# Flight Eye Tracking Assistant (FETA): Proof of Concept

Christophe Antony Lounis, Vsevolod Peysakhovich, Mickaël Causse

# Flight Eye Tracking Assistant (FETA): Proof Of Concept

Christophe Lounis[1], Vsevolod Peysakhovich[1], Mickaël Causse[1]

[1] ISAE-SUPAERO, Université de Toulouse, FRANCE

**Abstract.** Accident investigations show that piloting errors (e.g., incorrect trajectory) often result from an inadequate monitoring of the cockpit instruments. Recent improvements of the eye tracking technology now allow a reliable and rather accurate recording of eye movements in ecological environments. The present study investigates how the integration of eye tracking in the cockpit could help pilots performing an efficient surveillance of their instruments. We developed FETA, an embedded system that evaluates online the visual monitoring of the cockpit. The system compares the current visual scan of the pilot with a database of "standard" visual circuits established thanks to eye-tracking recordings from 16 airlines pilots. If the current visual scan deviates too much from the database, e.g., the speed is not fixated during a too long period, FETA emits a vocal alarm to reorient attention. This paper presents the development of FETA and its preliminary evaluation with 5 airlines pilots. During an approach-landing phase in flight simulator; we assessed the impact of FETA on situation awareness, cognitive resources, flight performance, and visual scans. Results showed that FETA system efficiently redirected attention toward critical flight instruments. However, improvements must be performed to satisfy with operational requirements. For example, it seems important to take also into-account flight parameters in order to limit unnecessary alerts.

**Keywords:** Eye-Tracking · Aviation · Human Factors · Human Computer Interaction · Neuroergonomics · Flying Assistant · Assistive Technology.

## 1 Introduction

Over the past 50 years, continuous technical and technological improvements in commercial aviation made it the safest modes of transportation [1]. Progress in cockpit systems and aircraft design [2], in pilot training, in flight crew and air-traffic control procedures, are still essential to maintain a low accident rate despite an ever-increasing traffic [3]. Nevertheless, accidents always occur and a large part of them involve human error (approximatively 60 to 80 percent of accidents), as shown Figure 1. One critical solution provided by the industry to reduce crew's workload [6] and to deal with human errors [4,5] has been the introduction of automation. However, automation also shifted the role of the crew from controllers to supervisors [7]. Unfortunately, automation is not always fully understood and nor correctly monitored [8]. It can induce complacency, overconfidence, and airline pilots sometimes rely too much on it [9]. Therefore, increasing the pilot's ability to stay in the loop, in particular by promoting an appropriate monitoring of the cockpit instruments, is a main current safety challenge [10]. This is particularly true during the approach and landing phases, critical periods of a flight, in which safety margin and tolerance on flight parameters deviations are very low [11].

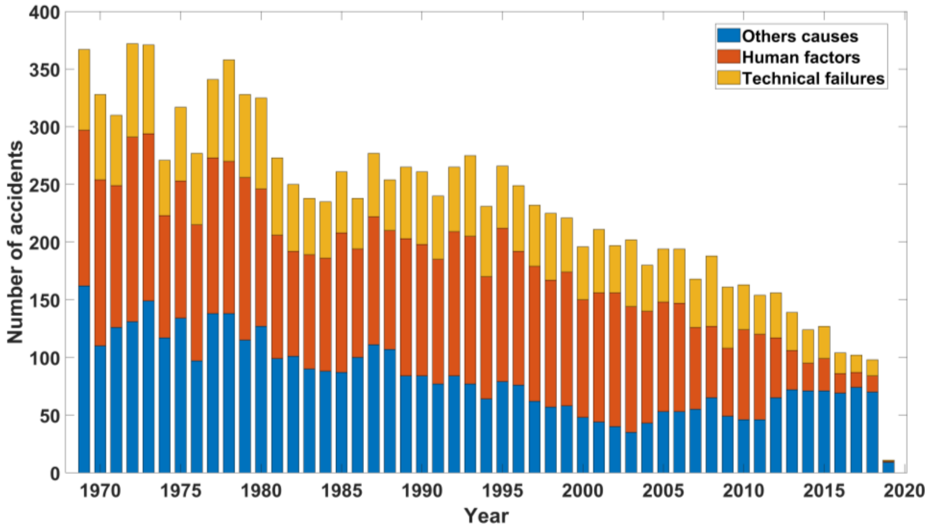


**Fig. 1.** Aeronautical accidents from 1969 to 2019 with human factors, technical failure and others causes as contributory factors. These data were retrieved from Bureau of Aircraft Accidents Archives (www.baaa-acro.com).

A report of the active pilot monitoring working group published by the Flight Safety Foundation [11] investigated 188 cases involving monitoring issues leading to accidents. They find out that 66% of monitoring errors occurred during dynamic phases of flights (e.g., climb, descent, approach, and landing). They also identified the failure to cross-check instruments as a major cause of those monitoring errors that resulted in excessive deviations of the flight parameters (e.g., altitude, trajectory, or speed deviation). Recently, the Federal Aviation Administration (FAA) published a final training rule that requires enhanced pilot monitoring training to be included into existing air careers training programs [12]. Furthermore, the French accident investigation bureau (Bureau Enquêtes et Analyses, BEA) suggested to analyze pilots' monitoring with eye tracking to improve piloting procedures [13]. In this sense, a recent paper described the different ways in which eye tracking could be used to assist commercial pilots during the flight [14].

Based on the latter studies, we propose the FETA system. It compares the current visual scan of a pilot with a database of "standard" visual circuits. If the current visual scan deviates too much from the database (e.g., the speed is not fixated during a too long period), FETA emits a vocal alarm (e.g., "*check speed*"). The current paper describes the development and the evaluation of the FETA (Flight Eye Tracking Assistant) system [15]. In particular, we evaluated the impact of FETA on situation awareness [16], subjective workload, flight performance, and visual scans.

## 2 FETA system development

The main purpose of the FETA system is to warn the pilot when he looks not sufficiently at an instrument. In order to calibrate the "not sufficiently", the threshold beyond which visual scans become "abnormal", we built a database of standard visual circuits in the cockpit with a sample of 16 airline pilots. They performed approach-landing phases in flight simulator while their eye movements were recorded. We also ensured that their flight performance remained in the standard safety thresholds.

### 2.1 Participants

Sixteen male professional airline pilots (ATPL: Airline Transport Pilot License or CPL: Commercial Pilot License) volunteered to participate in this study. Their mean age was 34 years old (range: 23-59). Their total flight experience ranged from 1,600 to 13,000 hours (M = 4,321.73 hrs, SD = 2,911.41 hrs). They were not paid for their participation. They had normal or corrected-to-normal vision. The experiment was approved by the Research Ethics Committee (CER, n°2019-131).

### 2.2 Procedure

Each participant signed a consent form and provided demographic information, their flight qualifications (type of aircraft), and their total flight experience hours. Pilots were briefed on the study and receive instruction about the flight scenario and the goal of this experiment. They filled a fatigue questionnaire. Next, pilots were installed in the flight simulator and were submitted to the eye-tracking calibration procedure. Participants took the captain position and performed a training consisting in two approach-landings scenarios in order to familiarize themselves with the flight simulator. Then, they performed the two experimental approach-landings scenarios.

### 2.3 Flight simulator

The study was conducted in the PEGASE (Platform for Experiments on Generic Aircraft Simulation Environment) flight simulator of the ISAE-SUPAERO (Toulouse, France), illustrated in figure 2. It simulates an Airbus A320 with a glass cockpit. The simulator includes pilots' seats, sidestick controllers, throttles, trim wheels, and rudder pedals.

**Fig. 2.** The PEGASE flight simulator used during FETA development and assessment.

### 2.4 Eye-tracking measurements

Eye tracking data was collected with a Smart eye System embedded in the cockpit. The Smart eye System consists of 5 deported cameras (0° - 2° of accuracy), with a sampling frequency of 60 Hz. Furthermore, the cockpit has been divided into several Areas of Interests (AOIs), as presented in figure 3. They correspond to the main flight displays. These AOIs are used by the FETA system to evaluate online current visual scans. We also used these AOIs during the human factor evaluation to examine the impact of FETA on visual scans. The threshold for detecting a fixation on an AOI was set at 100 ms [17].

**Fig. 3.** Cockpit Display with AOIs and Sub-AOIS: (**1**) Primary Flight Display (PFD), (**2**) Navigation Display (ND),(**3**) Electronic Centralized Aircraft Monitoring (ECAM), (**4**) Out of Window (OTW), (**5**) Flight Control Unit (FCU), (**6**) Flight Mode Annunciator (PFD.FMA) , (**7**) Speed Tape (PFD.SPD), (**8**) Attitude Indicator (PFD.ATT), (**9**) Vertical Speed Tape (PFD.VS), (**10**) Heading Tape (PFD.HDG), (**11**) VOR tag reading area in ND (ND-zone).

### 2.5 Experimental conditions

The 16 pilots performed two times the same flight scenario. The flight scenario consisted of a manual approach-landing task to Toulouse-Blagnac Airport Runway LFBO 14R. Flight began at coordinates 1.2159 longitude and 43.7626 Latitude. During the scenario, the pilot had to comply with some specific instructions. In particular: maintain a vertical speed between +500 ft/min and -800ft/min, a speed of 130 knots, and a heading of 143°.

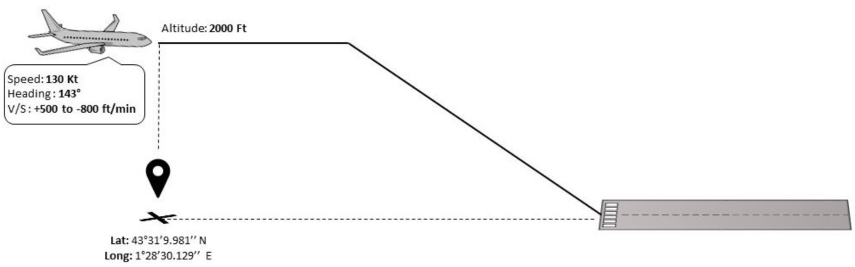


**Fig. 4.** The landing scenario with the flight parameter values that pilots had to maintain.

## 2.6 Flight parameters

Firstly, we checked flight performance of the pilots, assuming that correct flight performance is likely correlated to an efficient cockpit monitoring. Figure 5 shows the mean flight parameters deviation for vertical speed, speed and heading during the landing task. Flight performance of each pilot was adequate. Average vertical speed was in the correct range, and average speed and average heading were very close to the target values.

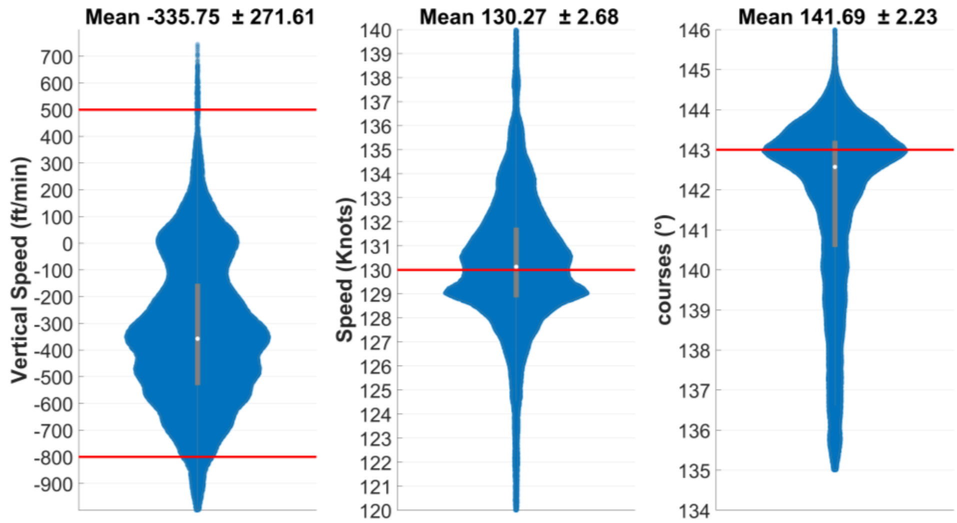


**Fig. 5.** Violin plot of flight parameters deviations during the landing task. The red lines correspond to target values, as given by the experimenter before the flight scenarios. N = 16.

## 2.7 Visual Behavior Database and notification threshold

The Visual Behavior Database (VBD) has been established with the eye recordings made on the 16 pilots that performed the two approach-landing scenarios. Mean non-dwell times were calculated for each AOI. While dwell times represent the time during which an individual gaze inside an AOI [18], non-dwell times correspond to the period of time during which an individual does not look at an AOI, see Figure 6.

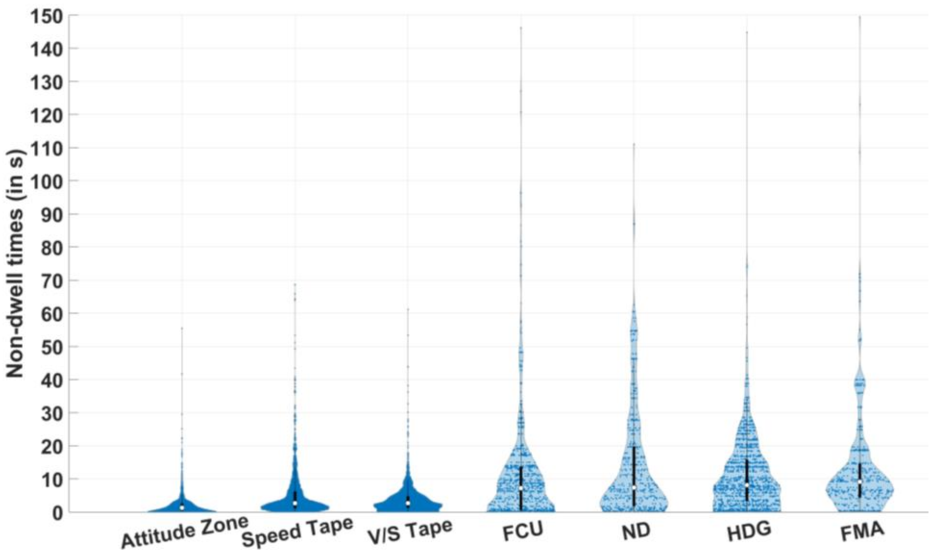


**Fig. 6.** Violin plot of non-dwell times during the landing task for the main AOIs. N = 16.

We used the "non-dwells times" of the 16 expert pilots as the metric for the FETA notification threshold. More precisely, the thresholds consisted of the averages of the non-dwell time for each AOI plus a standard deviation, as presented in (1).

$$\varphi_{threshold} = \mu_{NDT} + \sigma_{NDT} . \quad (1)$$

This metric indicates the maximum non-dwell time tolerance for each AOI (i.e., beyond which an insufficient monitoring is diagnosed).

## 2.8 FETA interface

Besides the Visual Behavior Database, the eye tracking system, and the vocal alarms, FETA also has an application permitting to visualize the activity from outside the cockpit. Coded in C#, the FETA interface has many features shown in Figure 7.

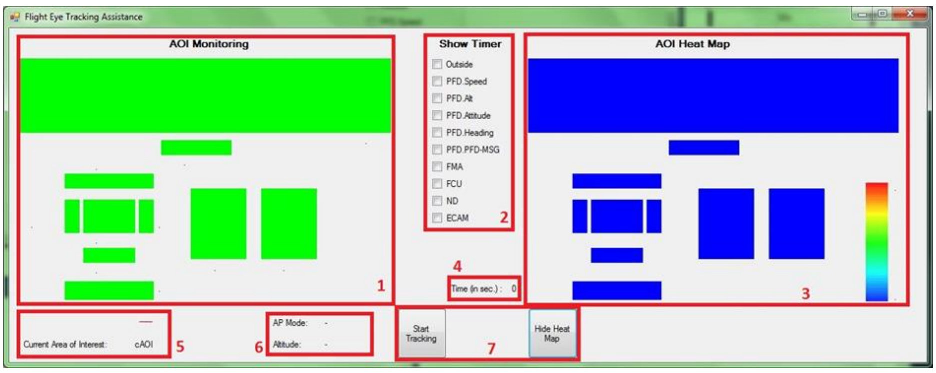


**Fig. 7.** FETA interface with its 7 different features.

The features of FETA are:

1. AOI Monitoring Panel (on the left of Figure 7)

It shows the state of each AOI. The color turns from green to blue when the AOI is not monitored enough according to the VBD.

2. Show timer (center of Figure 7)

User can tick the tick boxes of any of the AOIs to see the timer of each AOI. This timer shows the elapsed duration since the last monitoring (in seconds).

3. AOI Heat Map Panel (on the right of Figure 7)

This heat map panel indicates the proportion of fixation times on the AOI since the beginning of the flight.

4. Timer (center of Figure 7)

This feature shows the elapsed time until the beginning of the simulation in seconds.

5. AOI Text Alert and Current Area of Interest Annunciator (at the bottom left of Figure 7)

The AOI Text Alert will show the name of the AOI that needs to be monitored. If more than one AOI needs to be monitored, this alert will only show the name of the AOI with the highest priority. The Current Area of Interest Annunciator shows the currently monitored AOI.

6. Flight Parameter Indicators (at the bottom left of Figure 7)

This feature shows the several flight parameters that affect the dynamic of the database.

7. Start/Stop Tracking Button and Show/Hide Heat Map Button (centred at the bottom of Figure 7)

The Start/Stop Tracking Button starts or stops FETA, while the Show/Hide Heat Map Button shows or hides the AOI heat map.

8. Audio Alarm (cannot be shown)

FETA will emit an audio alarm that corresponds to the AOI Text Alert (e.g. "*check speed*").

## 3 FETA system assessment

The second part of this paper focuses on the evaluation of the FETA system. In particular, its impact on mental workload, situational awareness, flight performances, and cockpit monitoring. As a preliminary assessment, five pilots were submitted to three different scenarios varying in terms of monitoring difficulty.

### 3.1 Participants

Five male professional pilots (ATPL, CPL) volunteered to participate in this study. They had normal or corrected-to-normal vision. Their mean age was 39 years old (range: 33-50). Their total flight experience ranged from 2,500 to 8,500 hours (M =

3,176 hrs, SD = 2,645 hrs). Pilots were not paid for their participation. The experiment was approved by the Research Ethics Committee (CER, n°2019-131).

### 3.2 Procedure

Procedure was essentially the same than during the FETA calibration, except that the new 5 pilots performed four additional landings. During this evaluation, FETA auditory notifications (in case of abnormal monitoring) were restricted to three instruments: speed, vertical speed, and heading. These instruments were chosen because they corresponded to the flight parameter values that pilots had to maintain. Possible auditory alarms emitted by FETA were: "*check speed*"; *check vertical speed*", *check heading*".

### 3.3 Apparatus

This experiment also took place in the PEGASE flight simulator, using the same eye tracking system.

### 3.4 Experimental conditions

Pilots performed two times three different randomized landing scenarios. The first scenario (Scenario 1) was identical to the one performed by the pilots for the building of the VBD. In the second and the third scenarios, we increased monitoring difficulty. During these two scenarios, pilots were asked to read aloud the distance between the aircraft and a specific radio beacon (information displayed in the ND-zone) either every 0.5 Nm (scenario 2) or every 0.2 Nm (scenario 3). The pilots had to comply with the same speed, vertical speed, and heading constraints than during the VBD building. At the end of the simulation, pilots filled out 2 subjective questionnaires: situational awareness measures using SART [19] and workload Instantaneous Self-Assessment [20], which is a subjective scale ranging from 1 to 5. The latter allows assessing overall workload. After the flight scenarios, open interviews were conducted to garner the various opinions of the pilots according to the system.

### 3.5 Human factors assessment

Due to the low number of participants, we only present descriptive statistics for subjective assessments, and flight performance. However, eye tracking data allows us to use inferential statistics regarding the comparison with and without FETA. In particular by taking into account the difficulty of scenarios (1, 2, 3) as a covariate.

*Subjective results*

Figure 9 shows the SART results. A higher SART score indicates a better situational awareness. On average, FETA seemed to disturb the situational awareness when flying context was easy (scenario 1 and 2), but it tended to be the opposite when flying context was more complex (scenario 3). As presented in figure 9 (at right), ISA workload indicator did not show marked difference with or without the FETA system.

However, in an easy flying context (scenario 1), the FETA system seems to induce more workload and this trend is reversed when flying context is more difficult (scenario 2 and 3).

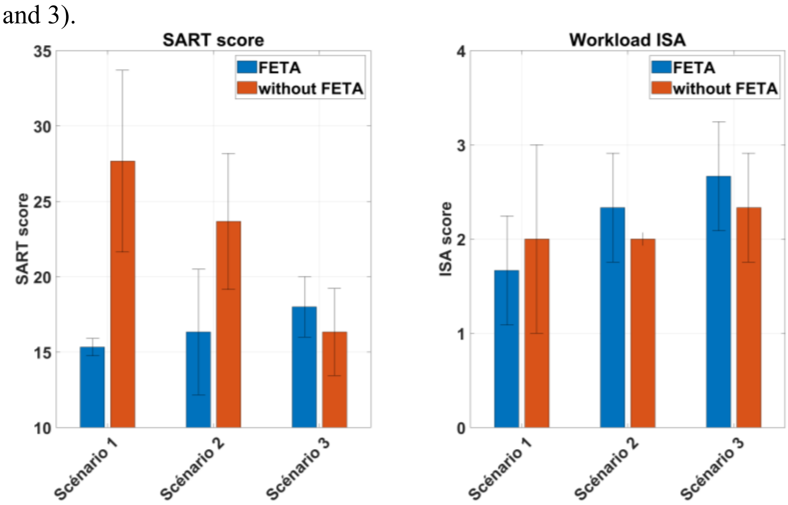


**Fig. 9.** Left: SART results (higher the values, better the situational awareness); Right: ISA results (lower is the value and lower is the subjective workload). All three scenarios with and without the FETA system are showed. N = 5.

*Flight performance results*

The figure 8 shows flight parameters deviations. During the easy scenario (scenario 1), pilots had higher speed deviations with FETA than without. Concerning the heading in the difficult condition (scenario 3), pilots had on average lower heading deviations with the FETA system than without.

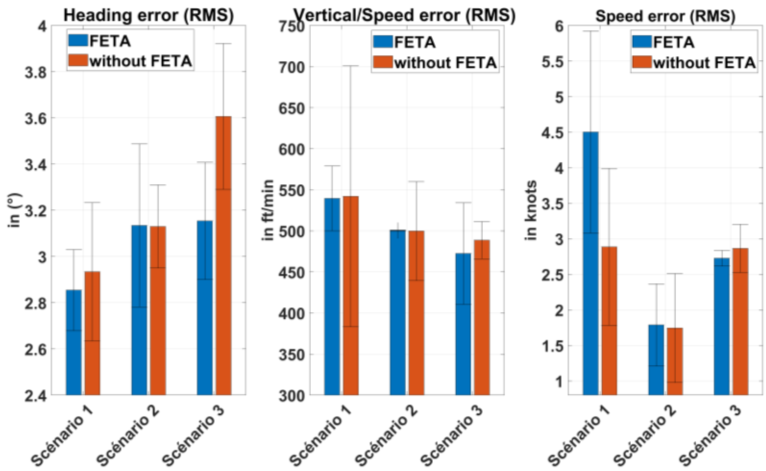


**Fig. 8.** Root Mean Square (RMS) of the flight parameters for each scenarios with and without FETA (the higher the value, the lower the performance). N = 5.

*Eye tracking results*

Figure 10 shows the percentage dwell times on each AOI for all scenarios with and without FETA system. The Wilcoxon-Mann-Whitney nonparametric test shows a significant effect ($p<0.05$) of FETA vs. without FETA condition on the AOIs according to speed, vertical speed, heading, flight mode annunciator and out the window.

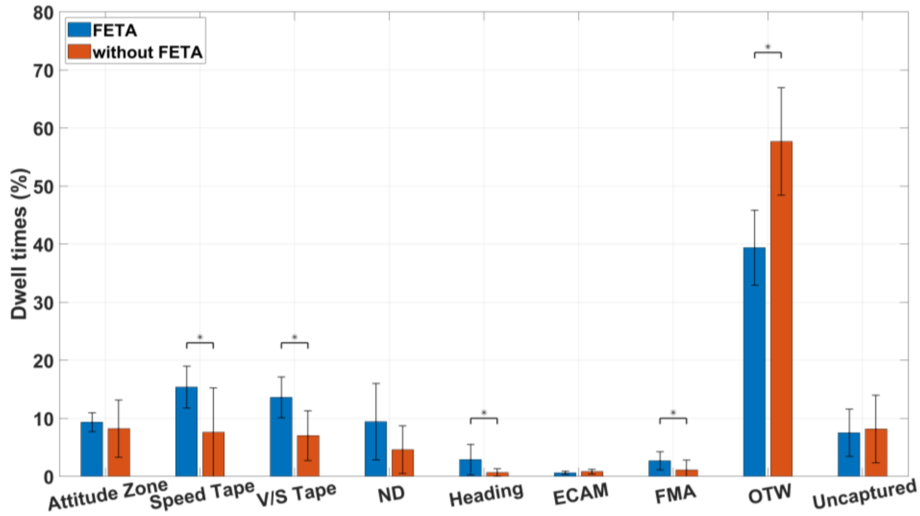


**Fig. 10.** Bar plot of percentage dwell times on each AOI. All three scenarios with and without the FETA system are showed. N = 5. (*$p<0.05$, Wilcoxon Mann-Whitney test).

## 4 Discussion and conclusions

The purpose of this study was on the one hand to present the concept and the development of a flight eye-tracking assistant (FETA) calibrated thanks to eye-movement recordings from 16 airline pilots. On the other hand, we also proposed a user-centered evaluation (e.g., situation awareness, mental workload) of the first version of this assistant together with an assessment of its impact on the cockpit monitoring. This evaluation was performed with 5 other airline pilots.

Overall, this first version of FETA demonstrated mixed results. First, results showed that there was no clear improvement of the flight parameters that had to be maintain during the landing (speed, vertical speed, and heading). There was an increased speed deviation during the easier landing and on the contrary an improvement of heading accuracy during the most difficult landing scenario. Consistently, subjective results tend to show that FETA was not detrimental only when flight scenario was difficult. In particular, situation awareness seemed slightly improved by FETA in the scenario 3. Eye tracking results were more favorable to FETA, with an increase of the time spent on some instruments subjected to the FETA audio notification in case of insufficient visual consultation. In presence of FETA, pilots checked more often the speed, the vertical speed, and the heading. This additional time gazing these instruments impacted the time spend on the window. Most likely, FETA was efficient to redirect attention toward the critical flight instruments thanks to the vocal alarm triggered when visual circuit deviated too much from the database. Despite this positive result, our experiment shed to light several issues that should be addressed in the future. Open interviews with the pilots allowed revealing some areas of improvements. For example, the use of the auditory modality is not necessary the best one. This channel is already used by the synthetic voice in the cockpit, and also during the exchanges between pilots and air traffic control. To overcome this problem, other notifications methods could be explored, such as visual and/or haptic modalities. Another important improvement would be to both integrate flight parameters values and eye movements in FETA. Indeed, it would be more appropriate to trigger notifications when both visual scans and flight parameters deviates too much from standards. For example, when speed decline too fast etc. This would help avoid triggering spurious notification (useless auditory notification from FETA), which was one of the main problems raised by the pilots during the debriefing. More generally, the FETA system should consider other eye tracking metrics when considering the landing task; for example, it could analyze the visual patterns (transitions between AOIs, not only the fixation on each AOI) and correct them when they deviate from established standards, using artificial intelligence. Furthermore, FETA could take into account other flight phases, automatically identified considering the flight data (e.g., altitude, speed, flight mode…). Then, this would enable to adapt eye-tracking metrics to the given flight phases. For example, cockpit monitoring is much less intense during the cruise, but this phase is more prone to drowsiness or fatigue. FETA could integrate metrics based on the percentage of eye closure [21] or considering the frequency of eye blinks [22]. Future study should consider these improvements and assessing FETA during complex flight phases with a higher number of pilots.

**Acknowledgments**

This work was supported by a chair grant from Dassault Aviation ("CASAC", holder: Prof. Mickaël Causse)". The Authors thank the PEGASE simulator technical team and all the pilots who participated in this study.

## 5 References

1. Boeing. Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations, 1959–2014; Aviation Safety. Boeing Commercial Airplanes: Seattle, WA, USA, 2015. Available online: http://www.boeing.com/news/techissues/pdf/statsum.pdf (accessed on 2 November 2017).
2. Sarter, N. B., & Woods, D. D. (1995, October). "From tool to agent": The evolution of (Cockpit) automation and its impact on human-machine coordination. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 39, No. 1, pp. 79-83). Sage CA: Los Angeles, CA: SAGE Publications.
3. Button, K. (2017). *Wings across Europe: towards an efficient European air transport system*. Routledge.
4. Chialastri, A. (2012). Automation in aviation. In *Automation*. IntechOpen.

5. Lee, J.D. Human factors and ergonomics in automation design. In Handbook of Human Factors and Ergonomics, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2006; pp. 1570–1580
6. Lee, J. D. (2006). Human factors and ergonomics in automation design. *Handbook of human factors and ergonomics*, 1580-1596.
7. Mumaw, R. J., Sarter, N., & Wickens, C. D. (2001, May). Analysis of pilots' monitoring and performance on an automated flight deck. In *11th International Symposium on Aviation Psychology, Columbus, OH*.8.
8. Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human factors*, *52*(3), 381-410.
9. Wickens, C. D., Clegg, B. A., Vieane, A. Z., & Sebok, A. L. (2015). Complacency and automation bias in the use of imperfect automation. *Human factors*, *57*(5), 728-739.
10. Berberian, B., Somon, B., Sahaï, A., & Gouraud, J. (2017). The out-of-the-loop Brain: a neuroergonomic approach of the human automation interaction. *Annual Reviews in Control*, *44*, 303-315.
11. Sumwalt, R., Cross, D., & Lessard, D. (2015). Examining How Breakdowns in Pilot Monitoring of the Aircraft Flight Path. *International Journal of Aviation, Aeronautics, and Aerospace*, *2*(3), 8.
12. Flight Safety Foundation (FSF). A Practical Guide for Improving Flight Path Monitoring: Final Report of the Active Pilot Monitoring Working Group. November 2014. Available online: https://flightsafety.org/toolkits-resources/flight-path-monitoring/.
13. Federal Aviation Administration (FAA). 14 CFR Part 121 Qualification, Service, and Use of Crewmembers and Aircraft Dispatchers; Final Rule; Federal Register: Washington, DC, USA, 2013; Volume 78, Issue 218.
14. British European Airways (BEA). Study on Aeroplane State Awareness during Go-Around. 2013. Available online: https://www.bea.aero/etudes/asaga/asaga.study.pdf (accessed on 2 November 2017).
15. Peysakhovich, V., Lefrançois, O., Dehais, F., & Causse, M. (2018). The neuroergonomics of aircraft cockpits: the four stages of eye-tracking integration to enhance flight safety. *Safety*, *4*(1), 8.
16. Lounis, C., Peysakhovich, V., & Causse, M. (2018). Intelligent cockpit: eye tracking integration to enhance the pilot-aircraft interaction.
17. Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human factors*, *37*(1), 32-64.17. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
18. Russi-Vigoya, M. N., & Patterson, P. (2015). Analysis of pilot eye behavior during glass cockpit simulations. *Procedia Manufacturing*, *3*, 5028-5035.
19. Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002, March). Eye tracking in web search tasks: design implications. In *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 51-58). ACM.
20. Tattersall, A. J., & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, *39*(5), 740-748.
21. Taylor, R. M. (2017). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational Awareness* (pp. 111-128). Routledge.
22. Sommer, D., & Golz, M. (2010, August). Evaluation of PERCLOS based current fatigue monitoring technologies. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (pp. 4456-4459). IEEE.
23. Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European journal of applied physiology*, *89*(3-4), 319-325.