-tracking integration to enhance flight safety. *Safety*, 4(1):8, 2018.

[12] Vsevolod Peysakhovich, François Vachon, and Frédéric Dehais. The impact of luminance on tonic and phasic pupillary responses to sustained cognitive load. *International Journal of Psychophysiology*, 112:40–45, 2017.

[13] Leandro L Di Stasi and Carolina Diaz-Piedra. Eye movements research in aviation: past, present, and future. In *Improving Aviation Performance through Applying Engineering Psychology, Advances in Aviation Psychology (Vol. 3)*, pages 143–162. CRC Taylor and Francis Group Boca Raton, FL, US, 2019.

[14] Paul M Fitts, Richard E Jones, and John L Milton. Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical engineering review*, 9(2):24–29, 1950.

[15] Sébastien Scannella, Vsevolod Peysakhovich, Florian Ehrig, Evelyne Lepron, and Frédéric Dehais. Assessment of ocular and physiological metrics to discriminate flight phases in real light aircraft. *Human factors*, 60(7):922–935, 2018.

[16] Nicklas Dahlstrom, Staffan Nahlinder, Glenn F Wilson, and Erland Svensson. Recording of psychophysiological data during aerobatic training. *The International journal of aviation psychology*, 21(2):105–122, 2011.

[17] Frédéric Dehais, Mickaël Causse, and Josette Pastor. Embedded eye tracker in a real aircraft: new perspectives on pilot/aircraft interaction monitoring. In *Proceedings from The 3rd International Conference on Research in Air Transportation. Fairfax, USA: Federal Aviation Administration*, 2008.

[18] Glenn F Wilson, Brad Purvis, June Skelly, Penny Fullenkamp, and Iris Davis. Physiological data used to measure pilot workload in actual flight and simulator conditions. In *Proceedings of the Human Factors Society Annual Meeting*, volume 31, pages 779–783. SAGE Publications Sage CA: Los Angeles, CA, 1987.

[19] Frédéric Dehais, Julia Behrend, Vsevolod Peysakhovich, Mickaël Causse, and Christopher D Wickens. Pilot flying and pilot monitoring's aircraft state awareness during go-around execution in aviation: A behavioral and eye tracking study. *The International Journal of Aerospace Psychology*, 27(1-2):15–28, 2017.

[20] Douglas D Boyd. Causes and risk factors for fatal accidents in non-commercial twin engine piston general aviation aircraft. *Accident Analysis & Prevention*, 77:113–119, 2015.

[21] Nichakorn Pongsakornsathien, Yixiang Lim, Alessandro Gardi, Samuel Hilton, Lars Planke, Roberto Sabatini, Trevor Kistan, and Neta Ezer. Sensor networks for aerospace human-machine systems. *Sensors*, 19(16):3465, 2019.