Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut Supérieur de l'Aéronautique et de l'Espace

Présentée et soutenue par :
Christophe LOUNIS

le mercredi 16 décembre 2020

Titre :

Monitor the monitoring : pilot assistance through gaze tracking and visual scanning analysis
Surveiller la surveillance : assistance aux pilotes à travers son regard et l'analyse des parcours visuels

---

École doctorale et discipline ou spécialité :

EDSYS : Informatique

Unité de recherche :

Équipe d'accueil ISAE-ONERA ACDC

Directeur(s) de Thèse :

M. Mickaël CAUSSE (directeur de thèse)
M. Vsevolod PEYSAKHOVICH (co-directeur de thèse)

Jury :

M. Franck MARS Directeur de recherche L2SN – Président
M. Graham BRAITHWAITE Professeur Cranfield University - Rapporteur
M. Mickaël CAUSSE Professeur ISAE-SUPAERO - Directeur de thèse
Mme Carolina DIAZ-PIEDRA Professeure Université de Grenade Espagne – Examinatrice
M. Julien DONNOT Enseignant-Chercheur CEAM - Examinateur
M. Andrew DUCHOWSKI Professeur Clemson University - Rapporteur
M. Stéphane DURAND Technical manager Dassault Aviation St-Cloud - Examinateur
M. Vsevolod PEYSAKHOVICH Ingénieur Chercheur ISAE-SUPAERO - Co-directeur de thèse

ABSTRACT

Au cours d'un vol, les pilotes doivent surveiller de façon rigoureuse des instruments de vol spécifiques (e.g., indicateur d'attitude, vitesse, altimètre, les paramètres moteurs) ainsi que l'environnement extérieur (e.g., repérer des éléments du relief au sol notamment lors de conditions météorologiques dégagées et à basse altitude) dans le but de mettre à jour leur conscience de la situation. Cette activité de surveillance (monitoring en anglais), critique durant les phases de vols dites évolutives (e.g., décollage, phase d'approche, et atterrissage), tient compte de l'observation et de l'interprétation de la trajectoire, des modes d'automatisation sélectionnés, et des systèmes utilisés à bord. Cela suppose une comparaison en temps réel entre les données affichées aux instruments et les valeurs attendues lors des phases de vols. Une surveillance appropriée du cockpit permet de prendre des mesures correctives (e.g., ajuster la trajectoire de l'avion lors de la détection d'une déviation observable sur la zone d'attitude) en temps opportun lors de la déviation d'un paramètre, garantissant ainsi un niveau de sécurité optimal. Cette activité de surveillance est structurée en séquence d'engagement et de réorientation de l'attention visuelle du pilote d'un instrument vers un autre. Les rapports d'accidents ont démontré que bien souvent les erreurs de pilotage, tels que des trajectoires incorrectes ou bien une survitesse à l'atterrissage, étaient la résultante d'une surveillance défaillante et/ou inadéquate des instruments du cockpit. L'enjeu de ce travail de recherche est d''améliorer la sécurité des vols notamment grâce à l'intégration d'un oculomètre et/ou la recherche de solution pour améliorer l'entrainement des pilotes en vue de réduire les erreurs de surveillance à bord. Les mouvements des yeux sont une fenêtre sur l'état cognitif du pilote et permettent de révéler les chemins attentionnels empruntés par l'opérateur à travers son parcours visuel. En lien avec les problématiques de surveillance dans les cockpits, nous avons élaboré un assistant de vol (FETA : Flight Eye Tracking Assistant) basé sur des comportements visuels d'experts (e.g., 24 pilotes avec plus de 1600 heures de vols). Cet assistant prévient les pilotes, grâce à une alarme auditive, quand ces derniers ne consultent plus suffisamment un instrument de vol en comparaison avec la base de données des mouvements oculaires experts. Une évaluation facteurs humains de cet assistant a soulevé plusieurs problématiques et a ouvert la voie à de nouvelles recherches concernant notamment l'utilisation de métriques reflétant aux mieux les parcours oculaires dans le cockpit et permettant précisément de quantifier l'attention visuelle d'un pilote à bord. Une partie de ce travail de recherche s'appuie sur une comparaison entre novices et experts dans le but de quantifier la marque de l'expertise. Une méthode utilisant le K coefficient appliqué aux AOI a permis de qualifier l'attention visuelle des pilotes (focal vs ambient) au cours de scenario en simulateur de vols présentant différentes charges d'activité visuomoteur. Des méthodes d'apprentissage machine basée sur des matrices de transition ont permis de classifier l'expertise avec une précision de 91%. Enfin, deux méthodes ont été utilisés pour qualifier et quantifier les stratégies visuelles dans le cockpit. Une méthode utilisant la Complexité de Lempel-Ziv (LZC), un algorithme de compression des données, permettant de mettre en lumière la complexité des sequences de balayage dans le cockpit. Ainsi que le méthode N-gram, a l'origine issue de la recherche sur les séquences ADN, permettant de quantifier les patterns communs au groupe d'expert et la longueur des patterns utilisés. Ces contributions

sont discutées à la lumière de l'amélioration d'un assistant basé sur des données oculométriques pour l'amélioration de l'apprentissage d'une part et pour éviter les problèmes de surveillances d'autre part. Finalement, l'évaluation du prototype FETA a soulevé des perspectives par rapport au choix de la modalité (e.g., auditive, visuelle, haptique) la plus pertinente concernant l'alerting.

Mots clés : Oculométrie ,Système de surveillance du pilot, Facteurs Humains, Strategies visuels, Interaction Homme-Machine, Apprentissage Machine ,Neuroergonomie, Mouvements Oculaires, Aéronautique.

During a flight, pilots must rigorously monitor specific flight instruments (e.g., attitude indicator, airspeed, altimeter, engine parameters) as well as the external environment (e.g., locate terrain features on the ground, especially in clear weather conditions by low altitude) to update their situational awareness. This monitoring activity, which is critical during dynamic flight phases (e.g., takeoff, approach phase, and landing), consist in observing and interpreting the flight path, the selected automation modes, and the systems used onboard. This involves a real-time comparison between the data displayed on the instruments and the values expected during the flight phases. Appropriate monitoring of the cockpit enables to take corrective measures (e.g., adjust the aircraft's trajectory when a deviation is detected in the attitude zone) promptly when a parameter is deviated, thus guaranteeing an optimal level of safety. This monitoring activity is structured in a sequence of engagement and redirection of the operator's visual attention from one instrument to another. Moreover, accident reports have shown that piloting errors, such as incorrect trajectories or overspeed during landing, are often the result of inadequate monitoring of cockpit instruments. The purpose of this research work is to improve the flight safety thanks in particular to the integration of an eye-tracker. Eye movements are a window on the pilot's cognitive state and reveal the attentional paths taken by the operator through his visual path. In connection with cockpit monitoring issues, we have developed a Flight Eye Tracking Assistant (FETA) based on expert visual behaviors (e.g., 24 pilots with more than 1600 flight hours). This assistant warns the pilots, thanks to an audible alarm, when they no longer sufficiently consult a flight instrument in comparison with the expert eye movement database. A human factors evaluation of this assistant raised several issues with such an assistant and paved the way for further research including metrics that best reflect the eye paths in the cockpit and the need to find the right metric to quantify a pilot's visual attention on-board. Part of this research work is based on a comparison between novices and experts in order to quantify the mark of expertise. A method using the K coefficient applied to the AOIs allowed to qualify the visual attention of the pilots (focal vs ambient) during a flight simulator scenario with different loads of visuomotor activity. Machine learning methods based on transition matrices allowed to classify the expertise with an accuracy of 91%. Finally, two methods were used to qualify and quantify visual strategies in the cockpit. A method using Lempel-Ziv Complexity (LZC), a data compression algorithm, to highlight the complexity of the scanning sequences in the cockpit. Another called N-gram method, originally derived from DNA sequence research, which quantifies the patterns common to the expert group and the length of the patterns used. These contributions are discussed in the light of the improvement of a flying assistant based on eye tracking data for improving learning on the one hand and avoiding monitoring problems on the other. Finally, the evaluation of the FETA prototype raised perspectives on the choice of the most relevant modality (e.g. auditory, visual, haptic) for alerting.

Keywords: Eye-tracking, Pilot Monitoring System, Human Factors, Visual Scanning, Human-Computer Interaction, Machine Learning, Neuroergonomics, Eye movements, Aviation.

Laboratoire de Neuroergonomie et de Facteur Humain, DCAS
Toulouse, France

REMERCIEMENTS

En somme, voici la partie la plus difficile à rédiger pour moi. Quelques mois que je m'essaie à rédiger cette partie en vain. Si vous lisez ces lignes, peut-être dévinerez-vous que finalement je touche à mon but. Après l'angoisse des remerciements passé, l'inspiration me revient comme un second souffle, une bouffée de mots frais qui oxygène la spontanéité dactylographique à la fin du marathon de la thèse.

Après ces flots de paroles superflues qui masquent la pauvreté de ma pensée. Je vais remercier dans un premier temps mes encadrants de thèse Mickaël Causse et Vsevolod Peysakhovich, je vous remercierais jamais assez pour votre disponibilité et vos conseils tout au long de la thèse. Si ma thèse s'est aussi bien déroulée c'est en partie grâce à vous. Votre encadrement a été une parfaite adéquation entre des moments d'encradrements rigoureux et des moments d'exploration qui m'ont reellement permis une grande liberté dans l'orientation de la thèse. Je sais à quel point ceci est rare et je vous remercie pour ça. Vous avoir tous les deux en encadrant était une parfaite complémentarité. Un spécial remerciement à mes rapporteurs et membres de Jury, Andrew Duchowski, Franck Mars, Graham Braithwaite, Carolina Diaz-Piedra, Stephane Durand et Julien Donnot. Merci pour la bienveillance et les retours qualitatifs dont j'ai pu bénéficier au cours des rapports et de la soutenance de thèse. J'en profite également pour remercier Jean-Paul Imbert et Cyril Camachon membres de mon comité de suivi de thèse c'était vraiment un plaisir de vous avoir tous les deux et d'évoluer durant 3 ans avec vos conseils et évaluations. Je remercie également Dassault Aviation mécènne et contributeur financier de cette thèse. Je remercie également Hélène Thirion-Lannister pour sa précieuse aide en fin de thèse.

Je souhaiterais remercier mes collégues du laboratoire de Neuroergonomie et de Facteurs Humains pour cette solidarité entre doctorant et post-doctorant je pense notamment à Emilie Jahanpour, Kevin Verdiere, Quentin Chenot, Almoctar Hassoumi, Rebai Soret. Merci pour toutes ces discussions, ces rires et parfois ces pleurs. Vous êtes mes amis, je serais toujours disponible pour vous j'ai hâte de voir ce que nous allons devenir. Un special remerciement aux chercheur(e)s du laboratoire de Neuroergonomie et de Facteurs Humains merci à Seb pour ta bonne humeur et les ballades moto, un grand merci à Fred pour les journées de surf et les cocktails parties, merci à Raphaelle pour m'avoir fait confiance sur l'enseignement, merci à Evelyne pour ton coup de main précieux en statistiques, merci à Bertille pour les discussions sur le ras-le-bol du système de la recherche française et autres sorties au bar. Merci à Nicolas pour ton aide sur les modéles d'IA et Machine Learning, Merci à Alex, Chama, Gregoire, Alban, Mélodie, Simon, Juan, Giuseppe, Jérome, Caroline (*2). Je remercie également les personnes que j'ai pu encadrer durant ce projet, je pense aux ingénieurs de recherches comme Dennis Citralokam, Caliani Hoarau, Thomas Vagneron, et les ingénieurs du CNAM-ISAE également. J'en profite également pour remercier Coralie Ludwig qui m'a donné l'idée du N-Gram dans la recherche de patterns. Un grand merci à Alma et Lily pour votre soutient dans les moments un peu difficile.

Enfin, je voudrais remercier ma famille, dans un premier temps mes colocs d'amour et autres amies qui gravitent autour: Mat, Kashia, Claire, Arno (*2), Virgule, Alban, Mathoug, Pookie, Gertrude, Toufika, Roberta, Estella, Thiphaine, Karolina et Xav, Jo et Flo, Lorenne, Marie, Kevin, Clement, Adrien, Coline j'ai vraiment grandi à vos côtés dans un milieu de babos qu'est le notre, nous avons pas toujours été d'accords d'un point de vue politique et philosophique. Je vous remercie infiniment parce qu'avec vous le débat à toujours été ouvert et (presque) bienveillant. Vous m'avez ouvert l'esprit sur tellement de domaine que je ne peux qu'être reconnaissant et esperer continuer à partager des moments avec vous que ce soit la musique, les ateliers raps dans le garage, la forge, les voyages à velo, les longs débats sur les sciences et le capitalisme/communisme et autres systèmes économiques, l'escalade, la fête chez les copains d'ici et d'ailleurs, les jeux de sociètés, le jardinage, le façonnage de pizza pour 30 personnes, les ateliers de repas gastronomiques pour 10 personnes, les repas à théme autour de la cuisine georgienne, libanaises, polonaises et j'en passe ... les soirées improvisées à danser dans le salon, les ateliers théatre et danse du regard avec Manu dans le garage, les ateliers jeu en VR dans le salon au lieu de coder mon XP de fin de thèse (déso les encadrants), le bout de code en python qu'on reçoit sur nos smartphone pour organiser le ménage, la sortie des poubelles et les courses au marché, les concerts de musique à la maison en acceuillant des musiciens qui jouent dans la rue et qui aujoud'hui font des tournées en France, les scénes de théatres improvisées dans le garage avec des troupes talentueuses.

Merci Maman et Papa ! Etant donné nos origines éthniques et le milieu social duquel nous venons j'étais pas "destiné" à me retrouver ici. Enfin... Je vous remercie pour m'avoir donné cette force de caractére et cette personalité que je dois en partie grâce à vous. Vous avez fait en partie ce que je suis, preparez-vous pour la suite.

Un énorme merci à mes amis d'enfance avec qui rien n'a changé depuis le temps, Jordy (Bald), Iyad, Yves, Loic (micha) et aussi tous les copains avec qui j'ai "partagé" la balle orange.

Seryl, Nadia et Manon merci à vous, vous êtes une source d'inspiration pour moi chaque jour et je sais que notre force et de reussir tout ce que nous voudrons bien entreprendre alors du fond du coeur merci d'avoir supporter mon caractére un peu spécial et merci pour votre soutient tous au long de cette thése.

A mes soeurs et frére ... d'ici et d'ailleurs...

CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE DES SIGLES ET ACRONYMES

EEG *Electroencephalography*

fNIRS *Functional near-infrared spectroscopy*

DMS *Driver Monitoring System*

FETA *Flight Eye Tracking Assistant*

FSF *Flight Safety Foundation*

FAA *Federal Aviation Administration*

NTSB *National Transportation Safety Board*

CA *Captain*

FO *First Officer*

Kt *Knot*

ILS *Instrument Landing System*

ft *feet*

AGL *Above Ground Level*

ATC *Air Traffic Control*

ATCO *Air Traffic Controllers*

FMS *Flight Management System*

PFD *Primary Flight Display*

PM *Pilot Monitoring*

AP *Autopilot*

FL *Flight Level*

FMA *Flight Mode Annunciator*

ECAM *Electronic Centralised Aircraft Monitor*

FD *Flight Director*

SA *Situation Awareness*

PARG Pertes de contrôle de la trajectoire en phase d'Approche lors de la Remise de Gaz

BEA Bureau Enquêtes et Accident

PNF *Pilot Non-Flying*

PF *Pilot Flying*

ECG *Electrocardiography*

ET *Eye Tracker*

PDT *Percentage Dwell Times*

INTRODUCTION

Aviation is an extremely safe mode of transport with, for example, less than one accident per million flights for its commercial component (Airplanes, 2016; Oster Jr, Strong, and Zorn, 2013). This low accident rate, which was 20 to 30 times higher in the 1960s, was largely made possible by technological advances in aviation systems (Billings, 2018). Technical problems are now the cause of only about 10% of accidents, leaving a significant portion to be attributed to human implication. The exact values vary over years and sources, but approximately 60-80% of accidents involve human error (Wiegmann and Shappell, 2017). Yet, to date, removing humans out of the system isn't an option since we still need human skilled on board, especially in case of failure or unexpected event. Far from removing the human operator, Neuroergonomics proposes to study the cognitive functions involved in complex activities in order to achieve better human-system interaction. In this context, the use of neuroimaging (e.g., EEG, fNIRS...) for the implementation of brain-machine interfaces seems promising to adapt systems to the mental or emotional state of the operator (Callan and Dehais, 2019; Dehais and Callan, 2019). However, and despite significant progress, the use of such techniques does not come without constraints and requires the wearing of helmets or electrodes, and the signal is still strongly impacted by some signal noises (e.g., muscular artifacts, electromagnetic pollution, etc.) (Aricò, Borghini, Di Flumeri, Sciaraffa, and Babiloni, 2018; Arico et al., 2017). An interesting alternative is based on the detection of the gaze and eye movements through the use of eye tracking. It offers the possibility of observing the visual attentional state of an operator, which is underlying behaviour and decision-making (Glaholt, 2014). For example, Lufthansa NetLine/Ops++ room's operators developed an eye tracking solution to ensure that notifications are seen, and the system dynamically adapts content based on gaze position and provides haptic feedback to the operator's wrists in the event of negligence (e.g., the omission of an important notification). Other use cases exist such as estimating fatigue and attention of operators of construction equipment (e.g., Caterpillar and Seeing Machines), or for example the DMS (Driver Monitoring Systems) developed by NXP (an American-Dutch semiconductor manufacturer), this driving monitoring systems includes a camera-based monitoring systems which provide a real-time evaluation of the presence and the state of the driver (e.g., Lenné, Roady, and Kuo, 2020). In the latter case, the system detects pupil size, blink frequency, the duration of moments where the eyes were closed or not directed towards the road. The driver is alerted using the haptic modality (e.g., chair vibration) when abnormal values are detected. More recently Hinfact, a startup company from Toulouse, relies on the work of the laboratory of Neuroergonomics and Human Factors to offer flight instructors the possibility to support their briefing session with objective data based on body and gaze tracking: distribution of attention over flight instruments: gaze patterns, drowsiness levels. Despite these advances, it is striking that the aeronautical world, despite being at the forefront in terms of safety and technology, does not yet integrate eye tracking in the cockpit. This is all the more legitimate when about half of human error's accidents include failure of the crew to properly monitor the cockpit. Indeed, pilots have to rigorously monitor specific flight instruments (e.g., attitude indicator, speed, altimeter, engine parameters) and the external environment (by clear weather) to build and update their situational awareness (Endsley, 1997). The monitoring activity, particularly critical during dynamic flight phases such as take-off and landing, includes the observation and interpretation of the flight path data, aircraft-configuration status, automation modes, and on-board systems. It supposes a real-time comparison of instrument data or system modes against the expected values according to the current phase of flight. An appropriate cockpit monitoring allows timely corrective actions in the event of a parameter deviation, ensuring an optimal level of safety. This monitoring activity is structured in sequences of attentional shifts from an instrument to another. Accident investigations show that piloting errors (e.g., incorrect trajectory, overspeed, altitude deviation, etc.) often result from inadequate monitoring of the cockpit instruments. One way to further enhance flight safety is by looking into the pilot's eyes. Eye movements are a window onto the pilot's cognitive state and eye tracking can reveal visual scanning strategies as a hint of these internal states. This Ph.D. work focuses on eye tracking techniques and metrics in the cockpit to avoid monitoring problems and enhancing flight safety. The manuscript is organized following this way. Through the monitoring issues in aviation, the Chapter 1 presents the context of this thesis. It explains what aircraft accident reports tell us about human limitations. It concludes by why eye tracking is a possible way to bring responses to monitoring issues. The Chapter 2 presents a state of the art of the eye tracking in the cockpit and its possible integration in modern cockpits. These two introductory Chapters are followed by our four contributions. Each contributory Chapter, except Chapter 6, has been fully or partially published or submitted to an international peer-reviewed journal or a conference. More specifically, the Chapter 3 presents a proof of concept of a first Flight Eye Tracking Assistant (FETA) based on eye movement of expert pilots. This assistant warns pilots when they do not watch sufficiently a flight instrument compared to expert eye movement database. This Chapter ends with an evaluation of this prototype from a human factor perspective and paves the way for further eye tracking measures to improve this assistant, it also raises the question of what the best alerting modality is. Chapter 4 focuses on the need to find other eye tracking metrics to qualify the visual behavior in the cockpit in order to enhance the flight assistant (FETA). This Chapter presents novel algorithms to assess dwell patterns complexity and to assess focal or ambient visual attention in the cockpit on the data from two experiments in flight simulators. Chapter 5 provides the reader with an overview of different eye tracking metrics to qualify these visual scanning strategies in the cockpit through a study involving novice and expert pilots. Chapter 6 tries to answer the question of how to warn pilots if visual monitoring behavior is inadequate. This Chapter focuses on a gaze-contingent study involving different modalities (auditory, visual, and haptic) to triggers notifications. This chapter present an experiment still in progress and will be completed after this Ph.D. Nevertheless, part of the results and the experimental protocol will be presented in this document. The Chapter 7 discusses these contributions in the light of eye tracking assistant to enhance learning on the one hand, and on-board, on the other hand, to avoid monitoring matters.

Part I

State of the art

CHAPTER 1

HUMAN ERROR IN AVIATION: MONITORING ISSUES AT THE ORIGIN OF ACCIDENTS

Sommaire

Commercial aviation is one of the safest transportation systems in the world, and the average annual accident rate is the lowest it has ever been (Safety, 2018). This impressive record is due to many factors, including improvements in aircraft systems, pilot training, flight crew skills, air traffic control procedures, etc. (Group et al., 2014). One of the characteristics of the aviation community that has contributed to this success is a commitment to continuously improve safety and operation. While since the beginning of the century safety has improved thanks to technical prowess. Progress in cockpit systems, aircraft design, 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. Nowadays, as said previously, a large part of accidents involves human error (Shappell et al., 2007) see Figure 1.1.

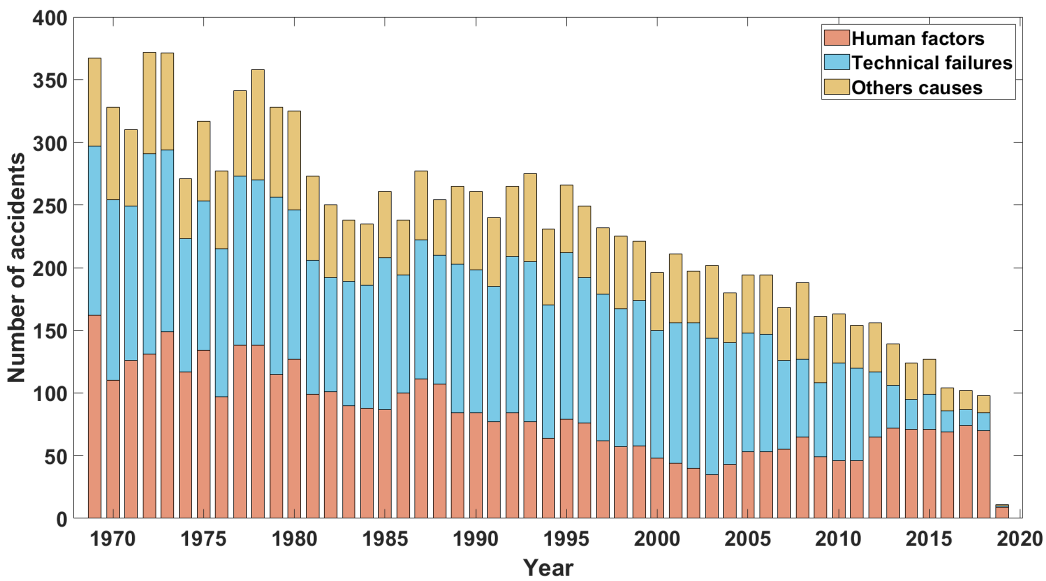


Figure 1.1: Commercial aviation accidents from 1970 to 2019 with human factors, technical failure, and other causes (Weather, . . .) as contributory factors. Data retrieved from Bureau of Aircraft Accidents Archives (www.baaa-acro.com).

One critical solution provided by the aviation industry in the 80s to deal with human errors has been the introduction of automation. And indeed, automation has made it possible to reduce human error, in particular by reducing the crew's workload through the integration of autopilots in the cockpit. However, automation has also shifted the role of the crew from controllers to supervisors (Parasurannan and Byrne, 2002). As a result of this role changing, the crews had to include in their expert panels another skill, that of monitoring systems (Parasuraman, 1987). During the flight, pilots must monitor rigorously their flight instruments to update their situation awareness. But, what does "monitoring" refer to ?

1.1 How and What to monitor ?

According to Flight Safety Foundation (Safety, 2018), "monitoring" is a word used quite liberally in aviation, and naturally, its meaning could be subject to confusion. In simple, plain language: Monitoring is adequately watching, observing, keeping track of, or cross-checking. In their report on "A practical guide for improving Flight Path Monitoring", the working group highlighted two components: "How to Monitor" and "What to monitor". The How to monitor part refers to actions required to perform the monitoring task. It can be divided into 3 subsections following the attention management (procedures/techniques for directing a pilot's attention to a particular place at a particular time), the deliberate checking (the active, disciplined and effortful action a pilot must take to look for something rather than just look at something), and the cross-checking (comparing separate, independent sources of information to confirm or refute understanding derived from the initial source). The "What to monitor" question is difficult to answer because all the skills above may be applied in different contexts. For example, the flight path could be monitoring as well as systems or the crew situational awareness. In order to investigate the question of "What to monitor?". We're going to rely on aeronautical accident examples to show how ineffective flight path monitoring leads to undetected errors. All the cases presented here are from a report of the Civil Aviation Authority on the development of pilot monitoring skills (Authority, 2013).

1.2 Aircraft Accident Reports

Accident: FedEx Flight 1478, July 26, 2002 (figure1.2)

On a flight from Memphis, Tennessee, U.S., a Boeing 727 descended through trees and impacted the ground 1,000 ft (305m) later. While sliding, the airplane struck construction vehicles that were parked on the field during the night, and burns marks on the ground indicated there was

Figure 1.2: FedEx Flight 1478 airplane wreckage

a fire on the airplane for the last 1,000 ft. The U.S. National Transportation Safety Board (NTSB) determined that the probable cause of the accident was the failure of the Captain (CA) and First Officer (FO) to establish and maintain a proper glide path during the night visual approach to landing.

Accident: King Air 100, October 25, 2002 (figure1.3)

Figure 1.3: King Air 100 airplane wreckage

On a flight from St. Paul to Eveleth, Minnesota, U.S., the flight crew failed to maintain an appropriate course and speed for the approach to Eveleth-Virginia Municipal Airport. During the later stages of the approach, the flight crew failed to monitor the airplane's airspeed and allowed it to decrease to a dangerously low level (as low as about 50 Kt below the operator's recommended approach speed) and to remain below the recommended approach speed for about 50 seconds. The airplane then entered a stall from which the flight crew did not recover. All occupants perished.

Accident: Empire Airlines Flight 8284, January 27, 2009 (figure1.4)

Figure 1.4: Empire Airlines 8284 airplane wreckage

While flying an ILS approach in light freezing drizzle, the crew experienced a flap asymmetry. The aircraft was at approximately 1,400 ft above ground level (AGL), just outside the final approach fix, when this occurred. The first officer continued flying the approach while the captain attempted to troubleshoot the flap problem by checking circuit breakers behind the first officer's seat and repositioning the flap handle several times. While the captain was attempting to troubleshoot the problem, airspeed decreased 35 Kt, culminating in a stall warning activation. The captain told the first officer, "Yeah, don't do that. Just keep flying the airplane, okay". The autopilot disconnected when the stall warning activated and the first officer advanced power. The stall warning ceased as airspeed increased. The first officer asked, "should I go around?" and the captain replied, "No, keep descending". The first officer was straining due to the amount of control wheel deflection caused by the flap asymmetry. The captain took over control of the aircraft as it passed through 700 ft AGL. He reduced power because the aircraft was now too fast. Airspeed decreased rapidly. As the aircraft passed through approximatively 150 ft AGL, the captain crashed about 300 ft (92 m) north of the runway threshold, slightly right off course, and skidded along the airport surface.

Accident: Asiana Airlines Flight 214, July 6, 2013 (figure1.5)

In July 2013, Asiana Airlines Flight 214 struck a seawall at San Francisco international Airport. The NTSB determined the airplane crashed when it descended below the visual guide path due to the flight crew's mismanagement of the approach and inadequate monitoring of airspeed. The NTSB also determined that the crew's insufficient monitoring of airspeed indications during the approach resulted from expectancy, increased workload, fatigue, and automation reliance.

Figure 1.5: Asiana Airlines 214 airplane wreckage

Accident: Colgan Air Flight 3407, February 12, 2009 (figure1.6)

Figure 1.6: Colgan Air Flight 3407 airplane wreckage

In February 2009, Colgan Air Flight 3407 crashed into a house in Clarence Center, New York, U.S., after experiencing an aerodynamic stall. The crew failed to detect a loss of 50 Kt of airspeed in 22 seconds. All of the above cases are examples of crews not properly monitoring the flight instruments and flight path deviation are relatively common, all operators are beset to some degree by the following types of deviations:

1. Altitude deviations;
2. Airspeed deviations;
3. Course deviations;
4. Taxi errors/ Runway incursions.

During monitoring, pilots are expected to carry out two distinct tasks. First, they monitor highly reliable automated systems over extended periods (such as in cruise flight). Second, they monitor complex aircraft flight path changes and system states while simultaneously completing several other flight-related tasks (e.g., communicating with air traffic control (ATC), cabin crew, passengers, programming approaches in the flight management system (FMS)). Even for highly skilled professional pilots, monitoring tasks are more challenging than they seem – especially when combined with other tasks, fatigue, or when flight parameters change very rapidly. This is why the dynamic phases are the most vulnerable to monitoring errors. In the report of the flight safety foundation involving 188 cases of airline monitoring errors the majority of monitoring errors – 66 percent – occurred while the aircraft was in a vertical phase of flight (e.g., climb, descent, approach, and landing). Across the 188 reports, a majority of the errors that resulted from inadequate monitoring were altitude deviations, course deviations, and speed deviations. Because modern aircraft have advanced autoflight capabilities and are highly reliable. Pilots often have little to do during the cruise for example, but they have to monitor occasionally unexpected flight path changes generated by the autoflight system and for system anomalies that rarely occur. Since pilots are humans, they are not immune to human failures. Indeed, the human brain has some limitations concerning monitoring and detecting events that rarely occur (Parasuraman, 1987).

1.3 Human cognitive limitation

There are sciences that study human functioning and cognition, attention span and limitations. The human brain processes a lot of complex information but is often unable to pick up information that arrives infrequently. Research in cognitive neuroscience and cognitive psychology has highlighted these human limitations Helmreich, 2000; Parasuraman, 2003; Parasuraman and Rizzo, 2008. There are some challenges and barriers to effective monitoring: many factors hamper monitoring, including system design, organizational factors, or external environment. But the biggest concern relates to human vulnerabilities, such as complacency/inattention, distraction, low attentional resource, low arousal, disorientation, tiredness, or stressors (i.e., workload, etc.). These concerns stem from historical accidents and incidents. Some relatively recent accidents and incidents, adapted from a report by the Flight Safety Foundation report, the Civil Aviation Authority report cited below, are illustrated here in terms of human limitations.

1.3.1 Inattention & distraction

The piloting activity is an attention-based activity with numerous tasks to be performed simultaneously. This activity is therefore extremely costly in terms of attention, cognitive, and motor resources. Attention is classically defined as a process integrated into the information processing system, the function of which is to enable the adaptation of human behaviour to its environment (Broadbent, 2013). It is therefore both what controls the processing of information, by selecting information relevant to the activity and inhibiting irrelevant information, and the resource that enables the information to be processed. Classically, we distinguish three main types of attention, each referring to an element of the above-mentioned definition: selective attention (referring to attentional control), divided attention (referring to the sharing of attentional resources on different tasks contributing to different goals), sustained attention (referring to the modulations of attention according to the duration of the activity) (Sarter, Givens, and Bruno, 2001). One example of an inattention problem with a concrete example during a routine flight from Faro to Bournemouth cleared for landing when the autothrust system disconnected on approach without the knowledge of the crew. On autothrottle disconnection, the visual warning, which was a flashing red annunciator, was activated for one-minute before being cancelled by manually disconnecting the autothrottle. The pilot monitoring (PM) was focused on the Primary Flight Display (PFD) and flight instruments and did not notice the flashing warning or the associated removal of the auto-throttle mode on the PFD. Because humans have limited attentional capacities, it is impossible to be aware of all elements of the environment. Attention plays the role of a filter directed toward the world; unexpected events (e.g., unsure, and infrequent) are difficult to be detected. Equipment failures are infrequent in modern commercial aircraft operations, and humans are inherently poor at monitoring for infrequent events (Dismukes and Berman, 2010). Distraction is one of the major factors which underlie most incidents and accidents, and this can be either physical or mental. Physical distractions occur due to unexpected problems in the cockpit or interruptions from cabin crew, ATC, etc. and mental distractions can refer to mind wandering (Baldwin et al., 2017) for example. Other fields also involve distraction issues such as automotive where Hendricks, Fell, and Freedman, 1999 shown that 37,6% of road accidents on 723 assessed resulted from distraction. In aviation, the report from the National Transport Safety Board (NTSB) concerning the famous crash of the Everglades in December 1972, cited "the failure of the flight crew to monitor the flight instruments during the final four minutes of the flight and to detect an unexpected descent soon enough to prevent impact with the ground. Preoccupation with a malfunction of the nose landing gear position indicating system distracted the crew's attention from the instruments and allowed the descent to go unnoticed".

1.3.2 High workload

Workload might simply be defined as the demand placed on the human operator (Moray, 2013). This definition, however, is overly limited because it only includes the requirements generated by external sources (e.g. task difficulty). In order to address workload completely, it is also necessary to consider demands generated internally that compete for an operator's resources. Therefore, an appropriate human factors definition of workload would be: workload is the demand placed on an operator's mental resources used for attention, perception, reasonable decision-making, and action. Because human resources are limited, the level needed for a specific task can exceed the available amount. Under these circumstances, workload can also be defined as the ratio of the resources required by the task to the amount of available resources. Inherent in this definition is the notion of resource differences among operators since each person will have a differing amount and nature of resources to apply to a task. Because of these differences, a given task will not produce the same workload level for all operators. Rather, workload depends on the operator's experience with the task, training, and relevant skill levels. A task can even produce different workload for the same operator at a different time depending on his or her state when the task must be performed. The automation is there to reduce and balance the workload during flights, but accidents have occurred where management of the Flight Management System (FMS) at critical phases of flight increases the workload resulting in errors going unnoticed. The Kahneman's resource model (Kahneman, 1973) captured elegantly the limitation imposed by brain capacities in this example. This model demonstrates that dual-task performance was less affected when two tasks required different processing structure, than when they used the same structure of processing. For this instance, two visual tasks are likely to interfere when being performed simultaneously because limited resources (Tversky and Kahneman, 1981) .

1.3.3 Low Arousal/Vigilance

Vigilance refers to the ability of observers to maintain their focus of attention and remain alert to stimuli over prolonged periods of time (Ballard, 1996; Lanzetta, Dember, Warm, and Berch, 1987; J. Warm, 1993). Arousal is that accounts for the vigilance decrement in terms of the lack of stimulation necessary to maintain alertness (Matthews and Davies, 1998). According to that models, the repetitive and monotonous aspects of vigilance tasks reduce the level of stimulation needed by elements of the central nervous system – the ascending reticular formation, the locus coeruleus, and the diffuse thalamic projection system – necessary to succor wakefulness and alertness (Joel S Warm, Matthews, and Finomore Jr, 2008). This aspect of human performance is an important concern due to the critical role that vigilance plays in aviation. The pilot's reaction time to any deviation will be slower and more error prone. This is particularly true during long haul flights were periods of routine flight management and weather surveillance are long. An example with a concrete case study, the aircraft was cruising at Flight Level (FL) 350 when it encountered turbulence and exceeded the target speed limit. The over speed alarm went off and the PM startled by the aural stimuli instinctively pulled back on the side stick for 6 seconds causing the AP to disconnect and to climb to FL 380. It took the crew 90s before they realized that the AP had disconnected (the over speed alarm had masked the AP disconnected alarm) and the FL had increased. The crew was distracted with the turbulence issue and no-one monitored the flight instruments. Important cues were missed (nose-up pitch of 12 degrees, high climb rate, excessive altitude, the position of the FD (Flight Directory) bars, FMA (Flight Mode Annunciator) indications, ECAM “AP OFF” message and AP light extinguished).

1.3.4 Disorientation

The human visual and vestibular systems are prone to illusionary inputs related to depth, height speed, and distance (Peters, 1964; Previc and Ercoline, 2004). This can seriously challenge the perception channels and result in incorrect decisions being made such as pilot induced oscillations (are sustained or uncontrollable oscillations resulting from efforts of the pilot to control the aircraft when the frequency of the pilot’s inputs and the aircraft’s own frequency are coupling). For example, the NTSB reported disorientation concerning the flight 4Q8 in January 2007. The aircraft was en-route from Surabaya, East Java to Manado, Sulawesi and was in cruise at FL350 with the autopilot engaged. The aircraft developed a problem with the inertial reference system and both pilots became so engrossed in sorting out the problem that they failed to respond to the increasing descent and bank angle. The pilots became disoriented and did not detect and appropriately arrest the descent soon enough to prevent the loss of control.

1.3.5 Lack of Attentional Resource

There have been many aircraft accidents and incidents where both crew members focused all their attention on dealing with a system malfunction to the detriment of other tasks (e.g., attention tunneling (Wickens and Alexander, 2009)). During a flight to Palmerston north, the crew encountered problems with the gear down mechanism. The captain instructed the FO to manage the malfunction while he kept an eye on the airplane. The FO started to go through the procedures but was interrupted by the captain who told him to skip through some of the tests. The captain was distracted from flying the aircraft and tries to help the FO sorting the problem out. The GPWS warning went off 4 seconds before the aircraft hit the ground.

1.3.6 Fatigue

Pilot fatigue is an insidious threat in aviation, but especially in operations involving sleep loss from circadian disruptions, increased sleep pressure from extended duty, and impaired alertness associated with night work (J. Caldwell and Caldwell, 2016). Furthermore, sleep inertia which occurs when one has just woken can hamper the monitoring task. Sleep debt and extended periods of wakefulness will impair the vigilance required under demanding flight conditions. A NTSB report (J. A. Caldwell, 2005; Marcus and Rosekind, 2017) has investigated an accident in which the captain failed to monitor the instruments and the FO failed to provide back-up and corrective input. They both failed to follow the stall procedures with the appropriate call outs and actions due to sleep deprivation as a major factor.

1.3.7 Situational Awareness

A recent study (Kharoufah, Murray, Baxter, and Wild, 2018) examined the different factors causations in a random sample over 200 commercial air transport accidents and incidents from 2000 to 2016. The objective of this study was to identify the principal human factor contribution to aviation accidents and incidents. The results showed that the most significant factor was situational awareness.

1.4 Theories of monitoring lapses

As we just saw with aircraft accidents, many factors are involved in monitoring lapses. The diagram in figure 1.7 summarizes the elements involved during monitoring lapses. Currently, there are two complementary theoretical currents that help explain attentional errors in aviation: mental load theory (Wickens, 1984) and attentional control theory (Posner, 1982; Spelke, Hirst, and Neisser, 1976). A first hypothesis considers attention as a resource reservoir with limited capacity. The attentional errors related to the performance of a secondary tasks, in piloting, would then be the result of mental overload. A second theory considers attention as a selective controller of information processing. Attention errors related to the performance of a secondary tasks, in piloting, would then be the result of a reorientation of the attentional control of the visual scene towards the items relevant to the performance of the secondary task (Simons, 2000; Strayer, Drews, and Johnston, 2003). Finally, a third complementary hypothesis, more specifically concerning perceptual control (Lemercier and Cellier, 2008; Noy, Lemoine, Klachan,

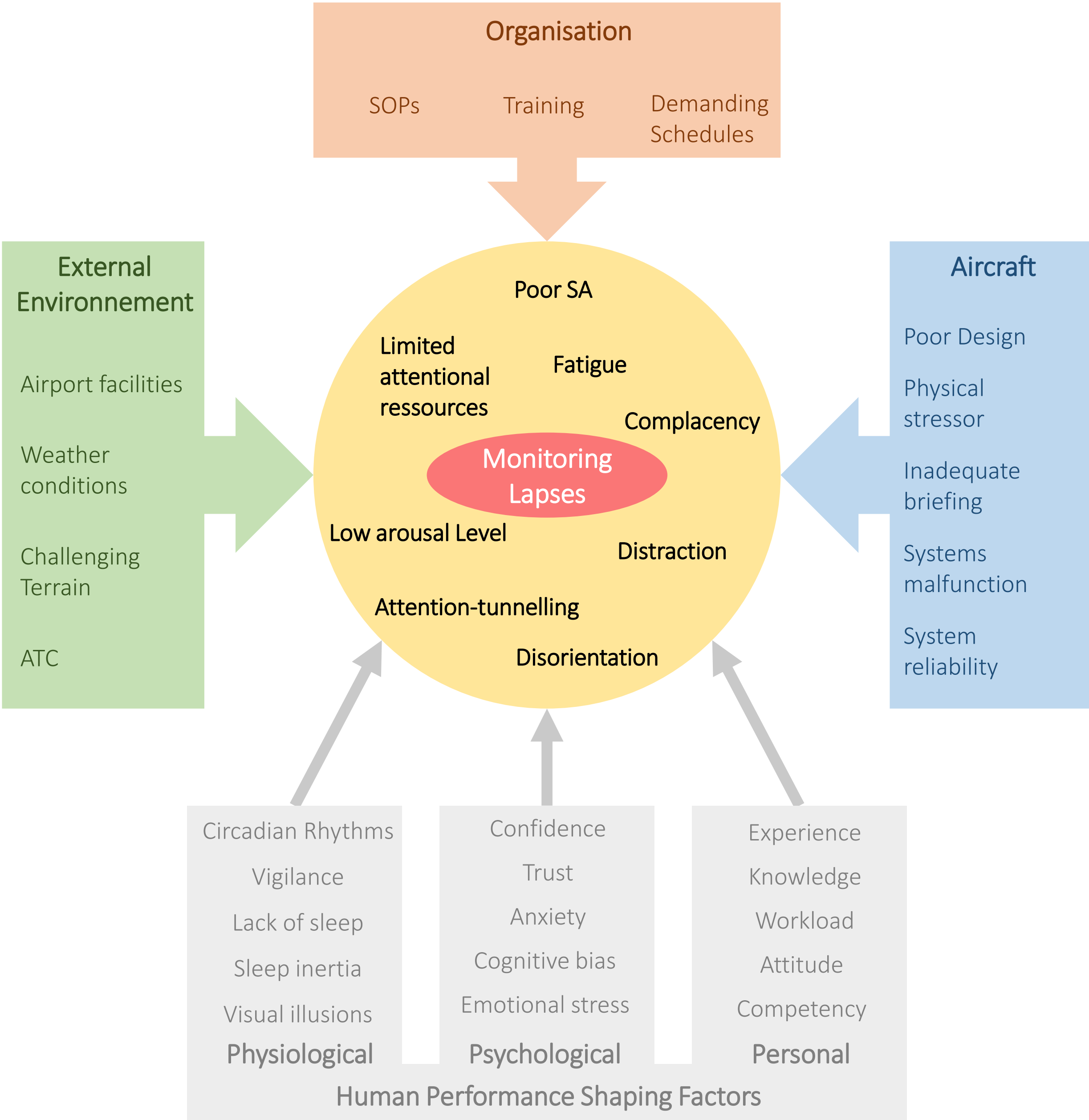


Figure 1.7: Monitoring lapses diagram (inspired from Authority, 2013)

and Burns, 2004), considers that the more a task requires visual control, the more it hinders visual control of the visual scene. These phenomena can promote the failure of the crew to monitor flight instruments properly.

1.5 Overcoming monitoring issues

As said previously, in numerous cases of aircraft accidents, pilots' visual scanning has been described as “inadequate”, “ineffective”, or “insufficient” (Jarvis, 2017). The National Transportation Safety Board reported in 1994 that inappropriate monitoring was involved in 84% of major accidents in the United States (Washington, 1978). The report was followed by numerous studies investigating the visual behavior of the pilots. However, in a “practical guide for improving flight path monitoring” by the Flight Safety Foundation (Group et al., 2014), which investigated 188 accidents with reported monitoring issues, it is underlined that many monitoring errors still occurred, most of them during dynamic phases of flight (e.g., climb, descent, approach, and landing). In 2013, the Federal Aviation Administration required airlines to include an explicit training program to improve monitoring skills Sumwalt, Thomas, and Dismukes, 2002. Following the PARG study, the Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation civile (French Investigation Agency) encouraged the use of eye tracker systems to finely analyze and improve crews' visual scanning. Interestingly, an extensive survey conducted on 931 pilots during the PARG study showed that most of the pilots need a better description of what a “standard” visual circuit in the cockpit can be. Similarly, in another recent survey (Lefrancois, Matton, Gourinat, Peysakhovich, and Causse, 2016), 75% of pilots deemed the knowledge of the required visual behavior for the different flight phases helpful to enhance their cockpit monitoring skills.

1.6 PM/PF roles

Initially, the supervision of the flight deck was performed by two pilots: the pilot flying (PF) and the Pilot Non-Flying (PNF). In response to a landmark accidents, FAA and Airlines redefined the role of the two pilots in the cockpit (AC 120-71A, 2003). From then, before the beginning of each flight sector, the aircraft captain decides which pilot will take direct responsibility for flying the aircraft for the complete flight or particular parts of it such as the Descent/Approach and Landing and they become 'Pilot Flying' (PF) for that sector or the specified part of it. The other pilot is then designated for that sector or relevant parts of it as 'Pilot Monitoring' (PM) and in that role must monitor the flight management and aircraft control actions of the PF and carry out support duties such as communications and check-list reading. The PM serves as the "second set of eyes" in the cockpit, occasionally assisting the pilot flying in his or her duties but primarily performing the job of keeping watch for anything amiss. The Operations Manual will specify fully the roles for the PF and PM, but one of the most important aspects of the duties of any PM is the cross-check of the actions of PF. Indeed, this part of the role represents one of the most important single reasons why a two-pilot flight crew is specified. As a result, checklist and monitoring are two essential defenses against pilot errors. However, it is not so easy to maintain excellent monitoring performance, Casner and Schooler, 2014 showed in their study of airline pilots that when pilots are free to monitor (i.e., without any particular defined task), monitoring lapses occur. Since monitoring lapses still occur, aviation industry thought may remove tasks done by human by dedicated systems more reliable to monitoring tasks. To figure this argument, have a look into the cockpits of the past.

1.7 Single Pilot Operations

The number of aircrew in cockpits was reduced over the last 60 years, going from 4 until the 1950s to 3 until the 1970s when the Navigator was removed - the radio navigator was dedicated to voice communication equipment - he was removed when inertial navigation systems were introduced. The crew was reduced from 3 to 2 until the 1980s when the Flight Engineer was removed (new monitoring equipment for engines and aircraft systems were introduced). Two pilot's aircrew have been the standard for three decades. This progressive elimination of technical crew members in commercial aircraft cockpits results from the replacement of human functions by system functions (Boy, 2014). Today, motivated by cost reduction, the shift from two pilot operations to single pilot operations (SPO) requires the aviation industry to investigate how cognitive processes will be redistributed among humans and systems. We put "humans" plural because even if the objective is to have a single pilot in the cockpit, there will be other human agents on the ground or onboard (e.g., flight planners, flight followers, and flight attendants) who could be involved. The single-pilot operation requires ensuring that the same level of safety is maintained despite the removal of a pilot on board. As history has shown us a few lines above, the SPO paves the way for the introduction of dedicated monitoring systems to counterbalance the crew reduction.

1.8 Neuroergonomics purpose

Human factors goal, particularly in aviation, is to reduce aircraft accidents (Salas, Maurino, and Curtis, 2010) by addressing the various causal factors. In this PhD work, we specifically address cockpit monitoring issues. The challenge is to improve pilot-aircraft interaction by considering the complex attentional and cognitive processes. Neuroergonomics approach proposes, using physiological tools, to find valid and robust measures of human behavior and cognitive processes (e.g., attention, memory, ...) (Parasuraman, 2003). Neuroergonomics generally promotes the use of brain imaging techniques or electroencephalography to measure the neural mechanisms underpinning human performance in complex real-life situations. Basic Neuroergonomics tools (e.g., EEG, ECG, fNIRS, ...) are often used in flight simulators to study cognitive processes (Callan and Dehais, 2019; Causse, Dehais, Péran, Sabatini, and Pastor, 2013). However, and despite significant progress, the use of such techniques does not come without constraints and requires the wearing of helmets or electrodes, and the signal is still strongly impacted by various source of noises (muscular artifacts, electromagnetic pollution etc.) (Aricò et al., 2018; Arico et al., 2017). Meanwhile, the possibility to perform remote measurement of the human attentional state in the cockpit is promising (Peysakhovich, 2016; Peysakhovich, Lefrançois, Dehais, and Causse, 2018). In this respect, an interesting technology is eye-tracking devices embedded in the cockpit. Since the last decade, the market of commercial eye trackers has been considerably democratized and such devices are non-invasive (can be used without any direct contact with the body), reliable, relatively inexpensive, and eye tracking data can be processed in near real-time. The data samples from an eye tracker device can serve to detect eye and gaze movements. Since the eyes are an important mediator between the environment and the brain, facilitating interaction with our everyday world. Eventually, the eye tracking techniques allow the monitoring of ocular behavior and give insight into the perceptual and cognitive processes underlying piloting (A. T. Duchowski, 2007; Glaholt, 2014).

CHAPTER 2

EYE TRACKING FOR RESEARCH AND PRACTICE

Sommaire

In chapter 1 we saw that a lot of accidents are due to human errors and a large part of them involves monitoring issues, particularly during landing. We also saw that those monitoring issues are the consequence of ineffective and inadequate visual circuits in the cockpit resulting in poor situational awareness. We assume that these issues can be helped by eye tracking integration in the cockpit. As a consequence, we propose in this chapter to present the eye tracking technology. After a brief presentation of how the eye operates, we introduce basic eye movements and present how eye tracking devices work with a theoretical background concerning eye tracking techniques throughout history. The human visual system is remarkable for the quantity and quality of information it provides us about our environment. A quick glance is enough for us to get information about the size, shape, color, and texture of the objects. The visual system is also able to identify whether objects are in motion, or to know their directions or relative velocities. This section describes the mechanisms behind the vision. In the same way, as in a bottom-up approach, a stimulus, goes through the cornea, the lens, projects onto the retina, and is finally processed by the brain in an afferent way. This section follows roughly this chronology. First, the light entering the eye, and hence we firstly describe the human neural hardware present in the eye. Then, following the optic nerve, this section ends with a simplified view of the brain to identify the neural mechanisms involved in visual attention and the orientation of the visual attention trough the eye movements.

2.1 Human neural hardware

The human eye lets light in through the pupil, turns the image upside down in the lens end then projects it onto the back of the eyeball-the retina see figure 2.1.

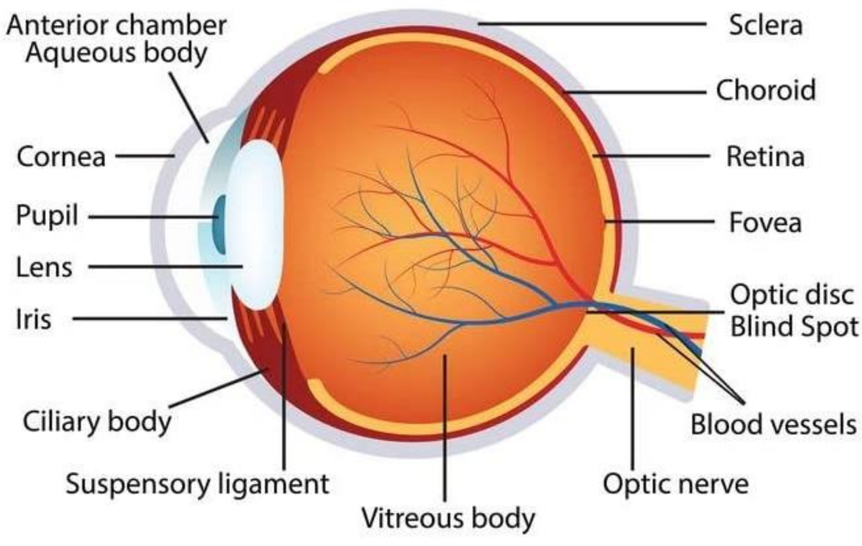


Figure 2.1: Human eye anatomy

At the front of the eye is the cornea, allowing light to enter the eye. After the cornea, the light rays pass through two liquid materials (e.g., aqueous body, and vitreous body). Between the cornea and the lens lies the anterior chamber, which is filled with a liquid (aqueous body) that provides nutrients for both structures. The other liquid material, the vitreous humor accounts for 80% of the eye's volume and helps to maintain the eye's shape. The formation of sharp images on the retina is largely due to the refraction of light by the cornea and the lens. The refractive power of the lens is smaller than the power of the cornea, although it is adjustable. These dynamic changes in the refractive power of the lens make it possible to focus objects at varying distances on the retina. Only the retina is filled with light-sensitive cells (e.g., as photoreceptors), called cones and rods (see figure 2.2), which transduce the incoming light into electrical signals sent through the optic nerve to the visual cortex for processing. The layer adjacent to the retina contains choroid, which is composed of blood capillaries, the main source of blood supply to the photoreceptors. Cones are sensitive to visual details (as high spatial density) and provide our color vision. Rods are sensitive to light and support vision under dim light conditions. There are three types of cones (e.g., red, green, blue). Each eye has about 7 million cones and 120 million rods. There is a small region (figure 2.2) called the fovea. Spanning less than 2 deg of the visual field, cones are extremely over-represented, see figure 2.2, while they are sparsely distributed in other areas of the retina.

As a consequence, the full acuity of the eye is only in this small area. Foveal information is prioritized in processing due to the cortical magnification factor, which increases linearly with eccentricity, from about 0.15 deg/mm cortical matter at the fovea to 1.5 deg/mm at an eccentricity of 20 deg (Hubel and Wiesel, 1974). As a result, R. L. De Valois and De Valois, 1980 have shown that about 25% of the visual cortex processes the central 2.5 deg of the visual scene.

The retina is composed of multiple layers of different cell types (R. L. De Valois and De Valois, 1993). It resembles a three-layer cell sandwich, with a connection bundle between each layer. These connectional layers are called plexiform or synaptic layers. The retinogeniculate organization is schematically illustrated in figure 2.3. The outer layer contains the photoreceptors (e.g., rods, and cones). The first connectional layer is the plexiform layer which houses the connection between receptors and bipolar nuclei. The next layer of cells is the inner nuclear layer containing bipolar (e.g., amacrine, bipolar, horizontal) cells. The next layer is the inner plexiform layer where the connection between inner nuclei cells, and ganglions cells are formed. The last layer is composed of ganglion cells. Rods and Cones are specialized types of neurons involving specific types of dendrites. In general, each neuron can connect to as many as 10,000 other neurons. This interconnected block behaves as a large neural circuit. Ganglions cells are of the threshold type, sending a signal when the activation exceeds a specific level.

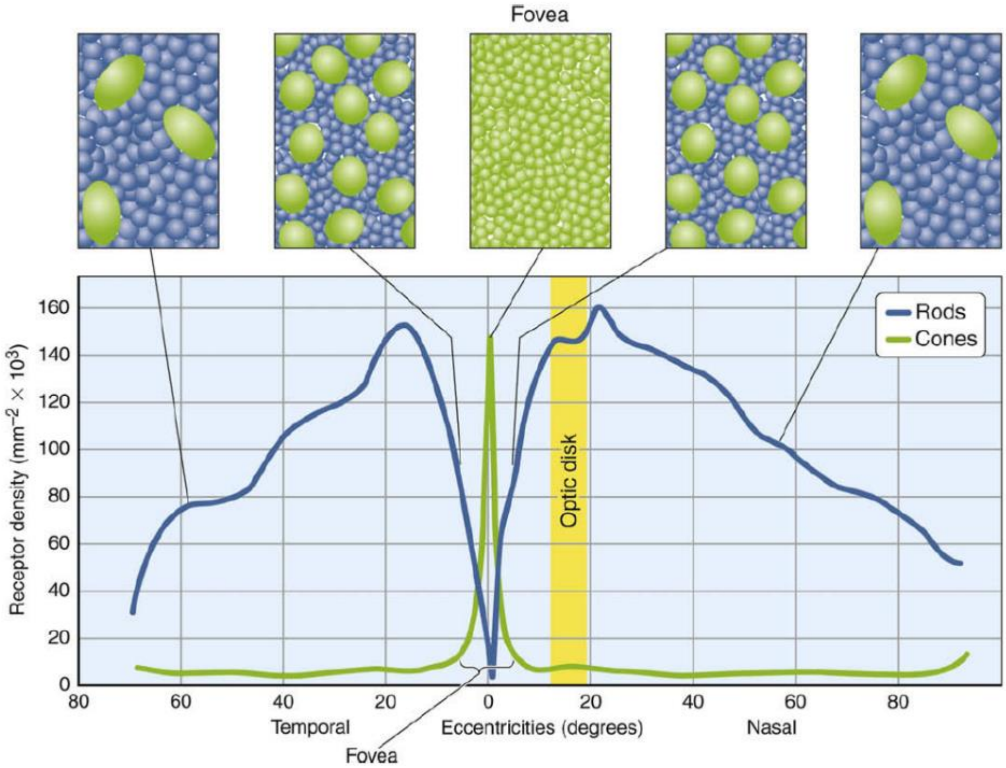


Figure 2.2: Distribution of the photoreceptors in the eye. In the fovea, the cone density is the highest and is correlated with visual acuity. Reported from (Mustafi, Engel, and Palczewski, 2009).

2.2 Eye movements

Eye movements are particularly important since visual acuity is only present in the fovea. Eye movements are used to orient the fovea to new areas of interest in the visual field (also known as "foveation"). These eye movements are controlled by three pairs of muscles illustrated in figure 2.4.

They control the three-dimensional orientations of the eye inside the head for horizontal (yaw), vertical (pitch), and torsional (roll) eye movements. According to Donder's law (Tweed and Vilis, 1990), only the activation of these muscles is responsible for the direction of the gaze, regardless of how the eye was previously. The eye tracking literature has long sought to measure and analyze the different eye movements (Holmqvist et al., 2011). Eye movements' classification is presented in Table 2.1 for each eye movement type, their durations, amplitudes, and velocities. Eye movements can be grouped into 2 functional categories: those used to stabilize the gaze

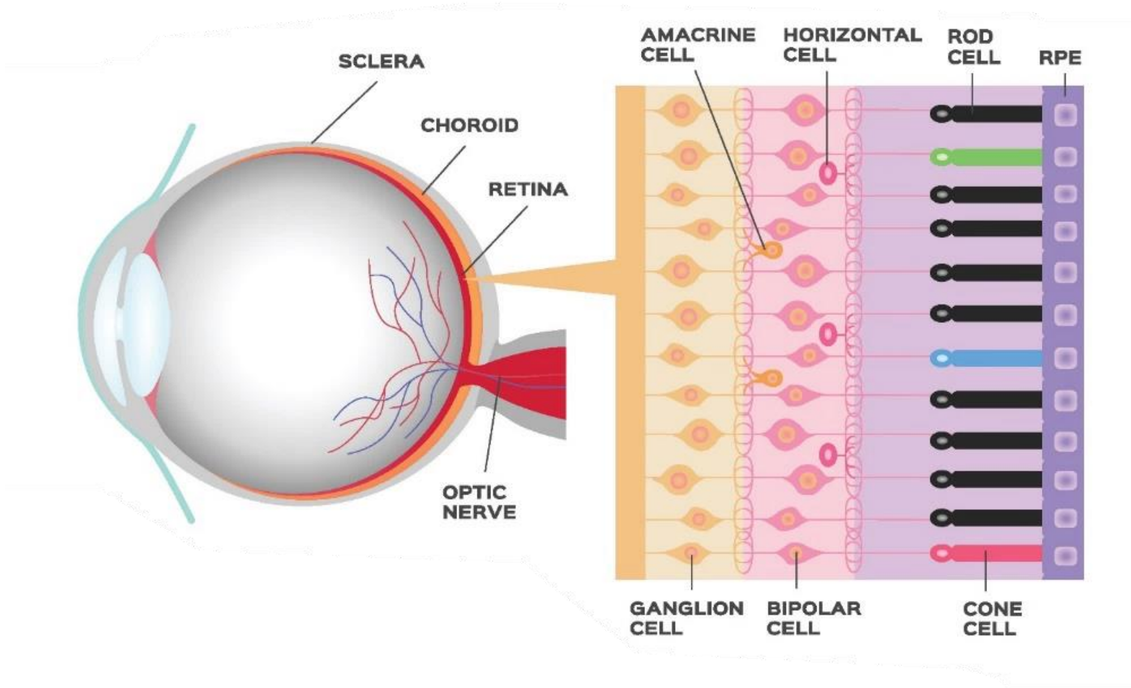


Figure 2.3: Schematic of the retina with their multiple layers (RPE: Retinal Pigment Epithelium)

during head movement or when the environment in the visual field is moving (e.g., vergence, optokinetic response, vestibulo-ocular reflex); and those used to direct the gaze when the fixated targets move (smooth pursuit) or when a new target is fixated (e.g., fixation, saccades, glissade). Other micro eye movements exist, because the eye is not completely stationary during a fixation, but has three types of micro-movements such as tremor: a small movement of frequency around 90Hz whose exact role is unclear (Martinez-Conde, Macknik, and Hubel, 2004), drifts: slow movements taking the eye away from the center of fixation, microsaccades: Their roles are to quickly bring the eye back to its original position. These intra-fixational eye movements (Holmqvist et al., 2011) are not studied in this manuscript. Other ocular motion is not subject to voluntary control. For example, pupil diameter is modulated by the antagonism of the parasympathetic and sympathetic nervous systems. Other reflexive movements include the optokinetic response (the smooth pursuit of an object as it travels through the environment followed by an immediate return of the eye to its original position; (Distler and Hoffmann, 2011)) and vestibulo-ocular reflex (the movement of the eye to maintain a stable retinal image due to vestibular activation; (Laurens, Strauman, and Hess, 2011)). When observing and following an aircraft in the sky, for example, our eyes perform one type of movement called smooth pursuit. Smooth pursuit is a slow movement to follow a moving stimulus by maintaining the image

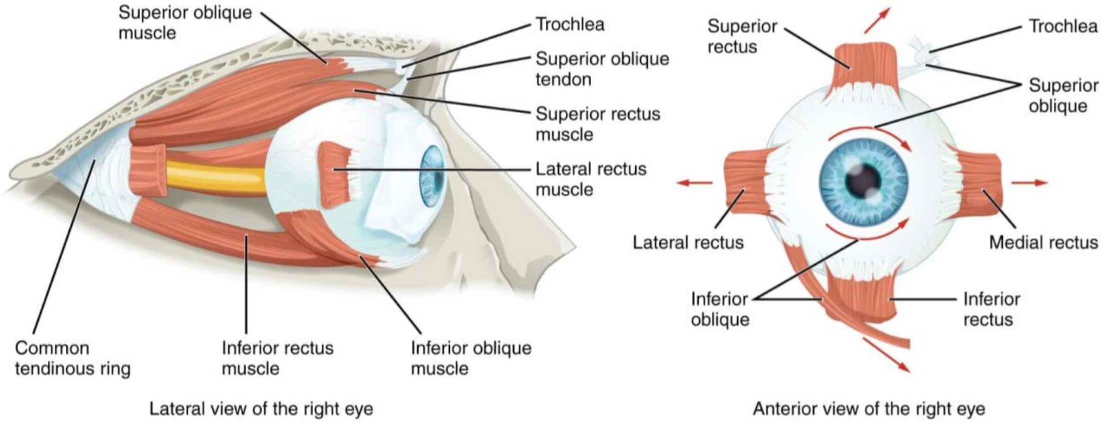


Figure 2.4: Extraocular muscles controlling eye orientation (Adduction/Abduction thanks to lateral and medial rectus , Depression/Elevation with superior and inferior rectus, and Intorsion/Extorsion with superior and inferior oblique).

on the retina. The rapid motion from one fixation to another is called “saccade”. Saccades are rapid, ballistic eye movements that cause an abrupt change in the direction of the gaze. Because saccades do not stop directly at the intended target, glissades are the post-saccadic movements that “wobble” towards the intended target. Indeed, the most reported event in eye tracking data is not related to a movement but rather to a state when the eyes remain "still" over a period of time, known as a fixation. It corresponds to the eye movements that stabilize the retina over a stationary object of interest.

As Yarbus (Yarbus, 1967) showed decades ago, eye movements are indicative of the strategies used to inspect a visual scene. Figure 2.5 shows the scan path of a participant examining the bust of Nefertiti for 2 minutes. The straight thin lines represent the saccades of the eyes, which are rapid ballistic movements that allow the fovea to be positioned in alignment with the elements of the picture. Along these lines, the densest areas correspond to the fixation points on which the participant has stopped to take the information. These results obtained by Yarbus show that seeing is an active process in which eye movements orient the gaze several times per second to direct the fovea towards particular regions of an image, in order to examine certain features that are particularly interesting or informative.

Eye movements types	Duration (ms)	Amplitude	Velocity
Fixation	>200	-	-
Saccade	30-80	4°-20°	30°/s-500°/s
Glissade	10-40	0.5°-2°	20°/s-140°/s
Smooth pursuit	-	-	10°/s-30°/s
Microsaccades	10-30	10'-40'	15°/s-50°/s
Tremor	-	<1'	20'/s (peak)
Drift	200-1000	1'-60'	6'/s-25'/s

Table 2.1: The most common types of eye movements events with their typical value (recovered from (Holmqvist et al., 2011)).

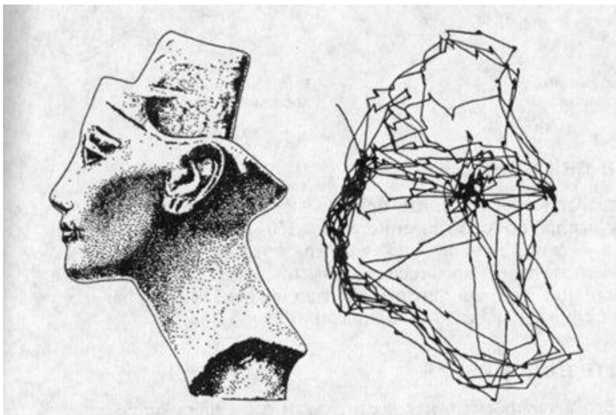


Figure 2.5: Eye movements of a participant looking at the bust of Nefertiti. The line on the right represents the eye movements performed during the 2 mins visual exploration (according to (Yarbus, 1967)).

2.3 Neural mechanism of the Visual attention

The human visual system is usually described by visual pathways corresponding to the connection between retina and brain regions. In a simple view of the brain (see figure 2.6), it is possible to identify the neural mechanisms involved in visual attention. There are three main neural regions implicated in eye movements programming and their functions (see Palmer and Palmer, 1999; for a review see A. T. Duchowski and Duchowski, 2017).

- Posterior Parietal Complex: disengages attention,
- Superior Colliculus: relocate attention,
- Pulvinar: engages, or enhances, attention.

These neural mechanisms are at the basis of the processes underlying visual attention and are responsible for the generation of eye movements.

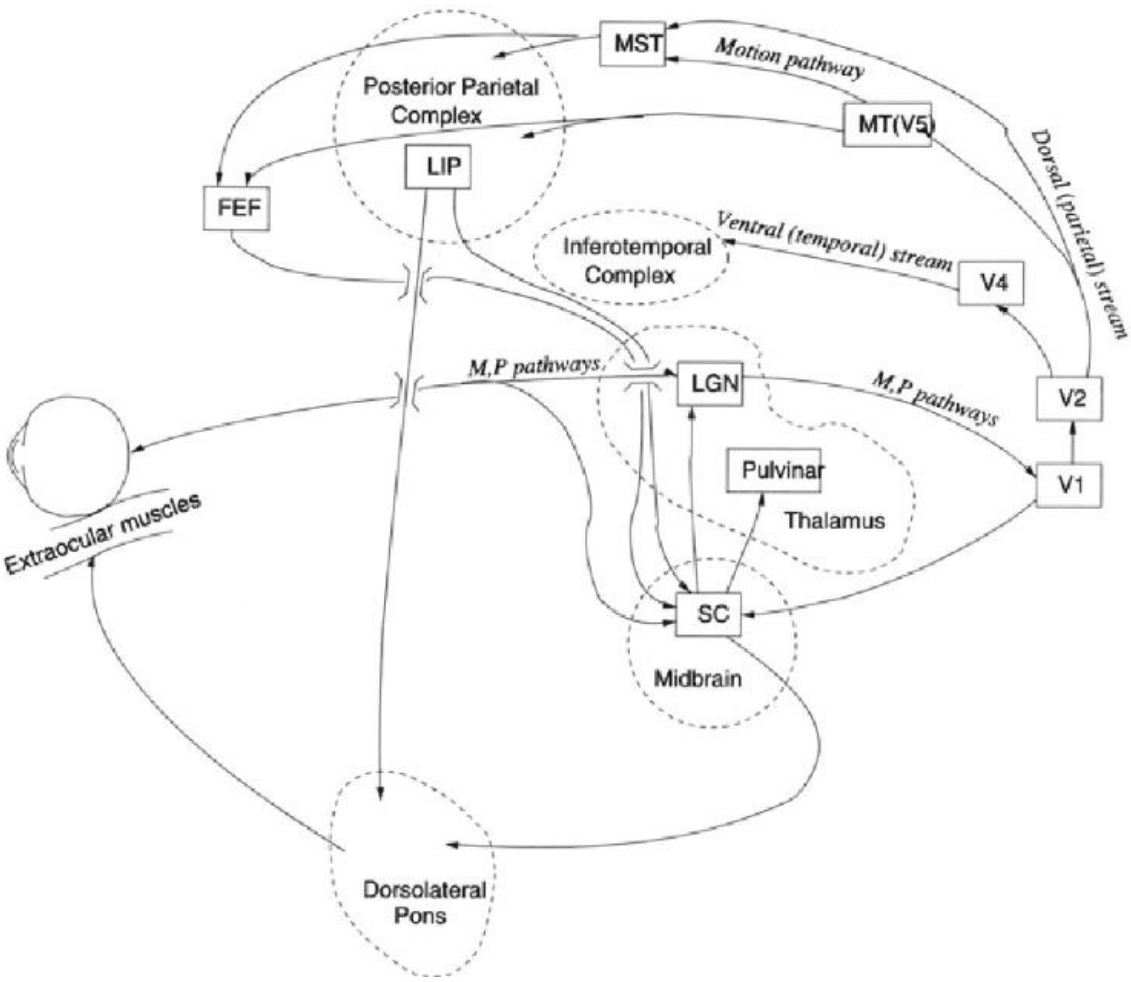


Figure 2.6: Simplified view of the brain and the visual pathways involved in selective visual attention (From (A. T. Duchowski and Duchowski, 2017)).

In this chapter, we have seen the anatomy of the eye and how it works. We have also highlighted the different eye movements and their implications for visual attention through the brain regions involved in the programming of eye movements. The next section discusses the subject of gaze detection and how eye movements can be deduced from it with eye tracking.

2.4 Eye tracking devices and technology

Eye tracking devices exists since a long time (see figure 2.7 for an overview, and see (Wade, 2010) for a complete review).

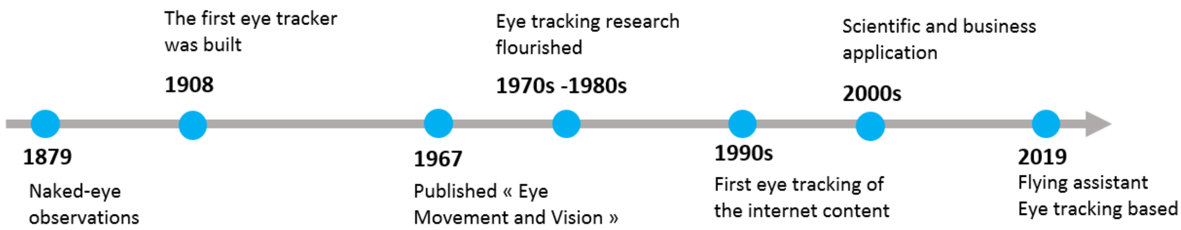


Figure 2.7: Quick overview of eye tracking technology and use case throughout the time

Origins of eye tracking date back to 1879 when the French physiologist Louis Emile Javal (figure 2.8) noticed that readers' eyes do not skim fluently through the text while reading but make quick movements (saccades) mixed with short pauses (fixations). In 1908, Edmund Huey (see figure 2.8) built a device which could track eye movement during the reading process.

This first eye tracker technique was intrusive as readers had to wear a contact lens with a small opening for the pupil. The lens was attached to a pointer which changed its position following the movements of the eye. The first recorded eye movements were not released until 1937, when Guy Thomas Buswell, an educational psychologist, used light beams (see figure 2.9) which were reflected in readers' eyes and recorded them on film. Buswell's research indicated that there is a significant difference between oral and silent reading, and that one person can read in two different ways.

Then, Alfred Lukyanovich Yarbus, a Russian psychologist, conducted several eye tracking studies in the 1950s and 1960s. The results showed that the readers' eye movement and fixation depend on their interest and the given task. For example, if the reader was asked several questions about the shown images, their eyes would focus on those parts which are relevant to the questions (see figure 2.10). In 1967, he published a highly influential book called "Eye Movements and Vision" (Yarbus, 1967). The research of eye movement and eye tracking

Figure 2.8: Louis Emile Javal (Left) and Edmund Huey (Right)

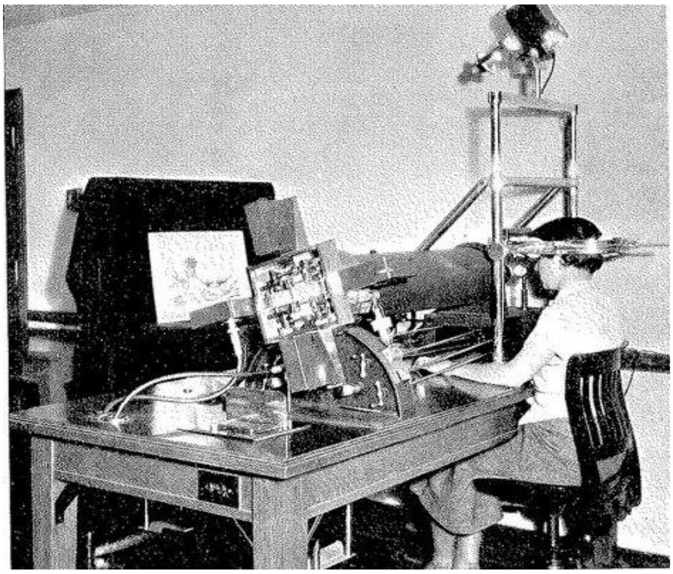


Figure 2.9: Apparatus used by Guy Thomas Buswell

thrived during the 1970s and 1980s. In the 1970s, the eye trackers became less intrusive, provided better accuracy, and were able to separate eye from head movements. At the same time, psychological theories started to examine the connection between eye tracking data and cognitive processes (Mele and Federici, 2012). The tracking and measures of eye behaviour were complex and expensive and were thus restricted to the military field and research laboratories.

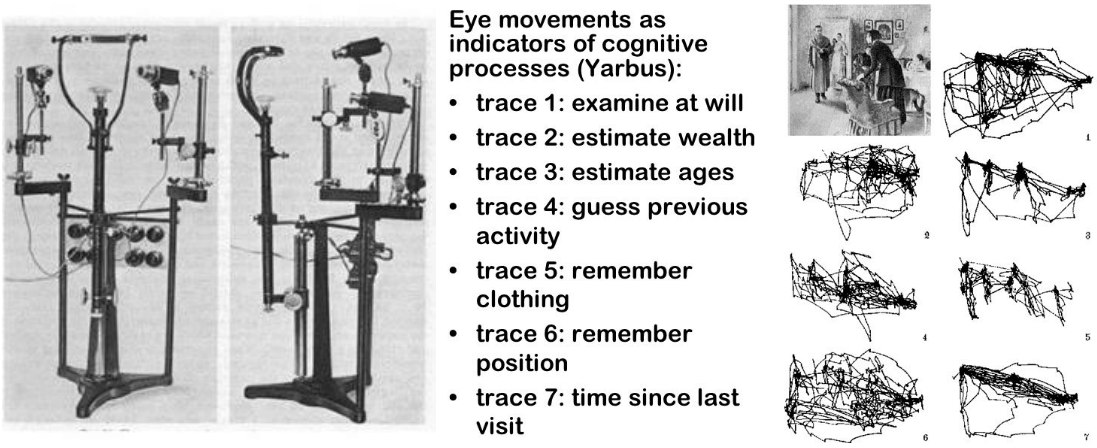


Figure 2.10: Yarbus eye tracker devices (left) and Yarbus eye tracker data reported from the book Eye Movements and Vision

Figure 2.11: Left: Example of video-based eye tracking technique (Pupil Invisible), the video with gaze tracking is stream on mobile. Right : Example of EOG tracking technique.

However, technological prowess, from the 1970s until nowadays, such as increased processor speed, and advanced digital video processing, have contributed to reduce the cost and increase the efficiency of eye tracking devices. The most widely used current designs are video-based eye trackers (see figure 2.11). Even if these techniques are predominant, we have to mention the electro-oculography (EOG) tracking technique to be complete. It is based on the fact that an electrostatic field can be modulates when eyes rotate (e.g., It can be described as a fixed dipole with positive pole at the cornea and negative pole at the retina.). By recording small differences in the skin potential around the eye, the position of the eye can be estimated. Also, since this is done with electrodes placed on the skin around the eye, this technique does not require a clear view of the eye. This technique is rather uncomfortable, and is not well-suited for everyday use, since it requires the close contact of electrodes to the user. The first method outlined below,

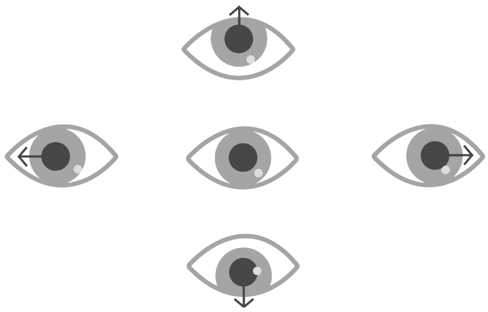


Figure 2.12: Example of corneal reflection at various gaze positions

video-based eye tracking technique, uses the pupil and the cornea properties. Because the pupil absorbs a large amount of light. This chromatic property allows it to be easily identifiable in image recognition (for a review see Hansen and Ji, 2009, and Hammoud, 2008). Concerning the cornea, it covers the outside part of the eye, and as presented in the section above, it is filled with a liquid that gives an aqueous appearance and transparency. Because of that, it reflects light. Usually, the reflection that you can see in someone's eyes comes from the cornea. This corneal reflection is known as 1st Purkinje reflection (P1) (Crane, 1994) illustrated in figure 2.12. The Purkinje reflection is shown in figure 2.12 as a small white circle in proximity to the pupil. Other light reflections coming from outside might create other interfering reflection. Basically, infrared light sources and infrared cameras are used to avoid this phenomenon by illuminating and recording eye movements in infrared spectrum, thus filtering natural light.

CHAPTER 3

STATE OF THE ART OF EYE GAZE TRACKING IN THE COCKPIT

Sommaire

We have seen in the previous chapter how the eye roughly works and the extent to which it is possible to use some of these characteristics to identify its position and interpret the direction of the gaze. In line with the previous chapter, an historical background of eye tracking in the cockpit will be detailed. Eventually, we finish this third chapter by stated on different eye tracking metrics useful for the aviation context, paving the way to the flying assistant of the next chapter.

3.1 Historical background of eye tracking in the cockpit

Scientists scrutinize the pilots' eyes since more than a century (for a review, Di Stasi and Diaz-Piedra, 2019), starting with balloon flight studies from the French physiologist Paul Bert, that examined among other things the effects of altitude on vision at the end of the XIX century. Later, in the context of World War I, Tiffin and Bromer (Tiffin and Bromer, 1943) conducted exploratory works on the visual scanning strategies of experienced and inexperienced pilots during the landing phase. They developed a photographic instrument allowing creating motion pictures based on the chronophotography technique (Marey, 1894). This method enabled the researcher to analyze the pilots' visual scanning during the last seconds of a flight, just before the touchdown. Although the project was not completed, and no final report is available, the investigation has served as an incentive for further studies by other groups. Thereafter, R. E. Jones, Milton, and Fitts, 1949 and colleagues examined the pilots' visual scanning to optimize the position of the flight instruments in the cockpit. They found that some instruments have a stronger relationship than others as indicated by the frequency of transitions between these particular instruments. They proposed to use such results and methods to improve flight instrument arrangements. Afterward, many studies have been interested in the analysis of visual behavior in the cockpit through the exploration of visual scanning.

3.2 Visual scanning in the cockpit

In general, pilots' visual behaviors are transcribed in a series of fixations and saccades. This makes it possible to divide their visual attention between the outside scene and the on-board

Figure 3.1: Eye tracking recording showing the visual path of a pilot in approach phase: 30 seconds recording. The larger the diameter of the red dot, the longer the fixation time.

instruments. Figure 3.1 is a representation of the visual path performed by a pilot in the cockpit during a recording of 30 seconds. As illustrated, fixations and saccades sequences allow to discern the path used by the pilot. However, from a technical point of view this data is only X,Y,Z coordinates and does not allow to add a semantic dimension, i.e. the pilot observed the PFD then the ND, and finally the ECAM. To overcome this problem, a widely used technique is to divide the cockpit panel into areas of interest corresponding to the cockpit instruments, as presented in figure 3.2. An area of interest (AOI) is a part of a stimulus that is of special importance (e.g., in our case flight instruments). They are created based on semantic information of the stimulus. Basically, the most important information given by AOI is related to transitions and dwells. A transition is a saccadic movement from one AOI to another, and a dwell is a temporal aggregation of fixations within an AOI (Blascheck et al., 2017). AOI based techniques provide additional information (e.g. semantics of the observed zone) in complement to the temporal aspect (e.g., order of transitions and time spent on it). Figure 3.3 gives an example of a recording of an airline pilot during 100 seconds of a landing. The AOIs presented in figure 3.2 are the same as the ones used in the figure 3.33.4. This recording allows distinguishing the moment in which the gaze is directed toward flight instruments, enabling to calculate the portion of time spent on a specific flight instrument (also known as Percentage Dwell Time (PDT)). Since information is not processed instantly upon the occurrence of a stimulus, a processing time is required for the brain.

Goldberg, Stimson, Lewenstein, Scott, and Wichansky, 2002 has defined this processing time of about 200–400 ms concerning ocular dwells. By using this temporal filter, it is possible

Figure 3.2: Cockpit view with flight instrument contouring by Area of Interests.

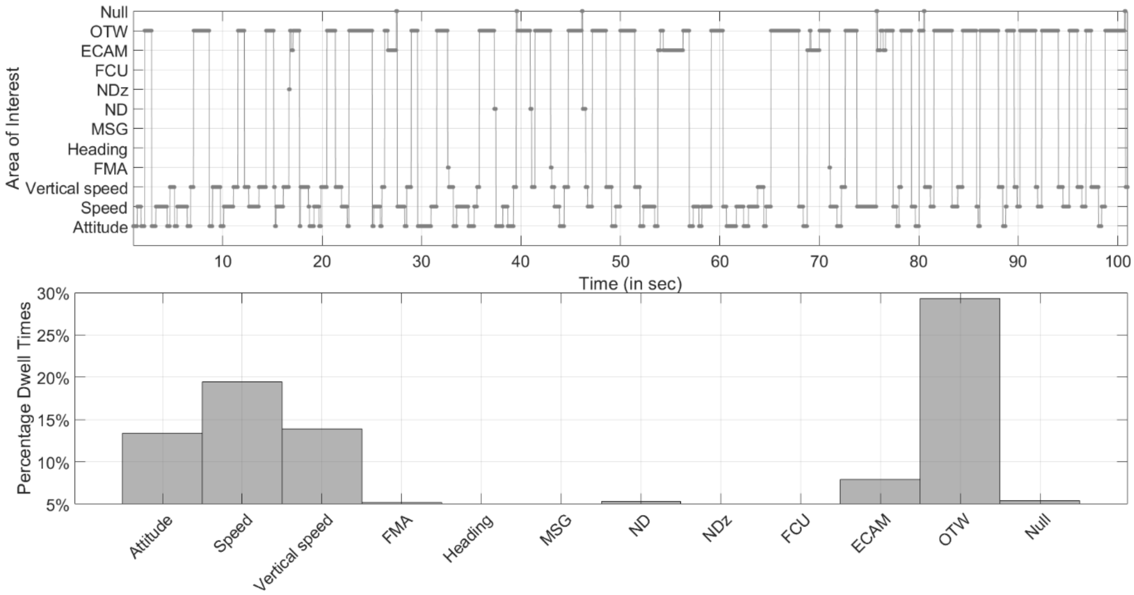


Figure 3.3: Top: AOI sequence chart; Bottom: Percentage Dwell Time (PDT) on various AOIs of a pilot during a landing approach for 100 sec.

to discern the attentional paths taken by the pilot. For example, we can see in figure 3.3 that most of the time is spent on the Out of the Window zone, the attitude zone, the speed

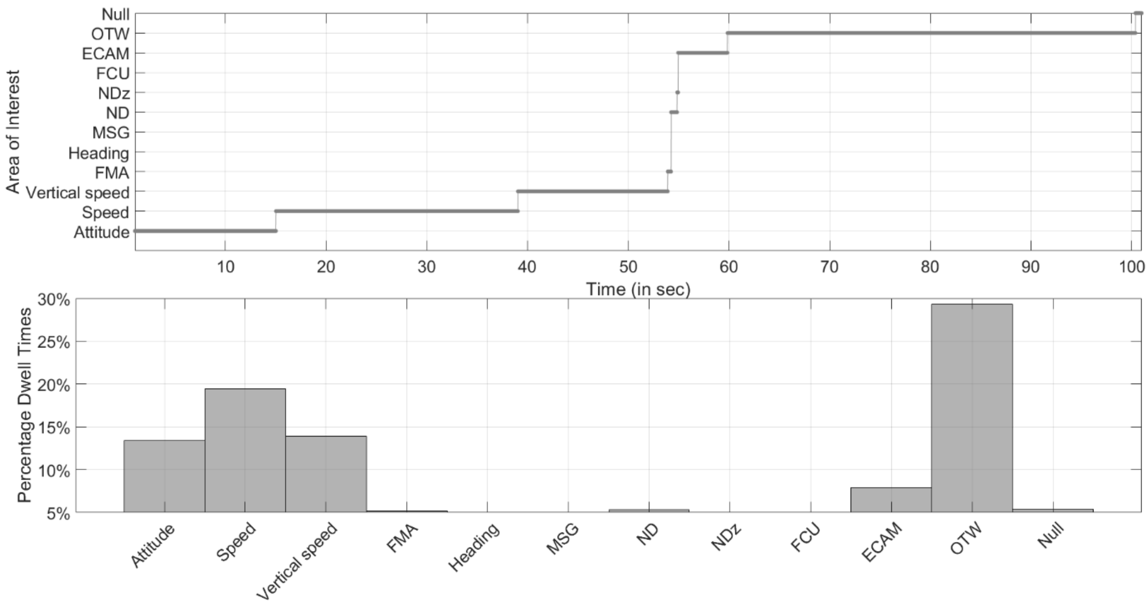


Figure 3.4: Top: AOI sequence chart; Bottom: Percentage Dwell Time (PDT) on various AOIs of a pilot during a landing approach for 100 sec.

tape, and the vertical speed tape. It is largely admitted that the time spent on AOIs are the reflection of the visual strategies employed by the pilots (Kasarskis, Stehwien, Hickox, Aretz, and Wickens, 2001; Ottati, Hickox, and Richter, 1999; C.-s. Yu, Wang, Li, and Braithwaite, 2014; C.-S. Yu, Wang, Li, Braithwaite, and Greaves, 2016). However, this is not enough to determine the visual scanning strategies underpinning the visual attention allocation. Figure 3.4 is a use case where the Percentage Dwell Times are the same as in figure 3.3. However, the transition frequency between AOIs is not equivalent, and this underlines a visual activity with longer fixation time on each instrument.

3.3 Why it is useful to investigate visual scanning strategies

The relationship between visual scanning skills and performance has been highlighted in experiences where participants were trained to gaze at relevant areas. Shapiro and Raymond, 1989 were among the firsts to link performance and efficient visual scanning techniques. In their study using videogames, three groups of gamers received a different training program: efficient scanning techniques, inefficient scanning techniques, and a control group without any training. The efficient" group was trained to minimize their eye movements and optimize their visual scan strategies. In contrast, the “inefficient" group was trained to increase the frequency of their eye movements without any established pattern. The participants exposed to efficient scanning training showed better performance and fewer fixations compared to two other groups. More recently in the air traffic control field, Kang and Landry, 2014 used the same method to enhance novices' performance in a conflict detection task. A novice group was exposed to experts' visual scans (“treatment"), overlaid on the radar screen during the task. Novices from this group outperformed novices from two other groups: the “control" group, which received no particular instructions, and the “instruction-only" group, which was verbally instructed to attend to altitude. These two last studies convincingly support the existing relationship between visual patterns, task performance, and the possibility to improve these patterns with adequate training. Furthermore, the links between the visual scanning strategies and the expertise were observed in fields such as radiology, driving, sport, or chess (e.g., Blignaut, Beelders, and So, 2008; Ganesan, Alakhras, Brennan, and Mello-Thoms, 2018; Memmert, Simons, and Grimme, 2009; Simon and Chase, 1988; Underwood, 2007). Gegenfurtner, Lehtinen, and Säljö, 2011 conducted a meta-analysis and highlighted that experts (compared to non-experts) generally demonstrate more fixations on task-relevant areas as well as shorter fixations. In their review of eye movements in medicine and chess, Reingold and Sheridan, 2011 have labeled this greater perceptual effectiveness of experts as "superior perceptual encoding of domain-related patterns". In addition to being indicative of performance and expertise, visual strategies can be used as a cognitive state marker (for a review see Peißl, Wickens, and Baruah, 2018; Ziv, 2016).

3.4 Visual scanning as a marker of the pilot's (Cognitive) Psychophysical State

3.4.1 Fatigue, Sleepiness, and stress

Diaz-Piedra et al., 2016 and Di Stasi, Catena, Canas, Macknik, and Martinez-Conde, 2013 found that saccadic velocity is a bio-marker related to aviator fatigue. Based on visual scanning, Wu, Wanyan, and Zhuang, 2015 provided a mathematical model connecting pilot's visual attention allocation and flight fatigue. Allsop and Gray, 2014 showed that under anxiety, the percentage of dwell time directed toward the Out of the window zone is increased.

3.4.2 Situation Awareness

van de Merwe, van Dijk, and Zon, 2012 measured SA by studying the pilots' search patterns (fixation rates on the display, dwell time on the display, and scanning entropy) in relation to information acquisition. The authors stated that this was done to assess level 1 as well as level 3 of SA via eye tracking (level 1 corresponding to the perception of an element in current situation and level 3 corresponded to the projection of future status (see D. G. Jones and Endsley, 1996 for a review of SA errors in aviation)). Specifically, pilots had to deal with a fuel leak that was expected to hamper SA. Pilots with high SA as assessed by their scanning measure (e.g., high fixation rate on the electronic centralized aircraft monitoring display) found the source of the malfunction earlier, showing more structured and predictable cockpit scanning. C.-s. Yu et al., 2014 suggested integrating eye-tracking devices into simulators for promoting SA training. They found pilots with better SA performance in the simulator showing lower perceived workload. Ryffel, Muehlethaler, Huber, and Elfering, 2019 developed a debriefing tool eye tracking based in upset prevention and recovery training (UPRT), this tool intended to improve the ability of pilots to recognize and avoid situations that can lead accidents.

3.4.3 Workload

Visual scanning can serve as a measure of the mental workload of pilots (Di Nocera, Camilli, and Terenzi, 2007; Li, Chiu, and Wu, 2012. Di Nocera et al., 2007 have shown that the dispersion of the eye fixations was higher during the takeoff and landing phases than during the cruise phase. Literature has indeed highlighted different mental workload levels attributed to these different flight phases. Furthermore, other research brings to light the effect of mental workload on flight performance (Dahlstrom and Nahlinder, 2009; Hankins and Wilson, 1998; C.-S. Yu et al., 2016). Increased workload conducted to a lower percentage of fixations Out of the Window and a higher percentage of fixations to the tactical display during low level-high speed combat aircraft simulations (Svensson, Angelborg-Thanderez, Sjöberg, and Olsson, 1997). Russi-Vigoya and Patterson, 2015 modulated the workload by adding an engine failure during experimental simulations. This experiment showed that the visual scanning changed (directed toward certain flight instrument) during the introduction of this unexpected condition that generated a high workload. In addition, Tole, Stephens, Harris, and Ephrath, 1982 found in an experiment involving aircraft pilot that the visual scanning on flight instruments varies as a function of the load and increased as a function of the estimated skill level of a pilot.

3.4.4 Expertise

According to Bellenkes, Wickens, and Kramer, 1997 the fixations of experts are shorter and fixations on instruments are more frequent. Similarly, Kasarskis et al., 2001 noticed that experts' pilots (1500 - 2150 flight hours) perform more fixations and have shorter dwell times than novices (40 - 70 flight hours). They postulate that experts have more structured visual patterns than novices. Lorenz and Biella, 2006 have shown that experts (from 3000 to 10300 flight hours) spend more time looking outside the cockpit compared to novices (13-500 flight hours) during a taxiing task. Furthermore, the importance of efficient visual scanning strategies was highlighted in a study involving fighter pilots flying high speed low altitude flights. In this study, the pilots who achieved the best flight performance made shorter fixations to the heads-down tactical display and alternated more frequently between the tactical display and the outside world. Similar results were found in experts ($>$ 1000 hours) and novices (200-400 hours) playing flight simulation games.

3.5 Eye Gaze metrics for aviation context

Classical eye movements measures such as fixation duration, dwell time, or the number of fixations, provide relevant data when analyzing visual scanning strategies. However, as we have seen in the previous section, these metrics are not satisfactory to reflect the visual scanning as these metrics often involve time-averaging operations, thus, neglecting the information regarding the sequence of instrument scanning. Consequently, a rich part of the data that reflects the dynamic of the deployment of the attention processes is lost or not fully exploited. Numerous other metrics are available to explore and characterize in more depth visual scanning strategies. One approach to analyze high-order patterns of eye movements is to analyze transitions between cockpit instruments (Fitts, Jones, and Milton, 1949; Glaholt, 2014; Hayashi, Beutter, and McCann, 2005, another one is to derive global pattern metrics such as gaze entropy (see Glaholt, 2014 for a review).

3.5.1 Transition Matrices

Several metrics allow quantifying whether visual scanning is narrow or wide. Transition matrix probabilities (figure 3.5) is one of them: they contain the information about how often a transition from one AOI to another occurred based on subsequent dwells of the visual scan.

This method provides a data representation that can also lead to the development of stochastic and queuing models (Goldberg et al., 2002) of the pilot's scanning in the cockpit.

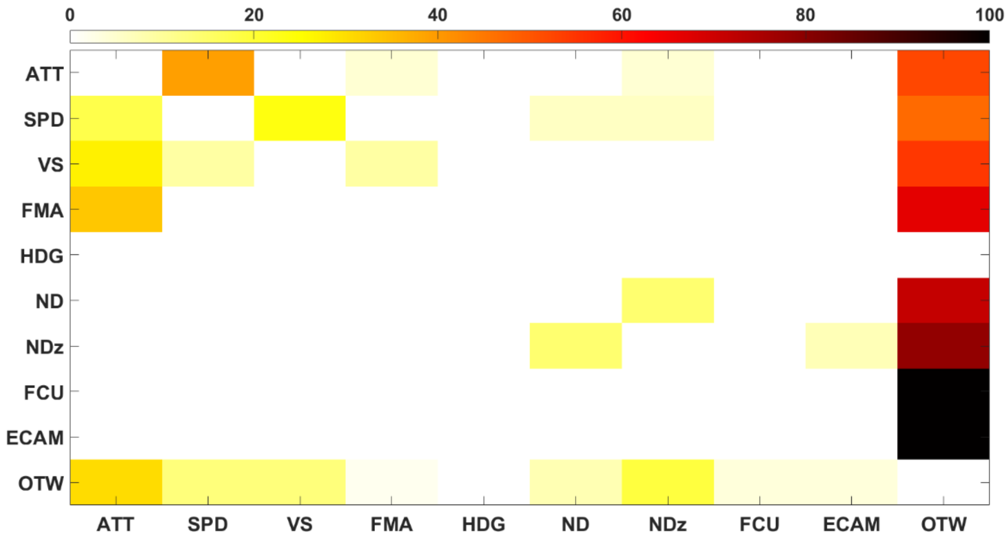


Figure 3.5: Transition matrix of one participant during landing. The colormap represents the probability (in %) of transitioning from one AOI to another.

This method can be extended to three dimensions by considering the location of the previous two dwells, which Norris, 1998 have described as a second-order Markov chain. D. H. Jones, 1985 showed that transitions matrices are sensitive to flight maneuvers in a study involving commercial pilots during various flight phases. Based on the transition matrices, Hayashi et al., 2005 proposed in 2005 a Hidden Markov Model approach corresponding to different flight tasks. Its works were used afterward to model the dwell patterns of the space shuttle crew (Hayashi, 2004). Another method for quantifying the visual scanning is the transition matrix density. Introduced by Goldberg and Kotval, 1999, the transition matrix density can describe the dispersion of attention over time Ognjanovic, Thüring, Murphy, and Hölscher, 2019. Transition matrix density provides a single quantitative value by dividing the number of active transition cells (i.e., those containing at least one transition) by the total number of cells. An unusually dense transition matrix (large index value), with most cells filled with at least one transition, indicates a dispersed, lengthy, and wandering visual scan (this can reflect an extensive search on a display). A sparse matrix (small index value) indicates a more efficient and directed search Holmqvist et al., 2011.

3.5.2 Sequence analysis

The AOI sequence analysis approach allows measuring the extent to which the time sequence of eye movements is ordered or random during a flight. We use the broad term "visual scanning" to describe visual scanning made up of an at least one dwell to one area of interest (AOI), followed by a transition, and a dwell to another AOI; "visual scanning pattern" is used when the visual scanning is made up of repeated sequences of a given "visual scanning". The entropy measure considers the sequence of dwells location without considering the dwell duration. Gaze entropy is one of the methods used to compare visual scanning behavior Diaz-Piedra et al., 2019; Krejtz et al., 2015; Shiferaw, Downey, and Crewther, 2019. When applied to eye tracking data, transition entropy describes the amount of information needed to describe the visual strategies. Harris Sr, Glover, and Spady Jr, 1986 also report that eye movement entropy rate decreases as pilot mental workload increases, and also that novice pilots exhibit a low entropy rate early in training, but as training progresses, their entropy rate increases to match that of experts. Another pattern metric is known as Nearest Neighbour Index (NNI) and was introduced by Di Nocera et al., 2007. Nearest neighbour index is a measure of spatial clustering and is computed by summing the distances of each fixation to its nearest neighbour and by dividing this sum by the average distance between fixations derived from a uniform random distribution. Values less than one indicate departures from a random spatial distribution. Note that while NNI detects departure from spatial randomness, it not sensitive to the degree of randomness in the sequence of fixation or dwell locations (c.f. entropy). Di Nocera et al., 2007 found that NNI varied across phases of simulated IFR flight, showing the least random (most clustered) distribution of fixations during cruising flight and the most random (least clustered) distribution during take-off and landing which are expected to have the highest mental workload.

3.6 A flying Assistant Eye-Tracking based to overcome monitoring issue

In the first part of this chapter, we have seen several physiological and anatomical characteristics of the eye, then, we have been interested in the exploitation of certain ocular features that allow to identify the eye and to deduce the projection of an operator's gaze from it. Finally, we focused on the interest that such a device could have in the aviation field, notably through the strategies of visual scanning in the cockpit. At the interface between Chapter 1 and Chapter 2, the next chapter addresses the issue of a flight assistant based on gaze tracking data to overcome the monitoring problems at landing. Peysakhovich et al., 2018 have defined and identified four stages of eye-tracking technology integration in the piloting activity. A flowchart of the four stages is presented in figure 3.6, where the eye and flight data are recorded (Stage I) and proceed to form a visual behavior database. This database can be used to enhance pilots' training (Stage II) and to check the consistency of the visual behavior according to the flight context. If an inconsistency is detected, we can adapt flight deck (Stage III) or aircraft systems (Stage IV).

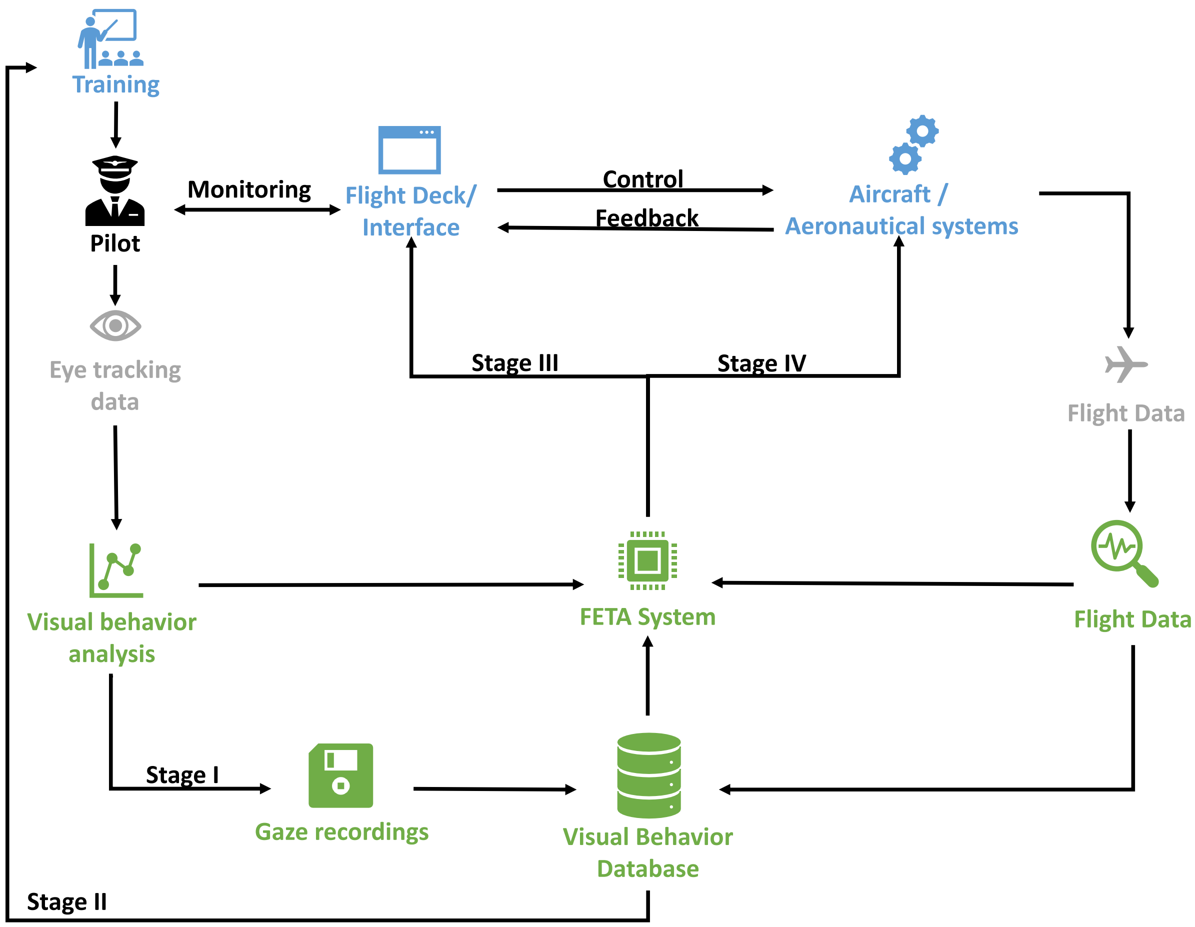


Figure 3.6: Flowchart of the eye-tracking integration reported and modified from Peysakhovich, Lefrançois, Dehais, and Causse, 2018, Blue rectangles correspond to the existing elements. Green color indicates the elements that are to be integrated.

3.7 Hypothesis

The objective of this study was firstly to establish standards related to the activity of monitoring the flight parameters displayed in the cockpit. The aim was to establish a database of standard visual behaviour. Secondly, a virtual assistant (FETA system) was implemented on board to notify the pilot when he deviates from these monitoring standards and to evaluate the effectiveness of this assistant and its impact on the operator. Third, this assistant was tested on 5 pilots. The majority of accidents in aviation concern human errors involving monitoring lapses, as seen in Chapter 1. These monitoring lapses lead to deviations in certain crucial flight parameters such as speed, heading and vertical speed.

- We assume that there are standard eye movements among expert pilots for consulting flight instruments, which would allow the construction of a database based on their eye activity,
- Secondly, we assume that it is possible to extract relevant information (considering dwell time on instrument) from this database to avoid deviation of flight parameters due to insufficient instruments consulting,
- Thirdly, this relevant data can be integrated into an assistant that can be used to notify the pilot when the current eye activity deviates from these standards,
- Fourthly, we assume that this assistant will reduce the occurrence of these deviations without having a detrimental impact on the operator.

Part II

Scientific contributions

CHAPTER 4

FETA: AN EYE TRACKING BASED FLYING ASSISTANT

Sommaire

In this chapter, we focus on the design of an eye-tracking based flying assistant. This system, called FETA (for Flight Eye Tracking Assistant), compares the current visual scanning of a pilot with a database of "standard" visual circuits. The pilots' data will be analyzed in more detail in Chapter 6. If the current visual scanning 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 chapter describes the development and evaluation of the FETA system. In particular, we evaluated the impact of FETA on situation awareness, subjective workload, flight performance, and visual scanning (e.g., Dwells Times in this case).

4.1 FETA system development and visual data base building

The main purpose of the FETA system is to warn the pilot when he does not look 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 (VBD) in the cockpit with a sample of 16 airline pilots. They performed approach-landing phases in a flight simulator while their eye movements were recorded. We also ensured that their flight performance remained in the standard safety thresholds.

4.1.1 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 hrs, SD = 2,911 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) (see Appendix A).

4.1.2 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 instructions 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.

4.1.3 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 4.1. It simulates an Airbus A320 with a glass cockpit. The simulator includes pilots' seats, sidestick controllers, throttles, trim wheels, and rudder pedals.

4.1.4 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° - 1° 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 4.2. 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 200 ms.

Figure 4.1: The PEGASE flight simulator used during FETA development and assessment.

4.1.5 Experimental conditions

The 16 pilots performed 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 (see figure 4.3). In particular: maintain a vertical speed between +500 ft/min and -800ft/min, a speed of 130 knots, and a heading of 143°.

4.1.6 Flight parameters

Firstly, we checked the flight performance of the pilots, assuming that correct flight performance is likely correlated with efficient cockpit monitoring. Figure 4.4 shows the mean flight parameters deviation for vertical speed, speed, and heading during the landing task. Flight performance of each pilot was adequate. The average vertical speed was in the correct range, and average speed and average heading were very close to the target values.

Figure 4.2: 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).

4.1.7 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 approach-landing scenario. Mean non-dwell times were calculated for each AOI. While dwell times represent the time during which an individual gaze inside an AOI [19], non-dwell times correspond to the period of time during which an individual does not look at an AOI, see figure 4.5.

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 equation 4.1.

$$\Phi_{\text{threshold}} = \mu_{\text{NDT}} + \sigma_{\text{NDT}} \tag{4.1}$$

This metric indicates the maximum non-dwell time tolerance for each AOI (i.e., beyond which insufficient monitoring is diagnosed). Non-dwell time can be considered as mean non-monitored duration for each instrument, see Table 4.1.

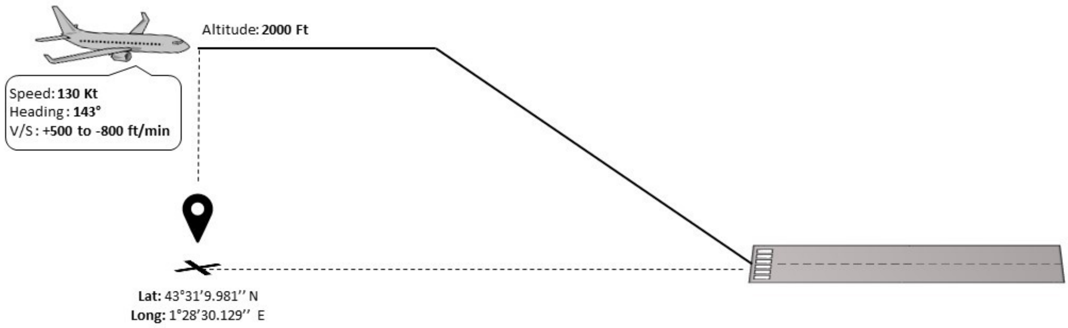


Figure 4.3: The landing scenario with the flight parameter values that pilots had to maintain.

AOI	μ_{NDT}	σ_{NDT}	$\Phi_{threshold}$
Attitude	1.91	2.11	4.02
Speed tape	5.28	7.50	12.78
Vertical speed	3.52	3.36	6.88
ND	12.93	18.22	31.15
HDG	14.73	17.33	32.06
ECAM	12.46	11.22	23.68
FMA	14	16.9	30.9

Table 4.1: Visual behavior database (VBD) containing Mean Non-dwell Time, Standard deviations and Thresholds calculated for each AOI in seconds.

4.1.8 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 4.6.

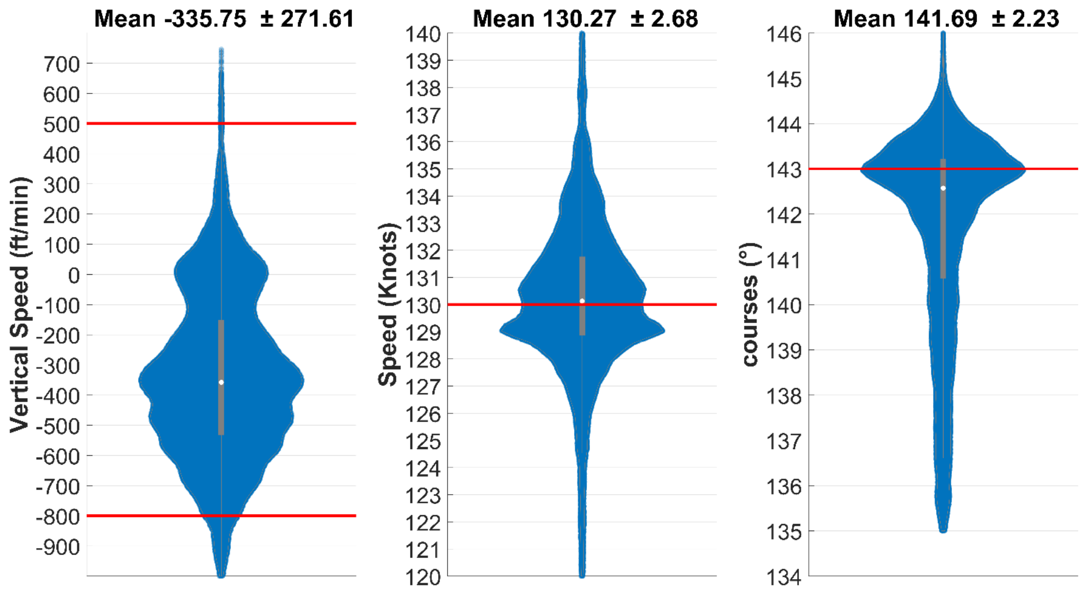


Figure 4.4: Violin plot of flight parameter deviations during the landing task. The red lines correspond to target values, as given by the experimenter before the flight scenarios. N = 16.

The features of FETA are:

1. AOI Monitoring Panel (on the left of figure 4.6). 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 4.6). 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 4.6). This heat map panel indicates the proportion of fixation times on the AOI since the beginning of the flight.
4. Timer (center of figure 4.6). 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 4.6). 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.

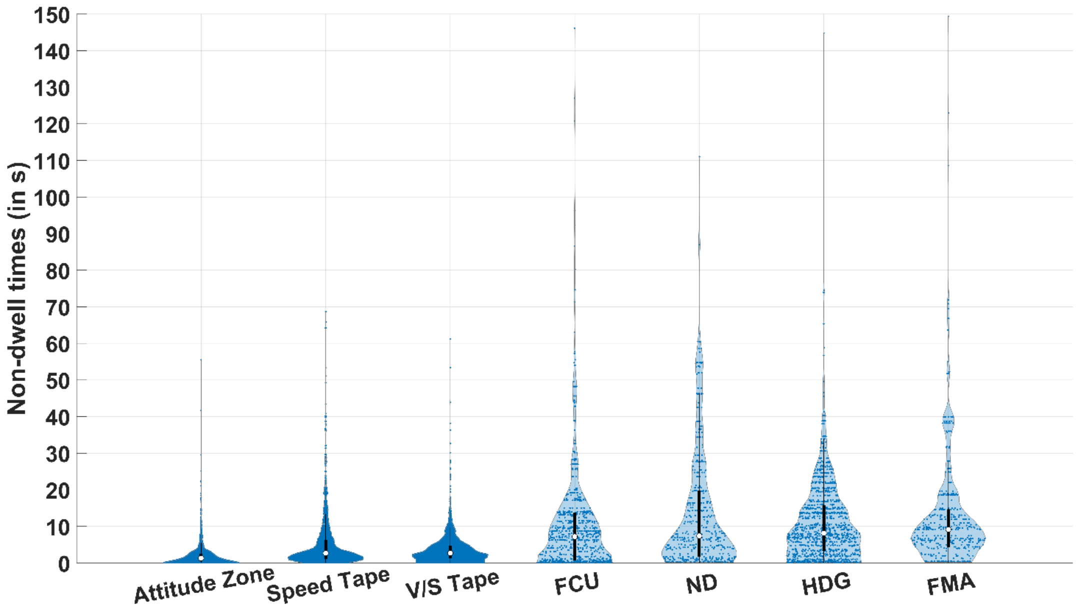


Figure 4.5: Violin plot of non-dwell times during the landing task for the main AOIs. N = 16.

6. Flight Parameter Indicators (at the bottom left of figure 4.6). 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 4.6). 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").

4.2 FETA Human Factors assessment

After the development of FETA and the building of the database, we conducted a separate experiment to evaluate the FETA system. In particular, we evaluated its impact on mental workload, situation awareness, flight performances, and cockpit monitoring. To perform a preliminary evaluation, five pilots were submitted to three different scenarios varying in terms of monitoring difficulty.

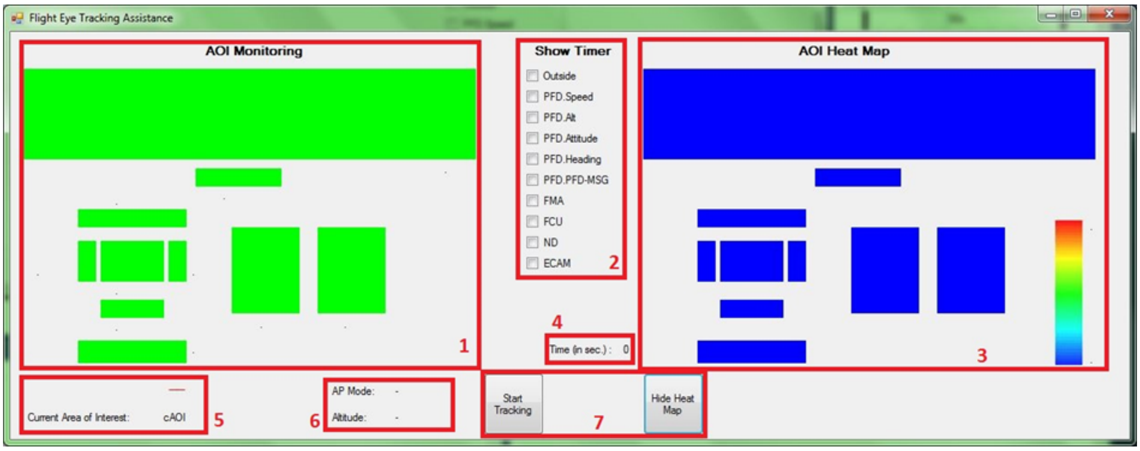


Figure 4.6: FETA interface with its 7 different features.

4.2.1 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 29 years old (range: 23-40). Their total flight experience ranged from 500 to 1,500 hours (M = 976 hrs, SD = 245 hrs). Pilots were not paid for their participation. The experiment was approved by the Research Ethics Committee (CER, n°2019-131, see Appendix A).

4.2.2 Procedure

The procedure was essentially the same as during the FETA calibration (building of the database), 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".

4.2.3 Apparatus

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

4.2.4 Experimental conditions

Pilots performed two times three different randomized landing scenarios. The first scenario (control scenario) 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 (easy dual-task scenario) or every 0.2 Nm (hard dual-task scenario). 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: situation awareness measures using SART (Taylor, 2017), and workload Instantaneous Self-Assessment (ISA) (Tattersall and Foord, 1996), which is a subjective scale ranging from 1 to 5. The latter allows assessing the overall workload. After the flight scenarios, open interviews were conducted to garner the various opinions of the pilots according to the system.

4.2.5 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 (control, easy dual-task, hard dual-task) as an within-subject factor.

4.2.6 Subjective results

Figure 4.7 shows the SART results. A higher SART score indicates better situation awareness. On average, FETA seemed to disturb the situation awareness when flying context was easy (control and the easy dual-task scenario), but it tended to be the opposite when flying context was more complex (hard dual-task scenario). As presented in figure Tattersall and Foord, 1996) did not show a marked difference with or without the FETA system. However, in an easy flying context (control scenario), the FETA system seems to induce more workload and this trend is reversed when flying context is more difficult (both easy and hard dual-task scenarios).

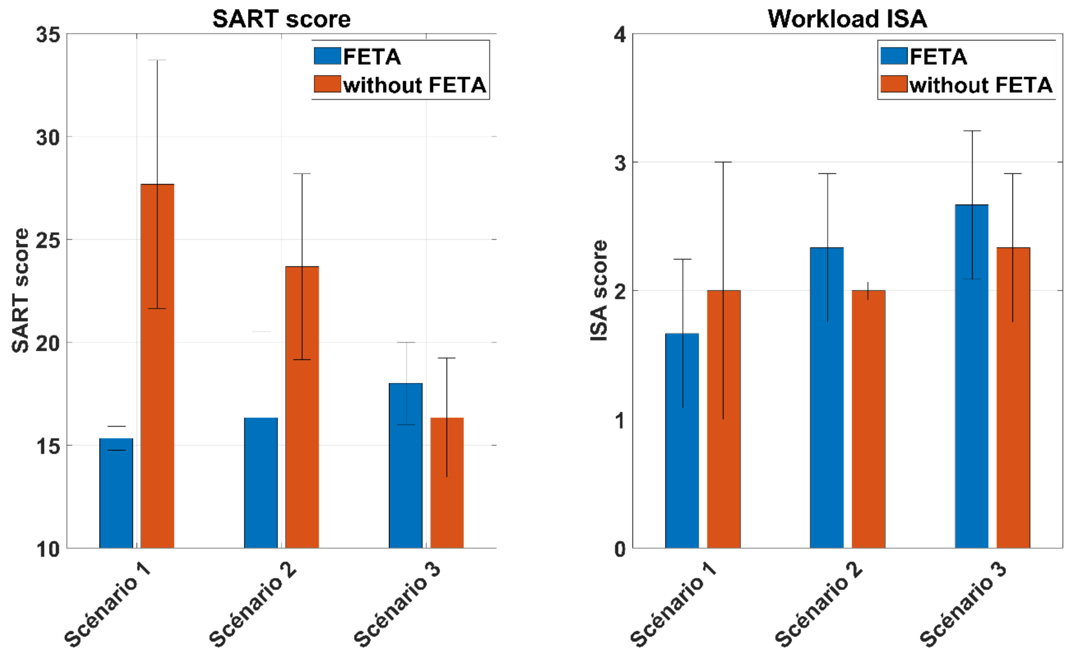


Figure 4.7: Left: SART results (higher the values, better the situational awareness); Right: ISA results (lower is the value and lower is the subjective workload). All three scenarios with and without the FETA system are showed. N = 5.

4.2.7 Flight performance results

Figure 4.8 shows flight parameter deviations. During the easy scenario (control scenario), pilots had higher speed deviations with FETA than without. Concerning the heading in the difficult condition (hard dual-task scenario), pilots had on average lower heading deviations with the FETA system than without.

4.2.8 Eye tracking results

Figure 4.9 shows the percentage dwell times on each AOI for all scenarios with and without FETA system. The Wilcoxon-Mann-Whitney non-parametric 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.

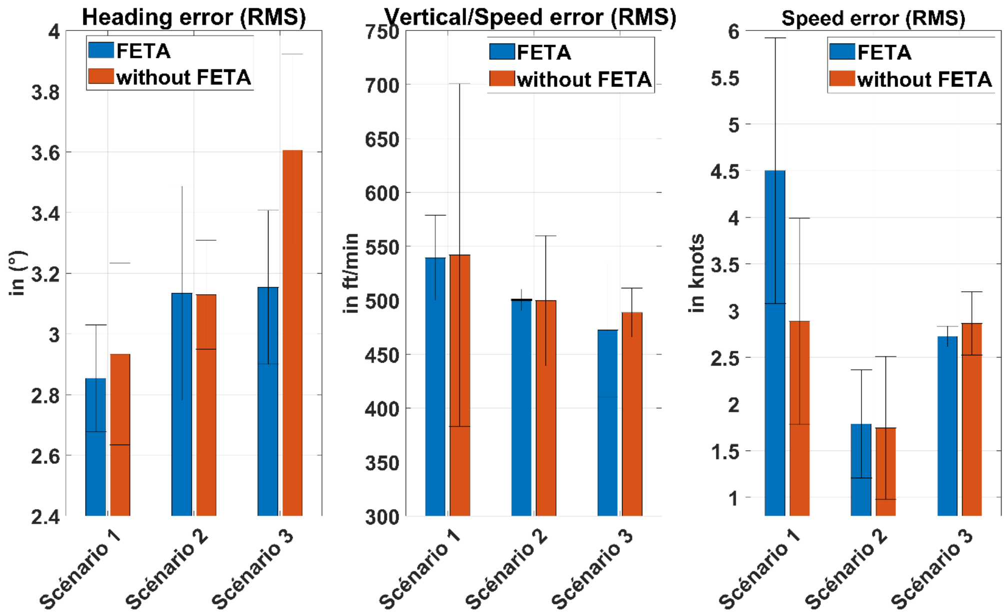


Figure 4.8: Root Mean Square (RMS) of the flight parameters for each scenario with and without FETA (the higher the value, the lower the performance). N = 5.

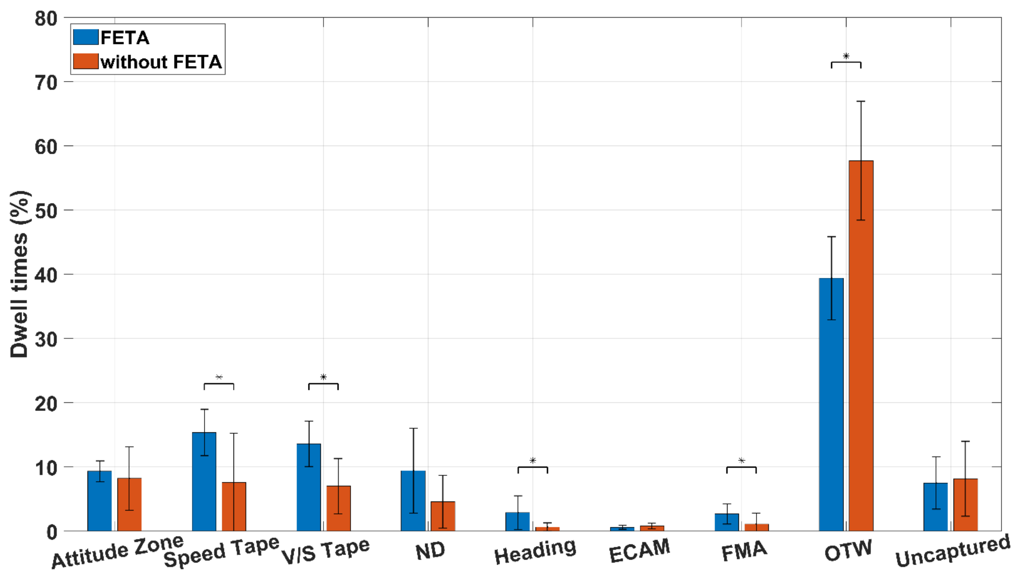


Figure 4.9: Bar plot of percentage dwell times on each AOI. All three scenarios with and without the FETA system are showed. N = 5. (*p<0.05, Wilcoxon Mann-Whitney test).

4.3 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 professional 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 professional airline pilots. Overall, this first version of FETA demonstrated mixed results. First, results showed that there was no clear improvement in the maintenance of the flight parameters during the landing (speed, vertical speed, and heading) when FETA was activated. 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 the flight scenario was difficult. In particular, situation awareness seemed slightly improved by FETA in the hard-dual task scenario. Eye tracking results were more favorable to FETA, with an increase of the time spent on some instruments subjected to the FETA audio notifications 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 at these instruments impacted the time spent on the window. Most likely, FETA was efficient to redirect attention toward the critical flight instruments thanks to the vocal alarm triggered when the visual circuit deviated too much from the database. Despite this positive result, our experiment sheds light on several issues that should be addressed in the future. Open interviews with the pilots allowed revealing some areas of improvement. For example, the use of the auditory modality is not necessarily 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 integrate both flight parameter values and eye movements in FETA. Indeed, it would be more appropriate to trigger notifications when both visual scans and flight parameters deviate too much from standards. For example, when speed decline too fast etc. This would help avoid triggering spurious notifications (useless auditory notifications from FETA), which was one of the main problems raised by the pilots during the debriefing. More generally, the FETA system should incorporate 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 Sommer and Golz, 2010 or considering the frequency of eye blinks Stern, Boyer, and Schroeder, 1994. Future studies should consider these improvements and assessing FETA during complex flight phases with a higher number of pilots. Consequently, as we believed that the FETA system should incorporate other eye tracking metrics, we explored in next chapter other metrics allowing us to qualify dispersion of visual activity (focal mode vs ambient mode) by modifying the well-known K coefficient defined by Krejtz, Çöltekin, Duchowski, and Niedzielska, 2017. In particular, this metric is important since overfocalization is one of the main problems in aviation, and it is difficult to capture with standard metrics such as dwell times this behaviour since it involves dynamic gaze comparison to detect it. Next chapter focuses on the establishment of this metric in aviation context. Chapter 5 proposes a framework for eye movement data analysis techniques in the cockpit to deeply quantify visual scanning strategies in novices and experts. The FETA system should incorporate these metrics.

CHAPTER 5

DISCERNING DYNAMIC OF VISUAL BEHAVIOR TO ENHANCE A FLIGHT EYE TRACKING ASSISTANT

Sommaire

We saw during the human factors evaluation of the FETA system, the need to establish new metrics to quantify the visual attention of pilots. Our evaluation showed that setting dwell times thresholds on flight instruments was too binary (too simple) to trigger relevant alarms. This chapter focuses on the search for a new metric that can be used in the cockpit and based on the eye tracking literature.

5.1 Context and Hypothesis

The distribution of the pilot's visual attention on various AOIs can be inspected by analyzing their eye movements during real or simulated flight tasks. Visual attention can be measured by analyzing fixations and saccadic eye movements Hoffman and Subramaniam, 1995, and researchers proposed methods to disambiguate ambient and focal attention modes. While the ambient mode is associated with scene exploration, the focal mode refers to a more local inspection of the information. For example, Follet, Le Meur, and Baccino, 2011 showed that early stages of picture viewing are associated with long saccades (probably to globally explore visual content), and then, the saccade amplitudes decrease over time. Similarly, Antes, 1974 showed that fixations become longer whereas saccade amplitudes decrease as time passed during free picture exploration. Participants tend to make many fixations at the beginning of the task to get an overall knowledge of the scene, and then focus progressively on some areas. Furthermore, the connection between saccade amplitudes and fixation duration is of particular interest due to its indication of the two dominant modes of visual processing. In the cockpit, it can be assumed that pilots switch from one mode to another, depending on the number of channels of information the pilot has to monitor during a given flight phase. Also, a highly dominant focal attention mode may reveal abnormal cockpit monitoring, analog to the attentional tunneling phenomenon. Considering Wickens' definition of attentional tunneling Wickens and Alexander, 2009, a straightforward way to identify this phenomenon is noting when operators omit unexpected events (e.g., they do not react to alarms) and persevere in their current action pattern. Such an expert approach requires analysis of the operators' behaviors to infer their attentional state (e.g., actions on the user interface reaction time). A complementary approach is to derive attentional tunneling from the measurement of physiological signals and ocular activity (Dehais et al., 2012; Regis, Dehais, Tessier, and Gagnon, 2012). Several authors have demonstrated that attentional tunneling results in fewer scanned areas of interest (AOI) on the user interface (Wickens, 2005), a decreased saccadic activity (Tsai, Viirre, Strychacz, Chase, and Jung, 2007), long eye fixations (Cowen, Ball, and Delin, 2002), and the absence of ocular fixations on relevant cues (Thomas and Wickens, 2004). Far from detecting this behaviour in a formal way, the aim of this study was rather to use the eye-tracking literature to qualify (focal vs ambient) the pilots' eye behaviour in the cockpit. This chapter is motivated by the many studies on visual information processing that attribute relationships between saccade amplitudes and fixation durations Unema, Pannasch, Joos, and Velichkovsky, 2005; Velichkovsky, Joos, Helmert, and Pannasch, 2005. Our goal is to extend works related to ambient and focal attention modes to a cockpit monitoring context. We compare the formerly defined $\mathcal{K}$-coefficient Krejtz, Duchowski, Krejtz, Szarkowska, and Kopacz, 2016 with a new approach using AOI and dwell time. The contribution of this work is three-fold:

- First, we investigated a modified $\mathcal{K}$-coefficient considering both the dwell times and the dwell transitions,
- Second, since visual behavior are different following flight phases (Scannella, Peysakhovich, Ehrig, Lepron, and Dehais, 2018; Wang and Fu, 2014), the proposed approach was tested on three different flight phases (i.e, Take-off, Cruise, and Landing),
- Third, we compare our approach with the standard $\mathcal{K}$-coefficient on data obtained from an experiment set up with professional aircraft pilots.

We provide the background of this work, and how our approach uses AOI and dwell time to alter and increase the benefit of the state-of-the-art metric. Since no study on the investigation of the $\mathcal{K}$-coefficient based on AOI and dwell time can be found, the present research attempts to further investigate this question. Then, we present the experimental design and method used to evaluate our approach. We finally discuss the possible generalizability for future works.

5.2 Visual searching task as background for aviation outcome

Fixation duration is a relatively changing variable that can range from about 100 milliseconds to many seconds (Pannasch, Helmert, Roth, Herbold, and Walter, 2008). This high variability is both intra-task and intra-individual Henderson and Luke, 2014. In many areas, searching for novel ways of analyzing eye movements data has gathered dozens of metrics (Sharafi, Shaffer, Sharif, and Gueheneuc, 2015). One important objective of eye movements analysis is to spot particular events from the large amount of recorded data (Holmqvist et al., 2011). We assume that the two modes of attention, ambient (exploration) vs. focal (inspection), can be routinely observed during a flight. During highly dynamic flight phases such as take-off and landing, cockpit monitoring generally requires the pilots to distribute their attention toward many channels of information (Velichkovsky et al., 2005). They perform short fixations on flight instruments and switch very frequently to other ones to update as much as possible their knowledge about critical flight parameters (altitude, speed, etc.) (van de Merwe et al., 2012). However, when an unusual event occurs or when the pilot is performing a demanding task, such as managing a system failure or performing a precision approach using course and glide path guidance systems, longer dwells and much lower saccade numbers could be observed (Di Nocera et al., 2007). As mentioned previously, an extreme and abnormal focalization of attention can be called attentional tunneling (Wickens and Alexander, 2009). Our work tries to uncover these ambient and focal modes in the cockpit monitoring context, assuming that they do not progress from ambient to focal during the chronology of the (i.e. from take-off to landing) flight such as over the time course of visual exploration (Pannasch et al., 2008), but the mode rather depends on the flight phase, the automation mode, or the particular tasks at hand. In a recent study, Krejtz et al., 2016 introduced a coefficient κ, a dynamic indicator which allows differentiating between ambient and focal attention. Thereafter, they used the coefficient $\mathcal{K}$ to investigate the dynamics of the visual patterns when operating cartographic tasks (Krejtz et al., 2017). Evidence for the accuracy of the coefficient $\mathcal{K}$ has been shown. Using the $\mathcal{K}$-coefficient, A. Duchowski and Krejtz, 2017 proposed a novel ambient/focal colorization which shows the dynamic interplay between the focal and ambient modes of visual information processing. Pannasch et al., 2008 evaluated the relationship between fixation durations and saccade amplitudes in free exploration of pictures and scenes during early and late phases of scene perception. Velichkovsky et al., 2005 studied scene perception in static and dynamic environments and explored the recognition of the focal mode of processing and ambient exploration. However, no investigation of this metric on AOI and dwell time has yet been performed, though many studies are based on parts and small areas in a scene. The flight deck of commercial aircraft has a dense amount of information displayed on different flight instruments. In contrast to paintings or maps, used as visual scenes in previously cited articles, information on the cockpit is not continuous, it is spatially distributed in several locations. The cockpit is not a uniform scene, but it rather consists of an aggregation of instruments, each having a particular and unique role. Measuring raw fixations times and saccades length does not allow focusing specifically on these instruments. Indeed, a fixation located outside a relevant area (i.e., that does not correspond to a flight instrument) is considered in the same way as a fixation or saccade directed towards a flight instrument. A coefficient based on AOI is, of course, a bit more complex to compute as it requires categorizing the visual scene (in different AOIs), but it somehow helps restricting the analysis to relevant data. While using the coefficient $\mathcal{K}$ with fixation duration and saccade amplitude, one seeks to find the optimal value that best discerns the ambient/focal attention. Our approach, in addition to incorporating the strengths of this metric, considers the different areas of interest, the corresponding dwell time, and the transitions between AOI.

5.3 Coefficient K and novel approach

The coefficient $\mathcal{K}$ allows collecting pairs of fixation duration and saccade amplitude and provides a value that classifies attention as ambient or focal. More precisely, it takes into account pairs consisting of the current fixation and the next saccade. Krejtz et al., 2016 developed a set of experiments where different metrics were evaluated, and concluded that to discern between ambient and focal attention, a coefficient of the form 5.1:

$$\kappa_i = \frac{\mathbf{d_i} - \mu_d}{\rho_d} - \frac{\mathbf{a_{i+1}} - \mu_a}{\rho_a}, \qquad \kappa = \frac{1}{n}\sum_{i=1}^{n}\kappa_i \tag{5.1}$$

could be used, where $\mathbf{d_i}$ is the duration of the i$-th$ fixation and $\mathbf{a_{i+1}}$ the amplitude of the saccade that occurs after the i$-th$ fixation. μ_d, μ_a are the mean fixation durations and saccade amplitudes, respectively, and ρ_d, ρ_a are standard deviations, respectively. The means coefficient is computed for all subjects and configurations (A. Duchowski and Krejtz, 2017; Krejtz et al., 2016). The negative values of the coefficient $\mathcal{K}$ indicate ambient attention (K<0) whereas the positive values indicate a focal mode (k>0). This way, the authors used the visual parameters defined by the fixation duration and saccadic amplitude to classify the viewing modes. Unlike the traditional $\mathcal{K}$-coefficient approaches that use the raw fixations, we used dwell times on defined AOI. All dwell times in the AOI were normalized and used to calculate the coefficient. Similarly, the amplitudes of the subsequent saccades were replaced with the distance between two subsequent AOI. Our goal is to propose a metametrics $\mathcal{K}$. For more clarity, Algorithm 1 shows the different steps of generating the modified $\mathcal{K}$-coefficient. The algorithm begins by calculating the mean dwell times and the standard deviations for each participant and all flight phases. In the same way, the mean distance between two subsequent AOI and the standard deviation is calculated. Then each normalized dwell time in an AOI is subtracted by the distance between this AOI and the next one. The algorithm repeats this procedure for all subsequent pairs. An important criterion of a good metric is to ensure that it could accurately compare to well-established metrics. Therefore, we evaluated the quality of the proposed metrics Modified $\mathcal{K}$ (M$\mathcal{K}$) to the standard $\mathcal{K}$-coefficient (see Krejtz et al., 2016 for more details).

Algorithm 1: Modified $\mathcal{K}$-coefficient (M$\mathcal{K}$)

Data: Dwell time T_i and distances between successive AOI D_i (e.g. $\mathbf{Data} = T_0, D_0, ..., T_n, D_n$)
Result: Modified Coefficient-$\mathcal{K}$ (M$\mathcal{K}$) of the eye movement sequences
begin
  Initialization
  MK = 0
  $\mu_T \longleftarrow$ mean(T_i)
  $\rho_T \longleftarrow$ sd(T_i)
  $\mu_D \longleftarrow$ mean(D_i)
  $\rho_D \longleftarrow$ sd(D_i)
  while $i \leq$ *Data Length* n **do**
    $\mathbf{MK_i} \longleftarrow \frac{\mathbf{T_i} - \mu_T}{\rho_T} - \frac{\mathbf{D_{i+1}} - \mu_D}{\rho_D}$
  end
  $\mathcal{K}\ \mathbf{MK} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{MK}_i$
end

5.4 Method

5.4.1 Participants and Apparatus

Fourteen airline professional pilots, qualified on Airbus A320, were recruited for the experiment. All pilots were males, with a mean age of 42.3 years (SD=3.8 years) and an average of 11,500 flight hours (FH) (SD=1,300 FH). They were briefed on the flight scenario (airport, aircraft weight, configuration, flight plan), but were not introduced to the exact purpose of the research. The pilots were recruited as part of the thesis work with Air France led by Olivier Lefrancois. All 14 pilots were pilots flying (PF), in charge of controlling the flight trajectory of the aircraft. We did not analyze the pilots monitoring (PM) in this study. The experiment was conducted in a full flight simulator Airbus A320 simulator (Thompson, see figure 5.1) provided by Air France group.

It is commonly used for the regular training of professional flight crews. Tobii Pro Glasses 2 (sampling frequency: 100 Hz) were used to record the eye movements of the pilots during the entire flight scenario duration (see Figure 5.1). The cockpit was split into 9 Areas Of Interest (AOI) corresponding to the main flight instruments and the Out of the Window (OTW). They corresponded to 1) the electronic centralized aircraft monitor system (ECAM), 2) the flight control unit (FCU), 3) the flight mode annunciator (FMA), 4) the navigation display (ND), 5) the out of window (OTW), 6) the attitude indicator (ATT), 7) heading tape (HDG), 8) the speed tape (SPD), 9) the vertical speed (VS) as presented in Figure 5.2. Acquisitions and pre-processing have been conducted using Tobii Pro Lab software.

Figure 5.1: Left: Thompson Air France full flight simulator used in this study. Right: View inside the cockpit with pilots wearing Tobii glasses.

Figure 5.2: Illustration of the nine different AOI (in red), they corresponded to 1) ECAM, 2) FCU, 3) FMA, 4) ND, 5) OTW, 6) ATT, 7) HDG, 8) SPD, 9) VS.

5.4.2 Procedure and flight scenario

All pilots performed the same flight scenarios. They started by a take-off from Toulouse runway (32R) and finished by a landing on the same runway. After take-off, flight crews were instructed to climb to 5000 feet, turn left, and intercept the Instrument Landing System (ILS). Then, they were cleared to the approach, performed with standard visibility conditions (runway visual range of 550 meters), and with a significant crosswind (15 knots), see Figure 5.3.

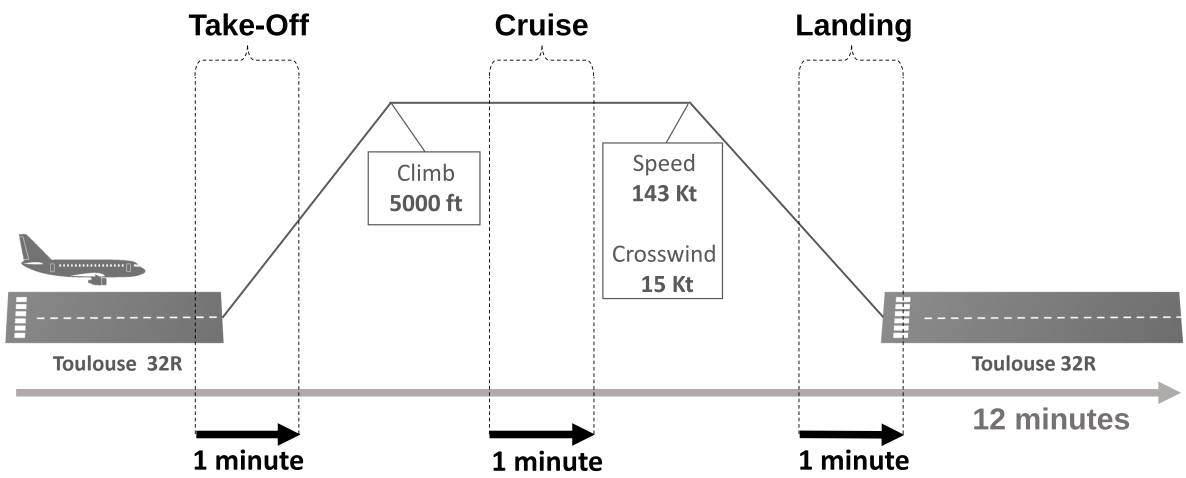


Figure 5.3: The flight scenario with different flight phases corresponding to the specific phases analyzed

Participants were required to fly in compliance with the flight crew operating manual and operator requirements regarding stabilization criteria. Pilots performed four flight scenarios with varying automation levels. The various automation levels have been manipulated by activating or deactivating flight director (FD) and/or auto-throttle at the beginning of the scenario. Flight director is a flight instrument that is overlaid on the attitude indicator, showing to the pilot the required attitude to follow a certain trajectory to which the flight is to be conducted. Auto-throttle is a system that allows to automatically control the power setting of an aircraft's engines. In this paper, we focused on a single flight scenario, during which automation level was high (both flight director and auto-throttle were set to on). We analyzed the eye behavior during portions of 1 minute of each of the different flight phases (take-off, cruise, landing) of this flight scenario, as shown in Figure 5.3. The duration of each phase was relative for each pilot. Especially during the take-off phase, which can be extremely fast (around one to several minutes), depending on the vertical speed chosen to reach the 5,000 ft step. We decided to compare the different phases at the same time scales.

5.4.3 Data processing

Concerning data processing of basic coefficient $\mathcal{K}$, Tobii pro lab software was used for fixations and saccades detection with a velocity-based event detection algorithm. The minimum duration of saccades was set to 22 ms, with a peak velocity threshold of 40°/s. In this paper, we applied a minimum fixation duration of 80 ms for the analysis (i.e., ignoring fixations with very short durations in range [50:80] ms) as in the study of A. Duchowski and Krejtz, 2017, where the K coefficient was established. Concerning data processing of our extended coefficient $\mathcal{K}$, Tobii pro lab software was used to collected AOIs sequences. An in-house script was created to collected dwell times. A matrix corresponding to all possible transitions between AOI was created to measure transition amplitudes between one AOI to another AOI. Dwell times smaller than 150 ms were discarded from this analysis. As mentioned previously, the ambient/Focal coefficient $\mathcal{K}$ was computed by subtracting, respectively, the standardized dwell time on the current AOI from the standardized amplitude of the subsequent transition (transition to the next AOI). For technical reasons, we had access to the eye-tracking data of eleven participants out of the fourteen initial participants.

5.4.4 Statistical analysis

Data were analyzed using Matlab R2019b. Normality tests were conducted separately on fixation durations & dwell times, saccade amplitudes & transition amplitudes, and coefficient $\mathcal{K}$ & modified coefficient $\mathcal{K}$ using the Shapiro-Wilk test. The sphericity assumption was verified using Mauchly's sphericity test. Normality assumptions were respected in all conditions except for the fixation and dwell durations during cruise and landing. Sphericity assumptions were respected for all conditions excepted for fixation duration. We performed one-way repeated measures ANOVA and Bonferroni post-hoc test to examine multiple comparisons. In case of violation of the normality assumption, we used non-parametric Friedman's ANOVA with Wilcoxon signed-rank test (with a Bonferroni-adjusted p-value) to examine multiple comparisons. Greenhouse-Geisser adjustment was used to correct violation of the sphericity assumption when needed. All p-values are reported with this correction. The null hypothesis was determined with α=0.05.

5.5 Results

The Friedman's test revealed no significant effect of the flight phases on the fixation duration, Chi-square value = χ^2 (2) = 0.5454. An ANOVA did not reveal any significant effect of the flight phases on saccade amplitude, $F(2, 20) = 1.568, p = 0.23$. Similarly, ANOVA did not reveal any significant effect of the flight phases on the coefficient $\mathcal{K}$, $F(2, 20) = 1.977, p = 0.16$, see Figure 5.4. The Friedman's test revealed a significant effect of flight phases on dwell duration, with a Chi-square value of χ^2 (2) $= 16.54, p < 0.001$, $\eta^2 = 0.50$. Wilcoxon post-hoc test showed that dwell duration was lower during take-off $(M = 727, SD = 940)$ than during cruise $(M = 1738, SD = 2045)$, and landing $(M = 2109, SD = 2193)$. There was also a significant effect of flight phases on the transition amplitude, $F(1.27, 12.72) = 41.99, p < 0.001$, $\eta^2 = 0.55$. Bonferroni Post-hoc test indicated that the transition amplitude was significantly lower during take-off phase $(M = 32.6, SD = 4.9)$ than during cruise phase $(M = 35.4, SD = 6.3)$, and landing phases $(M = 37.6, SD = 7.0)$. There was no significant effect between cruise phase $(M = 35.4, SD = 6.3)$ and landing phase $(M = 37.6, SD = 7.0)$. An ANOVA showed that the effect of phase on modified Coefficient $\mathcal{K}$ was significant $F(2, 20) = 0.77, p < 0.01$, $\eta^2 = 0.64$. Bonferroni Post-hoc test revealed that the Coefficient $\mathcal{K}$ was significantly lower during take-off $(M = 0.12, SD = 0.28)$ than during cruise $(M = 0.56, SD = 0.44)$, and landing phases $(M = 0.58, SD = 0.38)$. No significant effect was found between cruise phase $(M = 0.56, SD = 0.44)$ and landing phase $(M = 0.58, SD = 0.38)$ on Coefficient $\mathcal{K}$.

5.6 Discussion

We have examined the characteristics of commercial pilots' visual attention during a realistic flight scenario using a modified coefficient $\mathcal{K}$ applied to AOI. Instead of using fixation duration and saccade amplitude, we used dwell times (in AOI) and transitions amplitudes between AOI. As in the original paper of Krejtz et al., 2017, the coefficient $\mathcal{K}$ was used to detect two different modes of visual attention, with ambient mode corresponding to a higher dispersion of the attention, in our current study, a dispersion of the visual attention toward a high number of flight instruments, these instruments being rather distant the ones from the others (high transition amplitudes). On the contrary, the focal mode corresponds to a lower dispersion of attention, namely in our study, the maintenance of the attention on a lower number of flight instruments, mainly located near each other. Using dwell times and transition between AOI (instead of fixation duration and saccade amplitude) helps focusing the analysis on relevant gaze data, $i.e.$ related to piloting activity. It helps filtering irrelevant fixations, (e.g., on the other

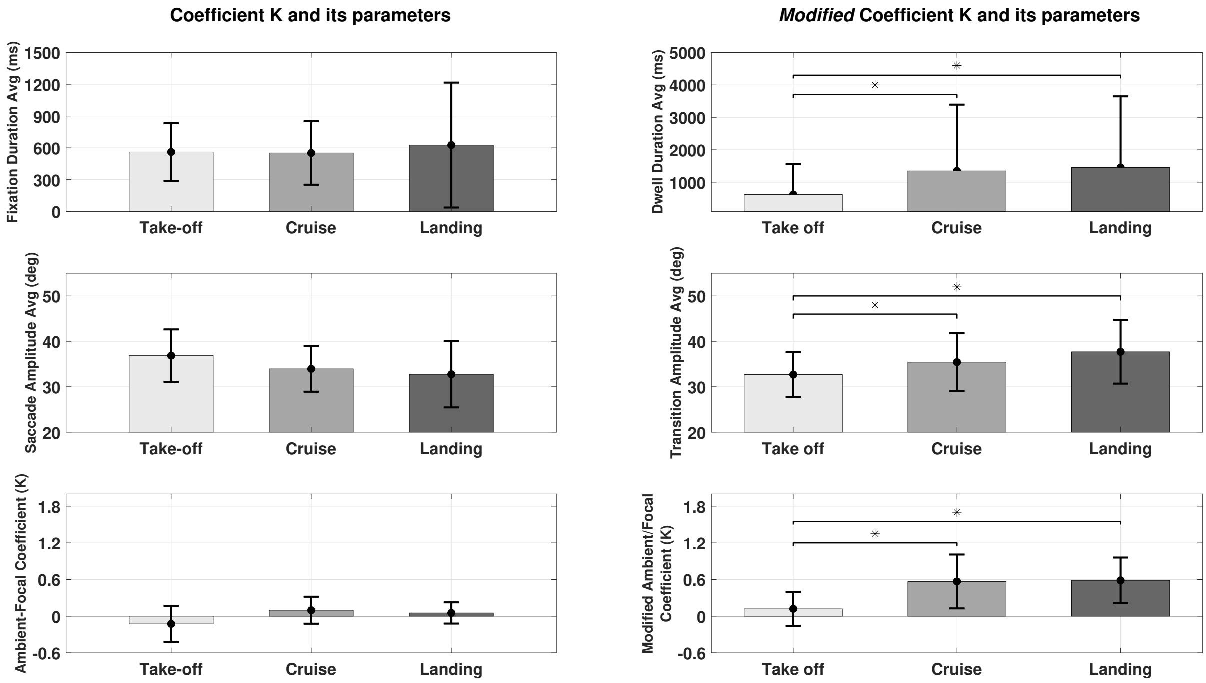


Figure 5.4: Marginally significant effects of different dependant variables on different flight phases (Take-off, Cruise, Landing). Bars represent standard deviations (SD). Left plots represent different metrics used to compute coefficient $\mathcal{K}$; Right plots represent different metrics used for modified coefficient $\mathcal{K}$; $\mathcal{K}$>0 indicate a focal visual attention, whereas $\mathcal{K}$<0 indicate an ambient visual attention.

pilot) and saccades (e.g., the pilot is performing short saccades inside an AOI but still consults the same information displayed on that AOI). Given the fact that the modified coefficient $\mathcal{K}$ was generally higher than 0, we assume that the focal mode dominates in the cockpit context when full automation is set. Results showed that this degree of focalization was lower during the take-off vs. cruise and landing phases. During the take-off, the pilot flying (PF) has to monitor a great variety of instruments (e.g., speed, auto-throttle, attitude) until the stabilization of the aircraft at 5000 ft. On the contrary, during the cruise, the monitoring of the flight parameters is less intense since the parameter values are quite stable. During the landing, the PF has to monopolize his attention toward the PFD and the runway, when the weather is clear (as indexed by the OTW AOI in our study), resulting in a higher coefficient $\mathcal{K}$. We did not find any significant result with the original $\mathcal{K}$ coefficient that might be more suitable for visual search tasks such as cartographic tasks or art viewing. In contrast to exploratory related tasks, pilots are familiar with their environments and do not need to seek information in the cockpit. As a consequence, visual attention is guided by endogenous task-related control (Klein, 2004; Posner, Snyder, and Davidson, 1980) rather than exogenous stimuli Andrews and Coppola, 1999. While this study provides evidence on the benefit of the modified coefficient $\mathcal{K}$, an in-depth evaluation with different AOI coupled with different tasks (Feit et al., 2017) would extend this work. We have shown that using an AOI based $\mathcal{K}$-coefficient to investigate the visual attention of a pilot offers many benefits. The analyses of this coefficient suggested that applying this approach to the cockpit showed that focal attention mode was dominant when automation was high. This could be explained by the fact that flight parameters are automatically controlled by the systems, reducing the need for the pilot to consult many instruments in the cockpit. In this sense, high $\mathcal{K}$-coefficient value may indicate an over-focalization of the attention, an abnormal behavior that can result in poor situation awareness (Woods, 1991). Finally, this indicator makes it possible to qualify the pilot's behavior as focal or ambient. This metric gives precious indicators about the deployment of the visual attention of the pilot. Such information would make it possible to identify abnormal attentional behaviors that do not correspond to the typical visual activity of a particular flight phase or task (e.g., overfocalization). Concerning real-time assessment of the visual behavior, further studies should take into account a large visual database of modified $\mathcal{K}$-coefficient during different flight phases. This database may serve as a reference to which the current $\mathcal{K}$-coefficient could be compared, thus enabling the online use of such a metric. Furthermore, buffering/segmentation with running distributions could be of help to real-time assessment, other interesting metrics can be explored, such as visual scanning structuring, and an assistant must take into account the profile of the pilot, and therefore the sensitivity of the metrics to different profiles, such as their expertise. The next chapter explores this issue with a study comparing novices' vs experts' visual scanning strategies.

CHAPTER 6

PILOT'S EXPERTISE AND VISUAL SCANNING STRATEGIES

Sommaire

This chapter focuses on the need to feed the FETA system with better metrics to distinguish visual scanning strategies. A novice expert comparison was carried out to highlight the differences between novice and expert pilots regarding visual scanning strategies.

6.1 Standard and advanced eye movement metrics to examine pilots' visuals scanning strategies

Classical eye movements measures such as fixation duration, dwell time, or the number of fixations, provide relevant results when comparing novices vs. experts. However, statistical analyses of these metrics often involve time-averaging operations, thus, neglecting the information regarding the sequence of instrument scanning. Consequently, a rich part of the data that reflects the dynamic of the deployment of the attention processes is lost or not fully exploited. Numerous other metrics are available to explore and characterize in more depth visual scanning strategies. The broad term "visual scanning" is used to describe visual scanning made up of an at least one dwell to one area of interest (AOI), followed by a transition, and a dwell to another AOI; "visual scanning pattern" is used when the visual scanning is made up of repeated sequences of a given "visual scanning". In this work, we investigated more particularly three of them, we analyzed high-order patterns of eye movements by considering the 1) transitions between cockpit instruments (e.g.,Fitts et al., 1949,Senders, 1966) 2) the dynamic of visual behavior Krejtz et al., 2017, and 3) the AOI sequences analysis.

6.2 Markov chains

6.2.1 Transition matrix probabilities

They contain the information about how often a transition from one Area Of Interest (AOI) to another occurred based on subsequent dwells of the visual scan. This method provides a data representation that can also lead to the development of stochastic and queuing models (Goldberg et al., 2002) of the pilot's scanning in the cockpit. This method can be extended to three dimensions by considering the location of the previous two dwells, which Norris, 1998 have described as a second-order Markov chain. D. H. Jones, 1985 showed that transitions matrices are sensitive to flight maneuvers. Based on the transition matrices, Hayashi, 2004 proposed in 2004 a Hidden Markov Model approach corresponding to different flight tasks. Its works was used afterward to model the dwell patterns of the space shuttle crew Hayashi et al., 2005.

6.2.2 Transition matrix density

Introduced by Goldberg and Kotval, 1999, the transition matrix density describes the dispersion of attention over time (Ognjanovic et al., 2019). Transition matrix density provides a single quantitative value by dividing the number of active transition cells (i.e., those containing at least one transition) by the total number of cells. An unusually dense transition matrix (large index value), with most cells filled with at least one transition, indicates a dispersed, lengthy, and wandering visual scan (this can reflect an extensive search on a display). A sparse matrix (small index value) indicates a more efficient and directed search (Holmqvist et al., 2011).

6.3 Attentional modes

6.3.1 Modified K-coefficient

Another evaluation of the dispersion of the attention is a novel parametric scale called K coefficient introduced by Krejtz et al., 2016. This metric was created and developed during exploring artwork (e.g., painting) and map viewing (Krejtz et al., 2017) in order to investigate the dynamics of visual scan (focal vs ambient) when operating such tasks. Fore more details, see chapter 5.

6.4 Sequence analysis

The AOI sequence Analysis approach allows measuring the extent to which the time sequence of eye movements is ordered or random during a flight.

6.4.1 Gaze transition entropy

Defined by Shannon and Weaver Shannon and Weaver, 1948, entropy is a measure of lack of predictability in a sequence. Ephrath, Tole, Stephens, and Young, 1980 have noticed an increase of entropy with increasing pilots' mental workload (by adding a secondary task). More recently, van de Merwe et al., 2012 found that entropy increased as a result of cockpit instrument failure, conditions that most likely produce an increased mental workload. More recently, Diaz-Piedra et al., 2019 indicate that gaze entropy can serve as a sensitive index of task load in aviation settings. This metric assess the structuration of the gaze when applied to AOI sequences. When applied to eye tracking data, transition entropy describes the amount of information needed to describe the visual strategies.

6.4.2 Lempel-Ziv Complexity

Initially, Lempel-Ziv Complexity (LZC) was defined by Lempel and Ziv in 1976. This method was a data compression algorithm computing the minimum number of bits from which a particular message or file can effectively be reconstructed. This measure of complexity is different from entropy by its approach. The Shannon's approach is interested in the minimum expected number of bits to transmit a message from a random source of known characteristics through an error-free channel (Burgin, 2017). Although, a lot of study found correlation between both approaches (Grunwald and Vitanyi, 2004; Leung-Yan-Cheong and Cover, 1978). The concepts of entropy rate and Lempel-Ziv Complexity are closely related since time series with high entropy rate tend to generate more complex sequences. A large amount of studies investigated this algorithm on electroencephalographic data (Bai, Liang, and Li, 2015), electrocardiographic data (Zhou, Zhang, and Gu, 2011), electromyographic data (Talebinejad, Chan, and Miri, 2011), and magnetoencephalographic data (Fernández et al., 2011) to detect the arising rate of new patterns along time series.

6.4.3 N-gram/K-mer analysis

Counting n-gram or k-mers (substrings of length k in DNA sequence data) is an essential component of many methods in bioinformatics, including for genome and transcriptome assembly, for metagenomic sequencing, and for error correction of sequence reads (Melsted and Pritchard, 2011). Basically, an N-gram model predicts the occurrence of an AOI based on the occurrence of its N – 1 previous AOI. So here we are answering the question – how far back in the history of a sequence of AOI should we go to predict the next AOI ? For instance, a bigram model (N = 2) predicts the occurrence of an AOI given only its previous AOI (as N – 1 = 1 in this case). Similarly, a trigram model (N = 3) predicts the occurrence of an AOI based on its previous two AOI. By using this method, it is possible to count the occurrence of N-gram AOI and their occurrence for each pilots, and thus it allows to compare for each N-gram the intra-group patterns consistency.

Figure 6.1 presented an Overview of the different visual scanning metrics.

AOI based approaches	Visual scanning metrics	Formula	Definition	Strengh	Shortcoming
Markov chains	***Transition Probability Matrix (TPM)***	$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}, \quad i, j \in S,$	transition frequency from one AOI to another one occured based on X series of dwells.	allow you to visualize which AOIs have strong links with each other.	when more than 3 AOIs are displayed, it is particularly difficult to visualise.
	Transition Density Matrix (TDM)	$p_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}, \quad i, j \in S,$	ratio of non-null transitions cells to the total number.	allows to see if some AOIs are not observed (useful to compare which areas have been explored)	does not take into account the order of the visual scanning.
Attentional modes	***Modified K-Coefficient***	$\kappa_i = \frac{d_i - \mu_d}{\rho_d} - \frac{a_{i+1} - \mu_a}{\rho_a},$ $\kappa = \frac{1}{n}\sum_{i=1}^{n} \kappa_i$	dispersion of the visual attention (focal vs ambient), by calculating a score between dwells and preceding transitions. Positive value of K reflects focal processing whereas negative reflects ambient mode.	usefull to detect too much focusing on a flight instrument (e.g., attentional tunnelling) in a sliding time window	z-score comparison required exhaustive database of all dwells to be compute
Sequence analyses	***Gaze Transition Entropy***	$\hat{H}_t = -\sum_{i \in S} \pi_i \sum_{j \in S} p_{ij} \log_2 p_{ij}.$	amount of information needed to describe the dwell transitions.	useful for comparing between subject with a single quantitative value	loss of qualitative information, the same entropy value may be associated with visual scanning that do not take into account the same AOI.
	LZC	$C(n) = \frac{c(n)}{b(n)},$	complexity of the dwells patterns by data compression.	provide a code book with all the patterns encoutered and their occurrence, correlate with gaze transition entropy	accounts for almost similar visual scanning in different ways (e.g., 1 - 11 - 1 - 11 =/ 11 - 1 -11 - 1)
	Common N-Gram sequences	$P(w_n \mid w_{n-N+1}^{n-1}) = \frac{C(w_{n-N-1}^{n-1} w_n)}{C(w_{n-N-1}^{n-1})}$	number of common sequences in each group	useful for comparing visual scanning between group or for comparing with an expert group	difficulty in highlighting important visual scanning patterns

Figure 6.1: Overview of the different visual scanning metrics classified by approaches

6.5 Objectives and hypotheses

The objective of this chapter is to provide a framework for eye movement data analysis techniques to deeply quantify visual scanning strategies in novices and experts. These eye movement metrics and algorithms are examined in light of the results of an experiment involving novice and expert pilots during a landing scenario performed in a flight simulator. We examined the impact of expertise and the difficulty of the flight scenario on the visual attention allocation among flight instruments. The participants performed three times the same landing scenario with varying difficulty conditions. This is the same landing task than in the experiment chapter 4 for the design of the visual behaviour database. Two difficulty conditions incorporated a supplementary visual monitoring task, with different time pressure, to make cockpit monitoring more complex by increasing visuomotor activity. We analyzed the effect of the pilots' profile (pilot vs. novice) as well as the effects of the landing difficulty on numerous standard (number of dwells, average dwell times) and advanced eye movements metrics (LZC, transition entropy, dispersion of attention). Our main hypothesizes were that expert pilots should exhibit different visual behaviors than novices, including more numerous dwells and shorter dwell times, in accordance with the idea that superior perceptual encoding processing comes with expertise. To the best of our knowledge, there is currently no paper dealing with the LZC applied to eye-tracking data. We propose to use Lempel-Ziv Complexity algorithm on eye-tracking data. The complexity (i.e., the quantity and diversity) of visual scanning patterns can be assessed using Lempel-Ziv Complexity (LZC). We expected also a sensitivity of all metrics to expertise, with more visual scanning complexity (as evaluated by the Lempel Ziv complexity and entropy analysis), and a more regular visual scanning between experts (as evaluated by the n-gram analysis). We also assumed that the pilots' expertise could be classified using machine learning, in particular using the transition matrices feature that describe the way pilots switch from an instrument to another. We assume that n-gram similarity matched more between expert group that between novice group, suggesting a higher number of common patterns built by expertise. Finally, we hypothesized that the addition of a parallel monitoring task should also have an impact on ocular behavior, notably by increasing complexity, reducing the regularity level, and generating an ambient mode of attention (i.e. more diffuse attention).

6.6 Materials and methods

For reproducibility purpose, the protocol is available on protocols.io; DOI number: dx.doi.org/10.17504/protocols.io.zb5f2q6.

6.6.1 Participants

Thirty-two participants, all males, participated in this experiment. Data from the 16 pilots who participated in the experiment in Chapter 4 have been included in this study. They all had normal or corrected to normal vision. They were not informed about the exact purpose of the study. They were divided into two groups according to their flying experience. A first group called "novices" consisted of participants with no flight experience (n = 16, mean age = 25.65, SD = 5.47 years). They were recruited from the French aerospace engineering school (ISAE-SUPAERO, Toulouse, France) and all had education in aeronautics. A second group called "pilots" consisted of active professional airline pilots (n = 16, mean age 34.39 , SD = 8.86 years) with a minimum of 1,600 flight hours (mean = 4,321.73, SD = 2,911.41 hours). They were recruited from various airline companies. They all flew on A320 and were currently flying on A320 (68.75 %) or B737 (31.25 %) at the time of the experiment.

6.6.2 Ethics statement

This research project was approved by the local institutional Research Ethics Committee of the University of Toulouse (Comité d'Ethique de la Recherche de l'Université de Toulouse, code N° 2019-131) and was conducted in accordance with the Helsinki Declaration. Volunteers signed an informed consent prior to the experiment and were informed of their right to stop their participation at any time.

6.6.3 Materials

6.6.3.1 Flight simulator

We used an A320-like flight simulator ("PEGASE") located at ISAE-SUPAERO (Toulouse, France), see Figure X. Like in the A320 aircraft, flight instruments included a Primary Flight Display (PFD), a Navigation Display (ND), an Electronic Central Aircraft Monitoring display (ECAM), and an FCU (Flight Control Unit). The participants controlled the aircraft with a side-stick, two thrust levers, and a rudder. We recorded flight data to calculate flight performance during the landing.

6.6.3.2 Flight Scenarios

Participants manually (i.e., without the autopilot) performed three times the same landing scenario according to three different conditions. The "control scenario" was a nominal landing without a supplementary task. The "easy dual task scenario" and the "difficult dual task scenario" were similar to the "control scenario" except that participants were asked to perform a supplementary monitoring task. The purpose of this supplementary task was to increase the level of visuo-attentional effort: participants had to regularly check the ND Zone in the ND screen to say aloud the value at the right time. In the "easy dual task scenario", participants were asked to say aloud the distance between the aircraft and the airfield threshold every 0.5 Nm (information provided by a radio beacon localized near the airfield and displayed in the ND Zone. In the "difficult dual task scenario", they were asked to say aloud this distance every 0.2 Nm. The experimenter stayed in the cockpit during the entire experimentation. Each of the three-landing scenarios consisted of performing a landing to Toulouse-Blagnac Airport, Runway LFBO 14R. The flight began at coordinates 1.2159° of longitude and 43.7626° of latitude. During each scenario, the participants had to comply with the same specific instructions related to the flight. In particular: to maintain a vertical speed between +500 ft/min and -800 ft/min, a speed of 130 knots, and a heading of 143° (corresponding to the Runway 14R). Each landing scenario started at an altitude of 2000 ft and lasted approximately four minutes. The three scenarios were randomized across participants to avoid learning effects. Performance dependent variables were heading, vertical speed, and speed deviations. The number of omissions (i.e., the participant omitted to call out the distance) during the supplementary task was also calculated.

6.6.3.3 Eye movements recordings

Eye movements were recorded at 60Hz using a Smart Eye remote eye tracker (Smart Eye AB, Sweden). The system detects human face/head movements, eye movements, and gaze direction. Gaze direction and eyelid positions are determined by combining image edge information with 3-D models of the eye and eyelids. The system uses five cameras integrated into the cockpit. A major advantage of using several cameras is that eye and head tracking can be maintained despite significant head motions (translation and rotation) or occlusion of one of the cameras by the participant (e.g., by its hand).

6.6.3.4 World model and Area Of Interest

The cockpit was split into 10 AOIs, corresponding to the different flight instruments and displays that pilots can examine during a flight, see figure 6.2.

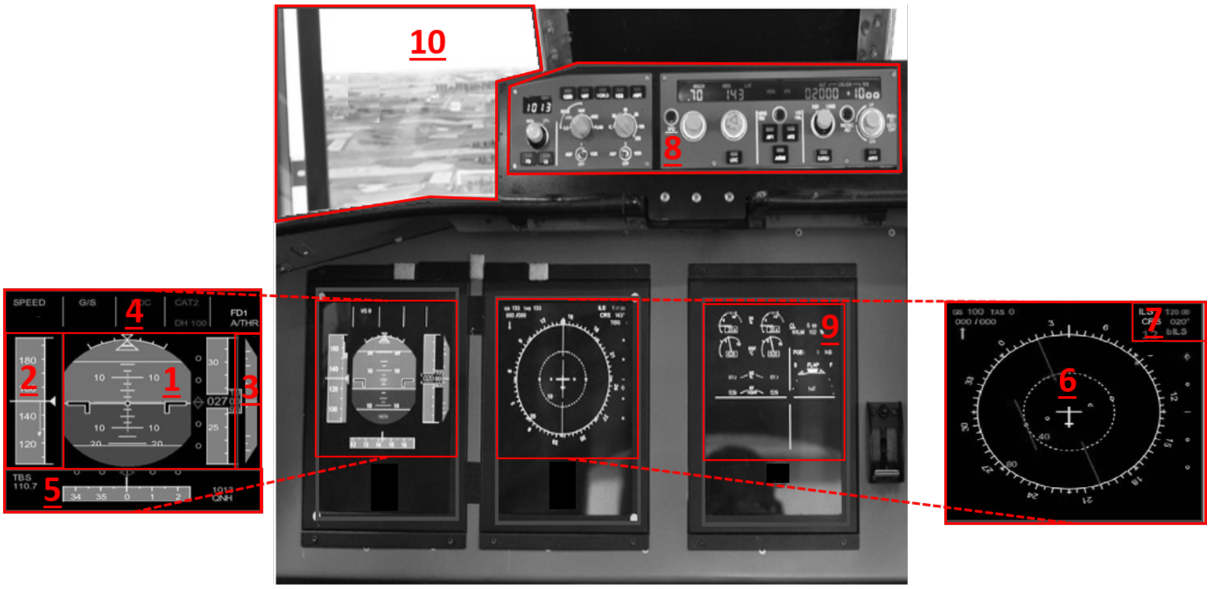


Figure 6.2: Overview of the ten different AOIs: (1) Attitude indicator, (2) Speed tape, (3) Vertical speed tape, (4) Flight mode annunciator, (5) Heading tape, (6) Navigation display, (7) ND zone (displays the distance to recall during the two landing scenarios with the supplementary task), (8) Flight control unit, (9) Electronic centralized aircraft monitoring, (10) Out of the window.

6.6.4 Procedure

At first, participants filled out the consent form and provided demographic information such as their flight qualification (aircraft type) and their flight experience (total hours of flight experience). Participants were briefed on the study and instructed about the different flight scenarios. Then, they were invited to seat in the flight deck at the captain position (left seat). The eye-tracking system was calibrated using an 11-point calibration. Following the Smart Eye manual recommendation, the 11 points were located in the vicinity of the AOIs. Participants performed a training session consisting of performing two times a landing scenario control scenario. Then, the pilots performed three times the same landing scenario than during the training, but with varying levels of complexity.

6.6.5 Data processing

Flight simulator and eye-tracking data were analyzed using MATLAB R2019b with custom homebuilt scripts. The data were recorded from the beginning of the landing scenario to touchdown. Because the landing duration depends on the pilot's actions, landing durations could differ by a few seconds. As a consequence, the beginning of the scenarios has been cut out to obtain the same duration for each participant, corresponding to 14,000 frames sampled at 60 Hz for the eye-tracking data and 233 frames at 1 Hz for the flight simulator.

6.6.5.1 Eye tracking data

Figure 6.3 shows the entire eye tracking pipeline analysis. Raw data were extracted from the eye tracker software, and only AOI-based data were used in this experiment. Each AOI was coded using numbers from 1 to 10. Data related to the dwells on the AOIs were extracted and concatenated to obtain two chronological vectors containing the indices of the visited AOIs (from 1 to 10) and the time spent on them. We discarded dwell times on AOIs inferior to 200 ms (Goldberg et al., 2002). Furthermore, consecutive fixations in the same area were merged (e.g., for 1, 1, 4, 4, 5, 5, 5, 6 we only consider 1, 4, 5, 6). From this data, the transition matrices were computed. We then extracted the transition vector (the vector containing the transitions between each AOI numbers) and used it to compute Lempel Ziv Complexity, entropy, and pattern identification. Transition matrices were concatenated into a single feature space for each participant. Given the high dimensionality of transition matrices, it is difficult to use classical inferential statistics. Therefore, we applied machine learning models on the concatenated transition matrices to compare the two groups of participants (novice vs pilot). The transition probabilities from one AOI to another were taken as a feature, thus raising the number of features to a total of 100 features (i.e., 10 AOIs $\times$ 10 AOIs). A principal component analysis (PCA) was used to reduce the features' numbers. This allowed us to restrict the model to only 35 features corresponding to the main transition probabilities of the matrices to optimize the machine learning models. Five-fold cross-validation was used, which is a good trade-off between bias and variance estimation (Friedman, Hastie, and Tibshirani, 2001). According to Combrisson and Jerbi, 2015 theoretical chance level for classification for $p < 0.05$ with two classes is around 58%. Concerning the LZC algorithm, Figure 6.4 illustrated its functioning. First, the AOI transcription consisted of assigning a unique letter (from A to K) to each AOI label. Then, the initial time series containing AOI label is substituted in time series with letter assigned. After this pre-processing step, the transition vector (or dwell sequence vector) containing time series with letters was extracted and the parsing step could be performed. Regarding the specificity of this algorithm, it identifies the new elements or series

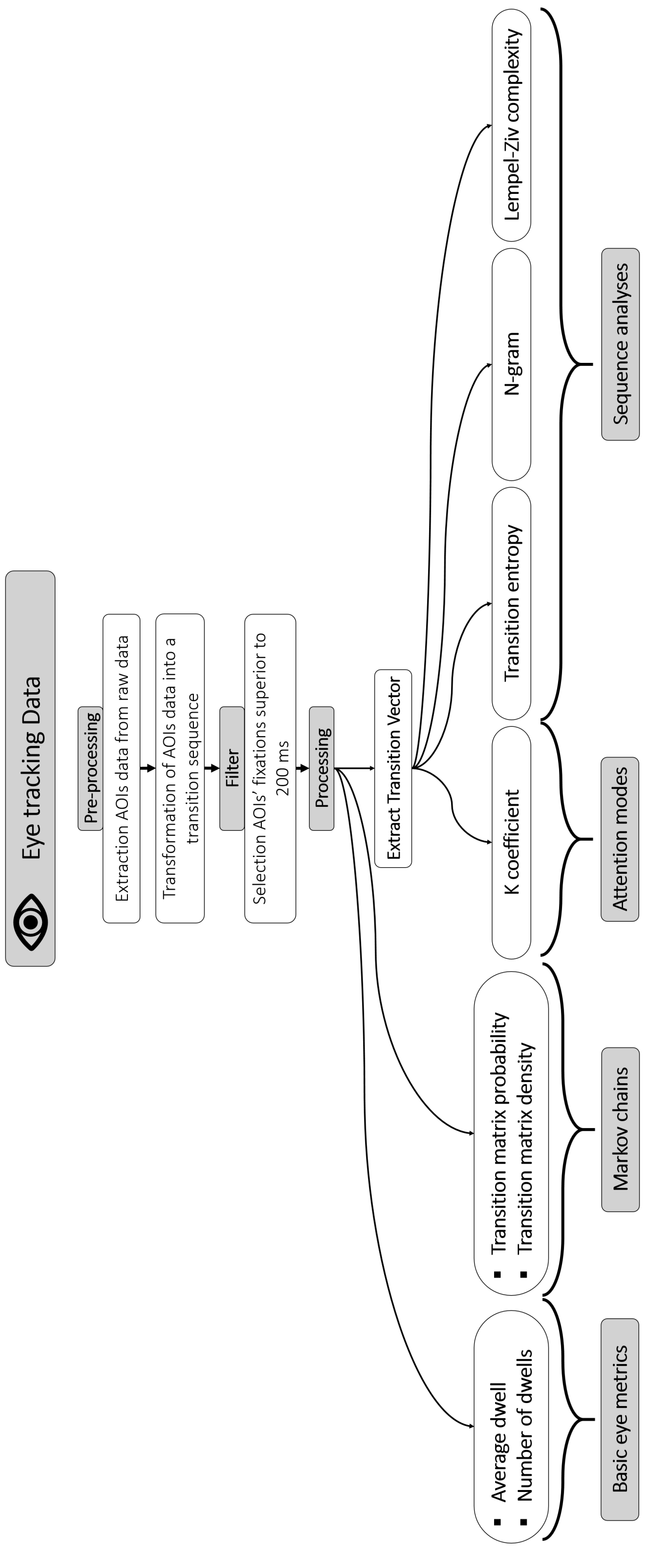


Figure 6.3: Analysis pipeline for eye-tracking data.

of elements in the transition vector by splitting the letter sequences. Then, it stores them in a code book those was not yet encountered. The code book provides an insight concerning patterns and pattern sequences coming from time series. Additionally, it contains index of each pattern encountered with the associated patterns, also a dictionary part which rely on index of basic patterns in case of elaborate patterns. Finally, the Complexity of Lempel-Ziv (LZC) is a numerical value computed by considering the number of different patterns in the code book. In this chapter, the term "complexity" refers to the numerical value given by the Lempel-Ziv Complexity.

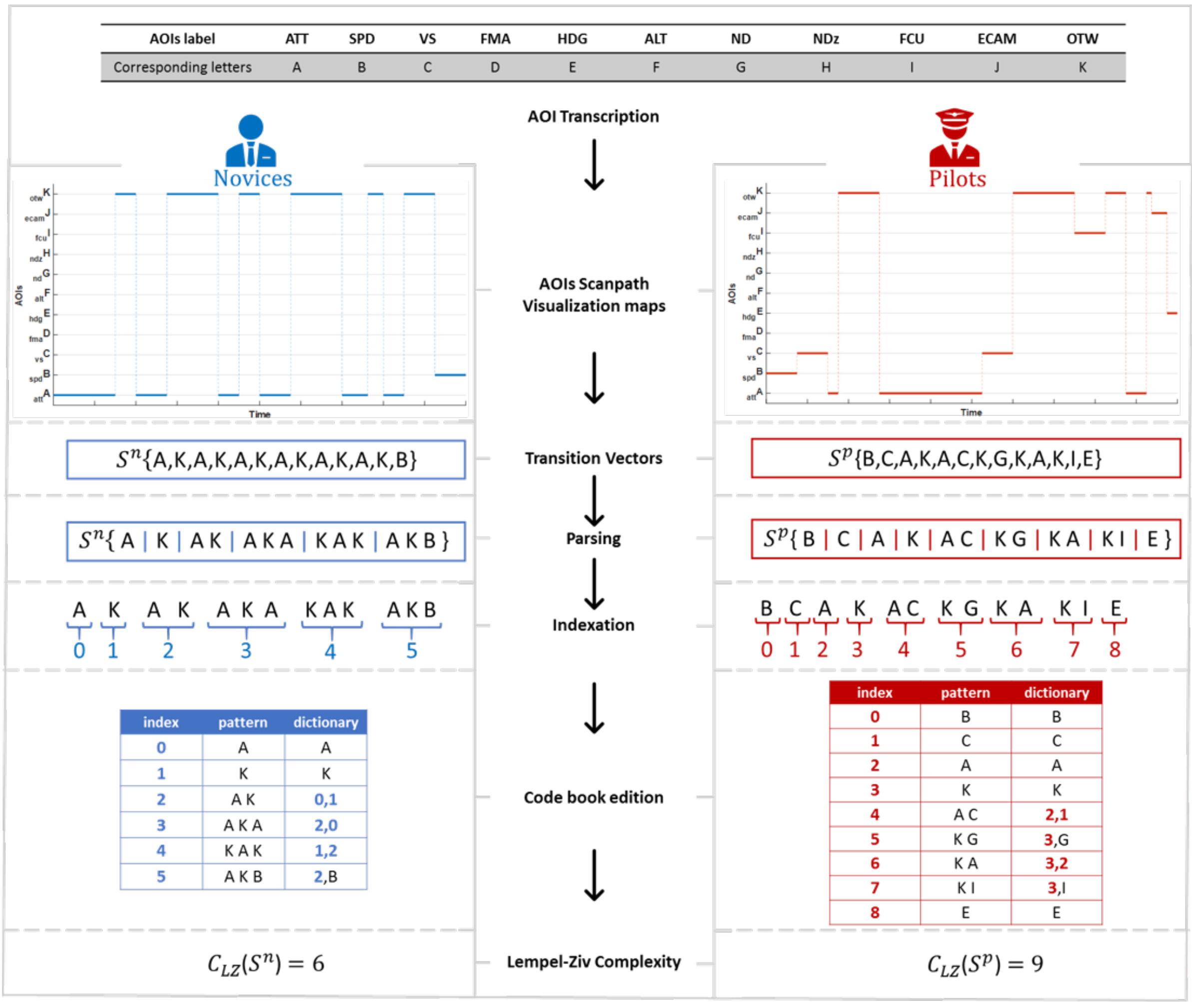


Figure 6.4: Overview of Lempel-Ziv Complexity pipeline applied to eye tracking data of novices and pilots

Concerning the K coefficient, and the transition entropy they were respectively computed following the methods of chapter 5, and Krejtz et al., 2015 respectively. The n-grams frequency-based method was used (Reani, Peek, and Jay, 2018) to identify the number of common 3, 4, 5, and 6-gram sequences in each group. For example, considering an AOI based time series of 1,5,6,4, the possible 2-gram are 15|56|64, and the possible trigrams are |156|564|. After counting the occurrence of given n-grams for each participant, the number of common sequences of each n-gram was calculated for each group.

6.6.5.2 Flight simulator data

The flying performances were examined to quantify the ability of the pilot to comply with the specific flying instructions given by the experimenter. As presented in 6.5, Root Mean Square Errors (RMSEs) were calculated for 3 different flight parameters: speed, vertical speed, and heading. In this experiment, the predicted values corresponded to the different specific threshold given by the experimenter (i.e., speed 130 Kt; vertical speed below -500 ft/min and above +800 ft/min; heading different from 143°) and the observed values corresponded to actual pilots' performances. The deviations were calculated following the formula 6.1:

$$\mathrm{RMSE}_{\mathrm{k,k+1}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2} \tag{6.1}$$

as where for n data points between points k and $k+1$, Pi was the predicted value and Oi the observed value.

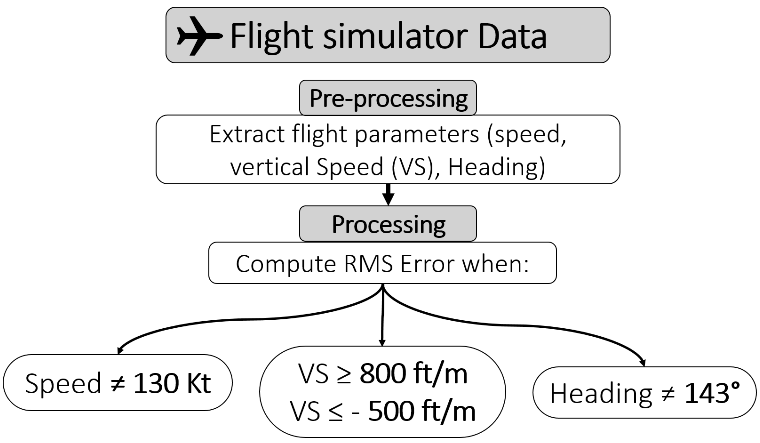


Figure 6.5: Analysis pipeline for the flight parameters.

6.6.6 Statistical analysis

We performed 2 X 3 repeated measures analysis of variance (ANOVA) for each dependent variable (i.e., dual task omission, average dwell time, the total number of dwells, LZC, transition entropy, K coefficient, RMSE heading, RMSE vertical speed, RMSE speed) to assess the effects of the group (novices, pilots) with scenario difficulty as the within-subjects factors (three levels: Control scenario, Easy dual task scenario, Difficult dual task scenario). The normal distribution for each dependent variable was also checked. We used the Greenhouse-Geisser and Huynh-Feldt adjustment to correct the violation of the sphericity assumption when needed. Bonferroni post-hoc tests were performed for multiple comparisons and reported Bonferroni post-hoc are only those with significant differences. The level of significance was set to $\alpha = 0.05$ and partial η^2 was used to estimate the effect sizes.

6.7 Results

6.7.1 Flight performances

The flight performances are shown in 6.6

6.7.1.1 Heading

There was no significant main effect of the group, F(1, 30) = 0.03, $p = 0.874$, nor main effect of the scenario, F(2, 60) = 0.9, $p = 0.39$, on heading deviations. The scenario X group interaction was not significant, F(2, 60) = 0.4, $p = 0.67$.

6.7.1.2 Speed

A significant main effect of the group on speed deviation was found, F(1, 30) = 4.3, $p < 0.05$, $\eta^2 = 0.13$, with the novice's group (M = 5.46; SD = 1.94) showing higher speed deviation than pilot's group (M = 2.66; SD = 1.97). Analyses also revealed a significant main effect of the scenario, F(2, 60) = 3.6, $p < 0.05$, $\eta^2 = 0.11$. Bonferroni post-hoc test showed that

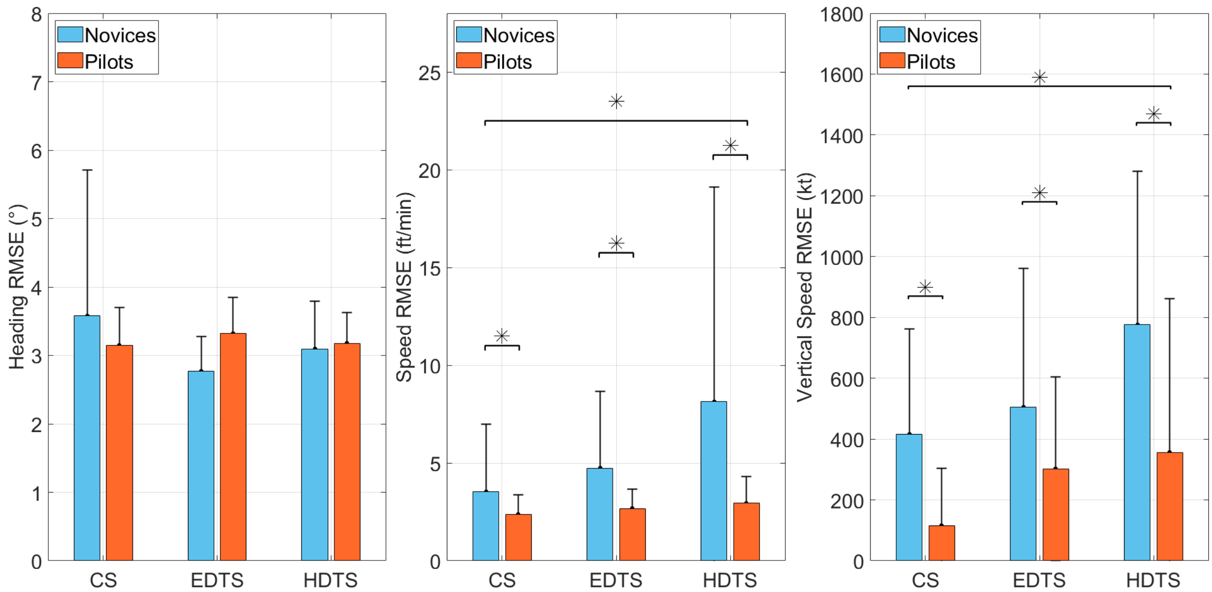


Figure 6.6: Flight performances for heading, vertical speed, and speed deviations among Novices and Pilots groups (error bars represent SD and * indicates main effects $p < 0.05$).

speed deviation was lower during the control scenario (M = 2.95; SD = 0.93) compared to the easy dual task scenario (M = 3.70; SD = 1.02) and the difficult dual task scenario (M = 5.53; SD = 2.79). There was a significant effect of scenario X group interaction, F(2, 60) = 3.3, $p < 0.05$, $\eta^2 = 0.09$. Bonferroni post-hoc test showed that the speed deviation was lower for the pilot's group in the difficult dual task scenario (M = 2.93; SD = 3.97) compared to the novice's group in the difficult dual task scenario (M = 8.13; SD = 4.02).

6.7.1.3 Vertical speed

Analyses revealed a significant main effect of the group, F(1, 30) = 11.4, $p < 0.05$, $\eta^2 = 0.28$, on vertical speed deviation, with the novice's group (M = 565; SD = 130) showing higher vertical speed deviation than pilot's group (M = 258; SD = 134). Analyses also revealed a significant main effect of the scenario, F(2, 60) = 5.1, $p < 0.01$, $\eta^2 = 0.15$. Bonferroni post-hoc test showed that the vertical speed deviation was lower during the control scenario (M = 265; SD = 103) compared to the easy dual task scenario (M = 403; SD = 141) and the difficult dual task scenario (M = 566; SD = 184). The scenario X group interaction was not significant, F(2, 60) = 0.7, $p = 0.52$, $\eta^2 = 0.02$.

6.7.2 Dual task omissions

Analyses showed (6.7) a significant main effect of the group on omissions, F(1, 30) = 35.3, $p < 0.05$, $\eta^2 = 0.54$. The novice's group had a higher number of omissions (M = 2.75; SD = 1) than the pilot's group (M = 0.68; SD = 0.5). Analyses also revealed a significant main effect of the scenario, F(1, 30) = 24.8, $p < 0.05$, $\eta^2 = 0.45$. Bonferroni post-hoc test showed that the difficult dual task scenario (M = 2.37; SD = 0.52) yielded more omissions than the easy dual task scenario (M = 1.06; SD = 0.3). The scenario X group interaction was significant, F(1, 30) = 16.2, $p < 0.05$, $\eta^2 = 0.35$. Bonferroni post-hoc test showed that there were more omissions during the difficult dual task scenario (M = 1.5; SD = 1) vs. easy dual task scenario (M = 3.95; SD = 2) in novices whereas the number of errors did not differ among the two scenarios for pilots.

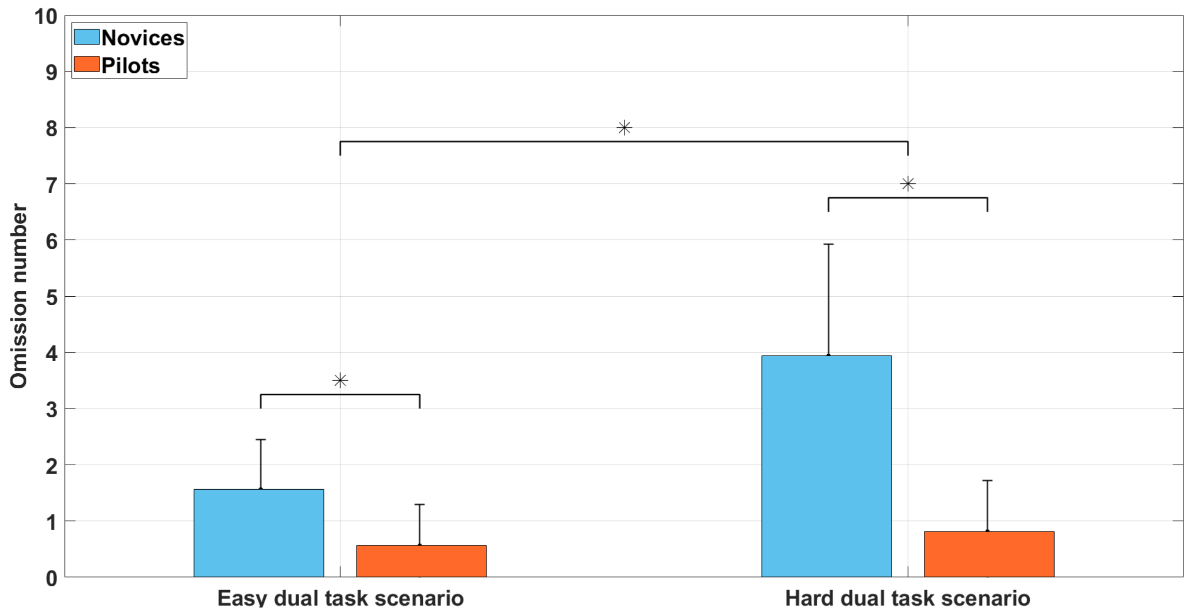


Figure 6.7: Omission number for the easy dual task scenario and hard dual task scenario among Novices and Pilots groups (error bars represent SD and * indicates main effects $p < 0.05$).

6.7.3 Basic eye metrics

6.7.3.1 Average dwell times

Analyses showed (6.8) a significant main effect of group, F(1, 30) = 8.1, $p < 0.05$, $\eta^2 = 0.22$, with short average dwell times for the pilot's group (M = 1.1; SD = 0.2) compared to the novice's group (M = 1.51; SD = 0.21). We also found a significant main effect of the scenario, F(2, 60) = 19.0, $p < 0.05$, $\eta^2 = 0.39$. Bonferroni post-hoc showed that the average dwell time was shorter during easy dual task (M = 1.16; SD = 0.12) and difficult dual task scenario (M = 1.16; SD = 0.17) than during the control scenario (M = 1.58; SD = 0.22). There was no significant scenario X group interaction, F(2, 60) = 2.3, $p = 0.11$, $\eta^2 = 0.07$.

6.7.3.2 Number of dwells

Analyses showed (6.8) a significant main effect of group, F(1, 30) = 13.3, $p < 0.05$, $\eta^2 = 0.31$, with a higher number of dwells for the pilot's group (M = 188; SD = 21) compared to the novice's group (M = 137.5; SD = 19.9). Analyses also revealed a significant main effect of the scenario, F(2, 60) = 13.2, $p < 0.05$, $\eta^2 = 0.31$. Bonferroni post-hoc showed that the number of dwells was higher during easy dual task scenario (M = 172; SD = 16) and during the difficult dual task scenario (M = 177; SD = 18) compared to the control scenario (M = 137; SD = 17). There was no significant scenario X group interaction, F(2, 60) = 0.7, $p = 0.50$, $\eta^2 = 0.02$.

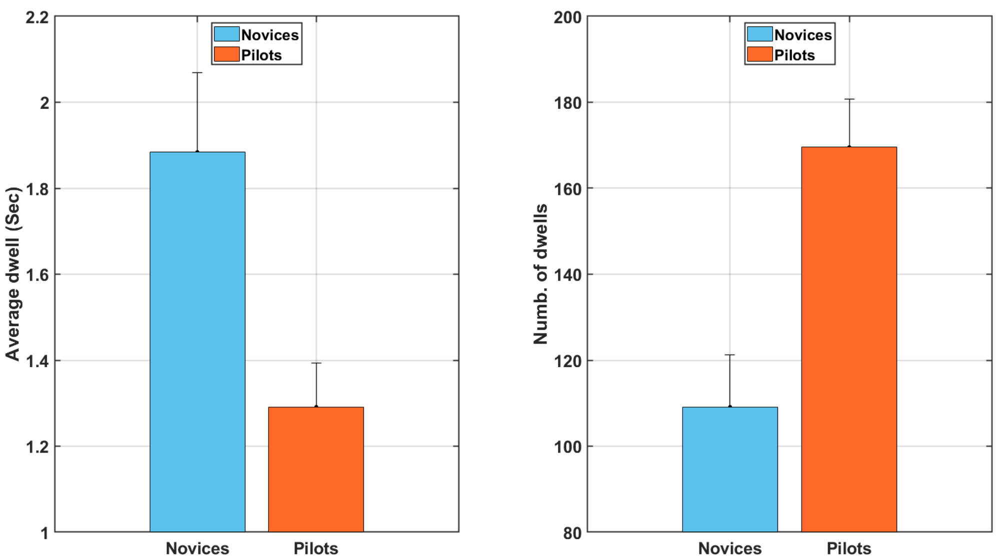


Figure 6.8: From left to right, respectively the Average Dwell and the Number of Dwells averaged over all scenarios among Novice and Pilot groups. (error bars represent SD and * indicates main effects $p < 0.05$).

6.7.4 Advanced eye metrics

6.7.4.1 Transition matrices

The confusion matrix presented in 6.9 show that approach based on Cosine KNN reached classification accuracy up to 91.7% to classify expertise based on transition matrices during the baseline scenario. As shown in 6.10, the differences in transition matrices between novice/pilots are mainly observed in a more homogeneous distribution of transition probabilities from one instrument to another for the pilots.

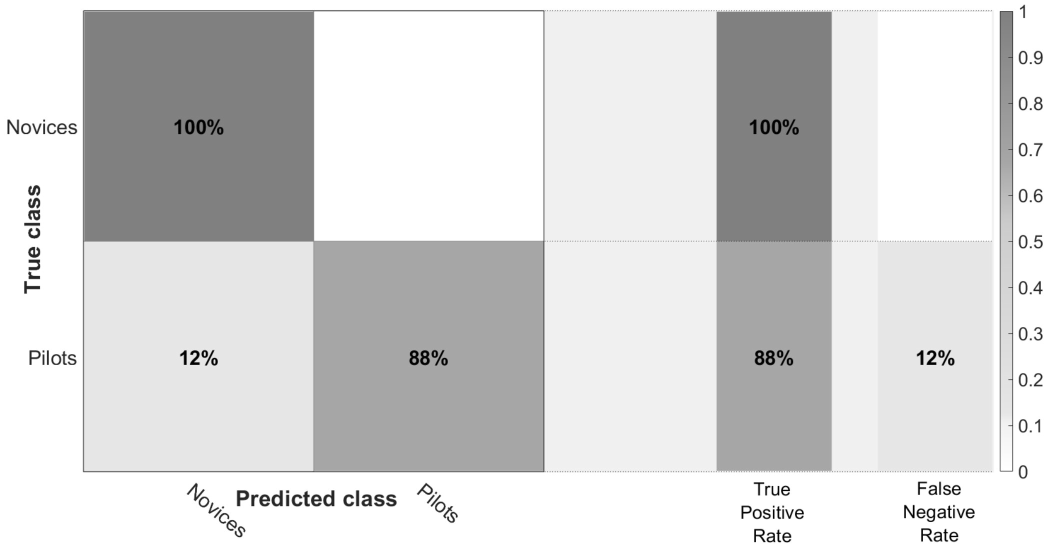


Figure 6.9: Confusion matrix of fivefold cross-validation using the Cosine K-Nearest Neighbors among Novices and Pilots groups during the baseline scenario.

6.7.4.2 Attention mode: K coefficient

Analyses showed (Figure 6.11) no significant effect of the group, F(1, 30) = 3.3, p = 0.07, η^2 = 0.10, on the K coefficient. However, the main effect of scenario was significant, F(2, 60) = 38.1, $p <$ 0.01, η^2 = 0.56. Bonferroni post-hoc test showed that K coefficient was lower during the easy dual task scenario (M = -0.12; SD = 0.06) and during the difficult dual task scenario (M = -0.01; SD = 0.12) compared to the control scenario (M = 0.28; SD = 0.10). There was also a significant difference between the easy dual task scenario (M = -0.12; SD = 0.06) and the difficult dual task scenario (M = 0; SD = 0.12). The scenario X group interaction

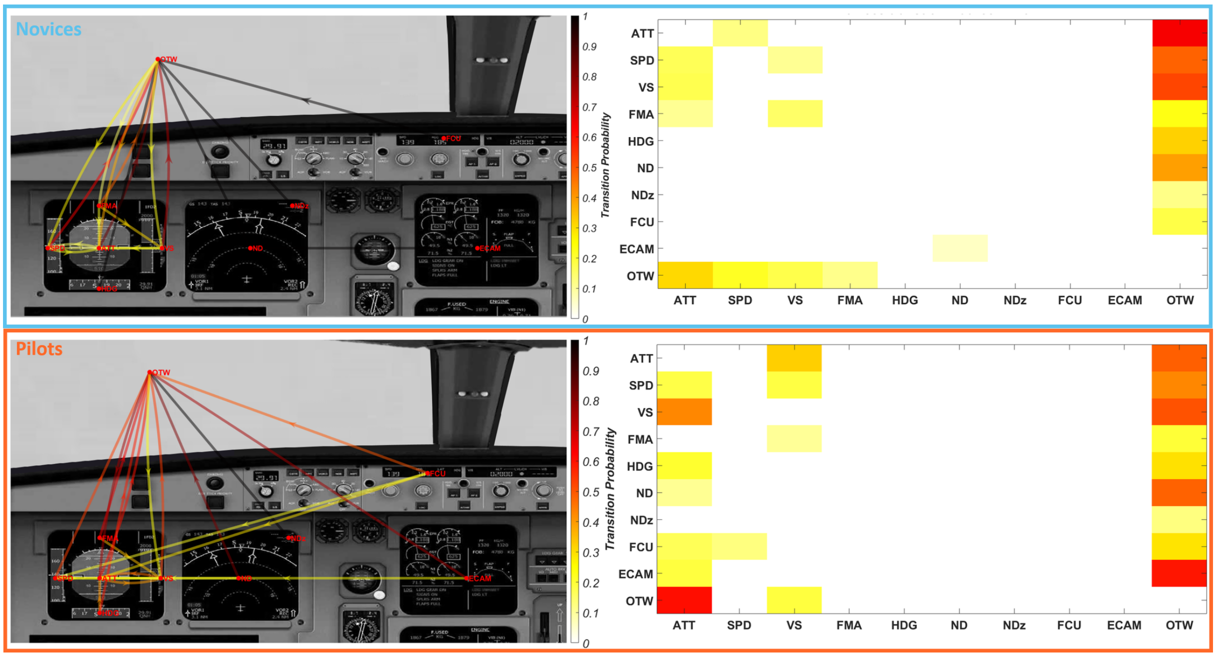


Figure 6.10: Markov chains (Left) and Transition matrices (Right) AOI-based representations among Novices (top) and Pilots groups (bottom) during the baseline scenario.

was significant, F(2, 60) = 4.8, p = 0.01, η^2 = 0.15. Bonferroni post-hoc test showed that K coefficient was lower for the pilot's group in the control scenario (M = 0.14; SD = 0.16) compared to the novice's group in the control scenario (M = 0.41; SD = 0.16). Bonferroni post-hoc test also showed that K coefficient was lower for the pilot's group in the difficult dual task scenario (M = -0.10; SD = 0.17) compared to the novice's group in the difficult dual task scenario (M = 0.09; SD = 0.16).

6.7.4.3 AOI sequence analysis

Transition entropy

Analyses showed (6.12) a significant main effect of group, F(1, 30) = 6.0, $p < 0.05$, $\eta^2 = 0.17$, with the novice's group (M = 1.22; SD = 0.2) showing lower transition entropy than pilot's group (M = 1.56; SD = 0.2). Analyses also revealed a significant main effect of the scenario, F(2, 60) = 8.4, $p < 0.05$, $\eta^2 = 0.22$. Bonferroni post-hoc test showed that the transition entropy was higher during easy dual task scenario (M = 1.50; SD = 0.16) and during difficult dual task scenario (M = 1.44; SD = 0.17) than during the control scenario (M = 1.23; SD = 0.15). The scenario x group interaction term was not significant, F(2, 60) = 0.2, p = 0.82, η^2 = 0.01.

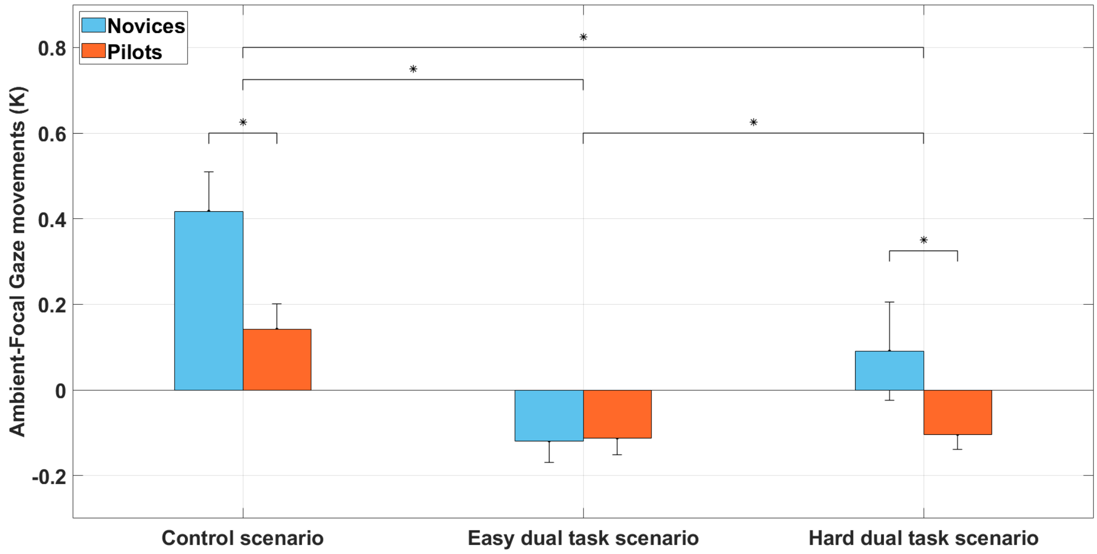


Figure 6.11: Ambient Focal K coefficient during the Control scenario, the Easy dual task scenario, and Hard dual task scenario among Novices and Pilots groups. K $>$ 0 indicates a focal visual attention, whereas K $<$ 0 indicates an ambient visual attention. (error bars represent SD and * indicates main effects $p < 0.05$).

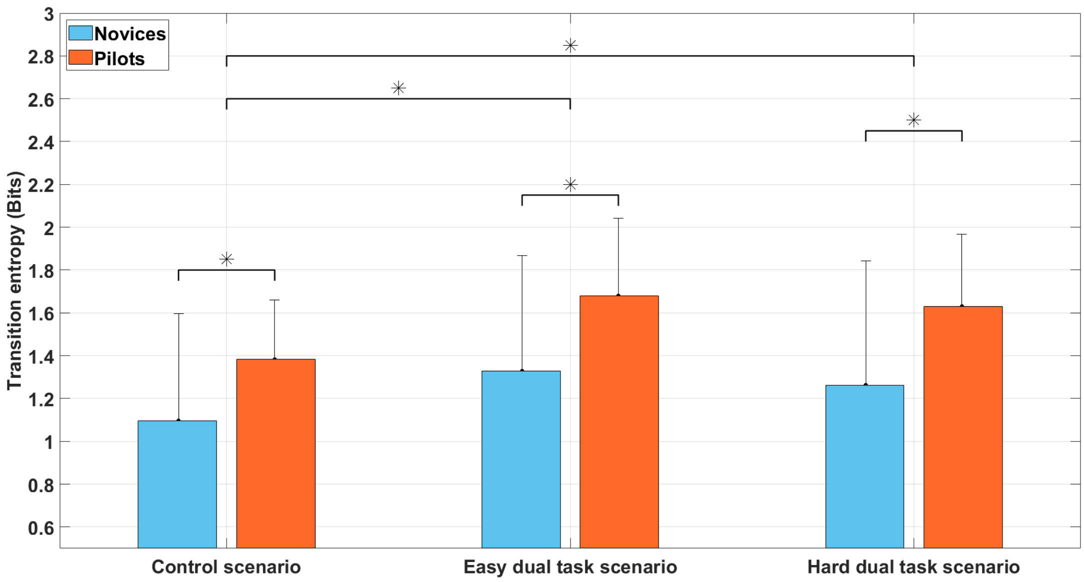


Figure 6.12: Transition Entropy during the Control scenario, the Easy dual task scenario, and the Hard-dual task scenario among Novices and Pilots groups. (error bars represent SD and * indicates main effects $p < 0.05$).

N-gram analysis

As presented in Figure 6.13, the count of common n-gram sequences revealed that pilots have more common sequences than novices during all scenarios (Control, easy dual task, and hard dual task). The easy dual task and hard dual task scenario yielded to more common sequences for both groups compared to the control scenario. Regardless of the n-gram length (3, 4, 5, or 6), during the control scenario, the pilots had more common sequences than novices. For example, the most frequent tri-gram pattern for the novices was OTW|VS|OTW – transition between out-of-the-window, vertical speed, and back. On average, it was repeated 6.4 times. For the pilots, the most frequent tri-gram occurred 17.4 times on average and it was OTW|ECAM|OTW. We also note that the ten most frequent n-grams included the same AOI at least twice (for instance, repeated transitions between same instruments). For novices, trigrams involving three unique AOIs were OTW|SPD|ATT repeated 2.6 times on average, OTW|VS|ATT – 2.1 times, and OTW|ATT|SPD – 2 times. For pilots, the trigrams involving unique AOIs were OTW|ECAM|ATT repeated 9.4 times on average, OTW|VS|ATT – 8.6 times, OTW|ATT|VS – 4.6 times, and OTW|HDG|ATT – 3.8 times. For the both easy and hard dual task scenarios, the most frequent trigram involved the ND zone display for both groups OTW|NDz|OTW. It occurred 17.6 times on average for novices and 19.1 for pilots during the easy dual-task scenario, and 21.1 times on average for novices and 22.2 for pilots during the hard dual-task scenario. For novices only one frequent trigram with unique AOIs found in the control scenario was also found during the easy dual task scenario (OTW|VS|ATT). However, this trigram was not found during the hard-dual task scenario. As for the pilots, among four trigrams with unique AOIs that were found in the control scenario, only 2 of them were found in the easy dual-task scenario, and only one in the hard-dual task scenario (see 6.1). Interestingly, the most frequent 5-grams among novices was OTW|SPD|OTW|SPD|OTW repeated on average 1.5 times whereas OTW|VS|ATT|OTW|ATT was the most frequent 5-gram among pilots repeated on average 3 times.

Table 6.1: Table representing the most frequent trigrams involving unique AOI in the pilot group during the control scenario (CS), the easy dual-task scenario (EDTS), and the hard dual-task scenario (HDTS).

Frequent trigram involving unique AOI	Average of occurrences in the CS	Average of occurrences in the EDTS	Average of occurrences in the HDTS
OTW|ECAM|ATT	9.4	0	0
OTW|VS|ATT	8.6	7.7	0
OTW|ATT|VS	4.6	5.5	0
OTW|HDG|ATT	3.8	0	0

Lempel Ziv Complexity (LZC)

Analyses showed (6.14) a significant main effect of group, F(1, 30) = 10.0, $p < 0.05$, $\eta^2 = 0.25$, with a higher LZC for the pilot's group (M = 40.3; SD = 5.6) compared to the novice's group (M = 33; SD = 5.2). There was also a significant main effect of the scenario, F(2, 60) = 13.2, $p < 0.05$, $\eta^2 = 0.30$. Bonferroni post-hoc test showed that LZC was higher during easy dual task (M = 40.46; SD = 4.4) and difficult dual task scenario (M = 37.9; SD = 4.97)

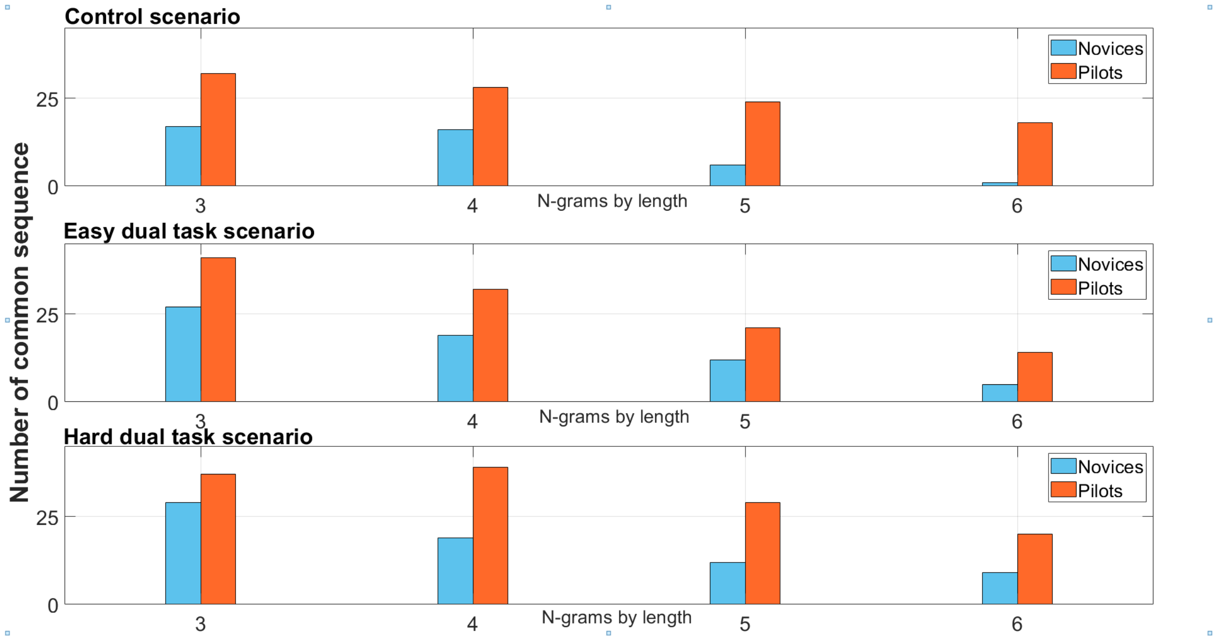


Figure 6.13: Number of common patterns sequence by N-grams length during the control scenario, the Easy dual task scenario, and the Hard dual task scenario among Novices and Pilots groups.

than during the control scenario (M = 31.7; SD = 3.76). The scenario X group interaction was not significant, F(2, 60) = 0.5, p = 0.62, η^2 = 0.02.

6.8 Discussion

Several previous studies have reported differences among pilots and novices in how they scan cockpit instruments using standard metrics such as fixation duration, dwell times, numbers of saccades, etc. Bellenkes et al., 1997; Kasarskis et al., 2001; Svensson et al., 1997; Xiong, Wang, Zhou, Liu, and Zhang, 2016 reported that experts visited more instruments and spent less time dwelling on each instruments compared to novices. Furthermore, one interpretation proposed by this study is that this visual behavior avoid visual "tunneling" and make sure the aircraft is performing as required. One of the limitations of these studies is the analysis of visual strategies across dwells and the number of dwells. Indeed, even if they reflect visual scanning strategies, they do not allow a comparison of the visual scanning strategies deployed and the occurrence of patterns. In this work, various eye metrics were analyzed in sixteen novices and sixteen professional pilots during landing scenarios involving different visuo-attentional effort. All the metrics used in this study allowed characterizing visual scanning in terms of gaze dispersion

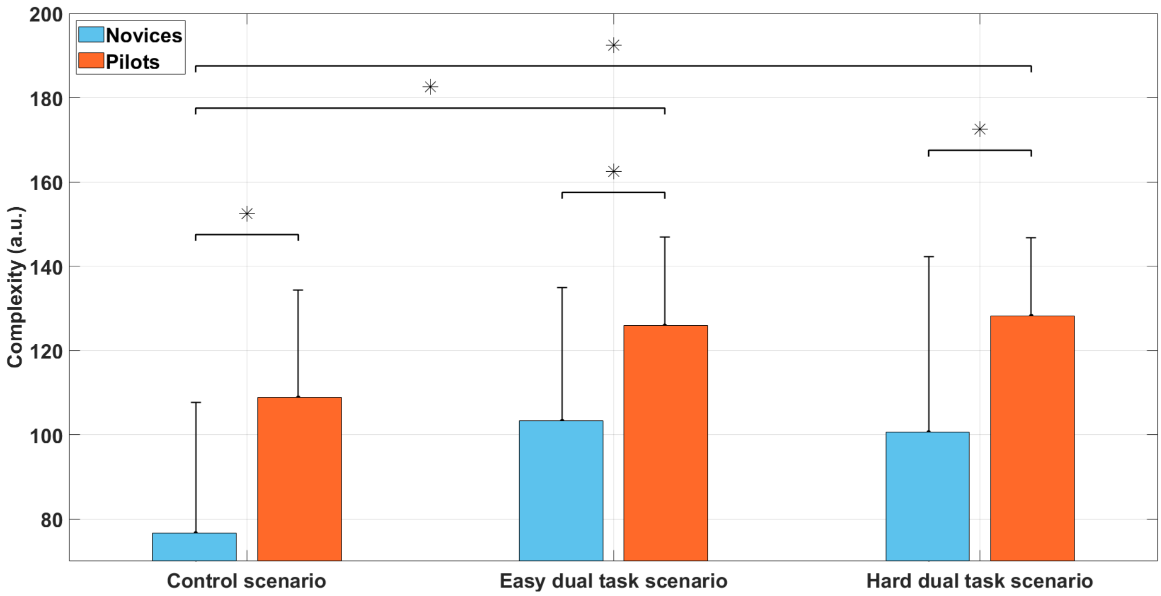


Figure 6.14: Lempel-Ziv Complexity during the Control scenario, the Easy dual task scenario, and Hard dual task scenario among Novices and Pilots groups.

and gaze patterning. We examined the impact of expertise and flying difficulty on the visual scanning strategies. As showed by our results, a large number of standard and advanced metrics was sensitive to these two factors.

6.8.1 Basic eye metrics

Our results show that experts' pilots had shorter average dwell times and a higher number of dwells compared to novices. This result has been interpreted in the literature (Charness, Reingold, Pomplun, and Stampe, 2001; Curby, Glazek, and Gauthier, 2009; Gegenfurtner et al., 2011) as an important sign of expertise, built on an optimization of the visual information processing, allowing faster extraction of information when consulting a flight instrument. This strategy allows consulting more often the various instruments, resulting in a better updating of situational awareness (C.-s. Yu et al., 2014). This outcome also supports the existence of a superior perceptual encoding of domain-related patterns (Goldberg et al., 2002; Shapiro and Raymond, 1989) in expert pilots.

6.8.2 Markov chains / Attentional modes

Based on transition matrices, a machine learning approach using Cosine KNN algorithm reached an accuracy of 93% to classify expertise. Expert pilots have more distributed transition probabilities when switching from an instrument to another: their visual patterns included more instruments. This suggest that experts include more flight instruments in their visual scanning and succeed to balance their time between them. The focal-ambient K coefficient showed that attention was dominantly focal (positive value) in both groups. However, the attention was more focal in the novice's group vs the pilot group. It can be assumed that experts' pilots have a greater spatial distribution of their visual attention than novices. The K coefficient also showed sensitivity to the task difficulty. By adding a monitoring task (Easy dual-task scenario) inducing a supplementary display to monitor, visual attention switched from focal to ambient for the 2 groups. Interestingly, by further increasing the time pressure of the monitoring task (hard dual-task scenario), we found that the induced dual-task changed the ambient-focal strategy of the novices by turning it into focal mode, while the pilot group kept their strategy consistent across the experimental dual task-scenarios (e.g., ambient mode).

6.8.3 Sequence analyses

As showed by the transition entropy analysis, more information (bits) was required to describe expert pilots' visual strategies than the novice group. Thus, the pilot group exhibited more complex visual scanning patterns. Within the professional pilot group, the n-grams analysis of common sequences highlighted the existence of a higher intra-group similarity probably built with expertise as well as more elaborate visual strategies considering common visual scanning patterns of size 6 (6-grams). Furthermore, this analysis revealed that some complex patterns (that include only distinct flight instruments) found in the control scenario were still present in both easy and hard-dual task scenarios. We expected that adding a double task would impact the visual scanning. Our results revealed that pilots kept their visual scanning strategies related to the manual landing task by the quite same variety of visual pattern (found in the control scenario) in the dual-task scenarios (easy, and hard). We back these results up with the dual task performances and flight performances where maintaining patterns related to the landing task (control scenario) during dual task scenarios would maintain relevant visual activity for maintaining flight performance and performing callbacks. Finally, AOI redundancies were also found in both groups, i.e. n-grams having twice several same AOI in an n-gram sequence. The complexity of the Lempel-Ziv demonstrated that redundancies was lower in the pilots' group. They displayed a higher complexity and richness of visual patterns, containing a larger variety of possible combinations.

6.8.4 Limitation

There were some limitations in this study. We compared professional pilots with non-pilots. A further researcher should consider participants with different levels of expertise from novice to expert (e.g., every 1000 hours) to finely examine the implementation of the visual strategies with expertise. A future study should also consider a full flight simulator to better fit with the operational context. This experiment could be also replicated with different meteorological conditions, and level of automation. Finally, the eye tracker devices are more and more mature and accurate (about 1° at a distance of one meter). However, the experts may succeed in taking information in peripheral vision allowing for example to maintain a constant speed by looking only at the attitude zone. This would explain why the "AOI SPD" corresponding to speed tape is not often found in the most frequent patterns (n-grams). Only few studies examined the effects of peripheral vision on performance. However, one study compared flight performance of instructor pilots and student pilots with or without the availability of information from peripheral vision (Fox, Merwin, Marsh, McConkie, and Kramer, 1996). This was realized by linking the eye tracker to the instrument display and fading all instruments not related to the gaze. Both instructors and students presented a degraded performance in condition when peripheral vision was inhibited. However, the instructors' performance suffering more than that of the students. Hence, it appears that peripheral vision can be processed by pilots and that the ability to process peripheral information is related to pilots' expertise.

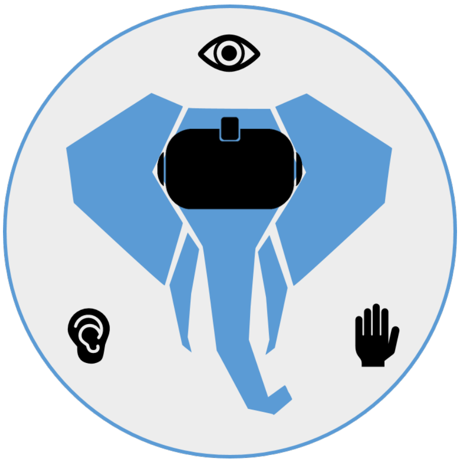
6.9 Conclusion

This study highlighted the differences between novices and expert pilots concerning visual scanning strategies and flight performances. Our result confirmed that expertise exerts a top-down modulation on gaze behaviour (Shapiro and Raymond, 1989). We used a wide variety of standard and advanced metrics to uncover the modification of the gaze behavior bring by expertise. Expert pilots have a more efficient perception of the information, a better dispersion of their attention, and more elaborate visual patterns. Expertise makes it possible, despite a dual-task costly in visuo-attentional resources, to maintain the type of visual patterns linked with the flying task (i.e. the irrelevant dual-task did not alter the nominal visual behavior). Overall, the eye metrics used in this research are relevant to finely assess pilot's gaze behavior in the cockpit and can contribute to better characterize visual scanning in the cockpit, an important topic for safety (Li, Chiu, Kuo, and Wu, 2013). These eye metrics can be used to evaluate pilots during their training program. For example, it might be possible to follow the evolution of their scanning strategies and determine whether they tend to resemble that of expert pilots. In the future, it might be possible to assess cockpit monitoring during real flight (Peysakhovich et al., 2018). In this way, a FETA system investigated the possibility to use an eye tracking assistant to warn pilots when they do not watch sufficiently an instrument (Lounis, Peysakhovich, and Causse, 2018, 2019). Our results suggest that such on-board eye tracking could be customized based on pilot experience. Finally, we believe that the eye metrics employed in this study can be also useful for practitioners and researcher in other fields such as air traffic control and automotive.

CHAPTER 7

GAZE-CONTINGENT MULTIMODAL WARNINGS DURING A PILOTING TASK IN VIRTUAL REALITY

Sommaire



As we have seen in Chapter 4, The FETA system compares the pilot's visual scanning to a database of standard visual behavior based on the average Non-Dwell Time metrics. During piloting, if the current pilot's visual scanning deviates from the database (one or several instruments are not sufficiently gazed), the FETA system triggers an auditory alarm to redirect the pilot's attention toward the relevant flight instruments (by priority order if several instruments are neglected). Our results showed that with the FETA system enabled, pilots looked more closely at flight instruments (e.g., speed, vertical speed, and heading) when the flight scenario was more difficult (high workload) and reported an improved situational awareness. However, in this pre-study, two issues were raised. The first one concerns the unnecessary notifications from FETA. The pilot may have seen a parameter that deviates but does not react because he is involved in a more important task or he does not want to correct the situation right now because he thinks that the deviation is tolerable. One solution to overcome this problem would be to rely both on eye tracking and flight parameter data to trigger the FETA notifications. The flight parameter data correspond to the values of the flight parameters, for example, speed or vertical speed. Thus, the pilot is only alerted if the parameter and its visual behavior both deviate from the database standards. The second issue raised in the pre-study concerns the way in which pilots are alerted when there is a lack of surveillance. Indeed, FETA alerts use the auditory modality. As the latter is already widely used in cockpits, other alert modalities such as vision or touch must be considered to avoid overloading auditory attention. Due to Covid-19 pandemic, this study could not be completed. Nevertheless, this chapter will briefly present the protocol and hypotheses.

7.1 Alertness modalities

The choice of the alert modality (visual, haptic, auditory) is important because there are situations in which the visual and/or auditory channels are highly overloaded, making the processing of visual and/or auditory information more complicated. This is the case for pilots who work in complex environments and who are sometimes at the limit of their visual and auditory processing capacity. Studies have shown that many accidents are due to a lack of response to alarms (Ancel and Shih, 2012; et d'Analyses et al., 2012; Team, 2008. Dehais et al., 2012; Dehais et al., 2014 in 2012 and 2014 showed that when pilots were overload, the perceptual processing of information not relevant to the main task at hand is impaired. For example, unexpected sounds may go unnoticed in an attention-demanding context. This is the phenomenon of inattentional deafness. It is possible that, during critical phases, the visual processing of information may interfere with the simultaneous evaluation of auditory alarms and thus induce inattentional deafness. Therefore, finding the right modality to alert pilots is essential. Multiple resource theory (Wickens and Liu, 1988) proposes that each modality has distinct attentional resources. The use of different modalities would thus make it possible to use different attentional channels, making information processing more efficient while improving task-sharing performance (Barbé et al., 2016; Wickens, 2002, 2008). In addition, introducing the haptic modality in cockpits could be a good solution for alerting pilots (Chang, Hwang, and Ji, 2011; Gaffary and Lécuyer, 2018; Salzer, Oron-Gilad, Ronen, and Parmet, 2011; Van Erp, Groen, Bos, and Van Veen, 2006; Young, Tan, and Gray, 2003). Indeed, at this moment, haptic signals are not used on board cockpits. Interpreting haptic signals would interfere less with the task at hand because they would have the ability to be perceived simultaneously with an auditory or visual signal (Baldwin et al., 2012; Sklar and Sarter, 1999).

7.2 Goal of the research

This project aims to develop and test a countermeasure system to poor system monitoring based on eye-tracking coupled with different warning modalities. To evaluate the feasibility of this principle, we planned to record the visual behavior of participants trained to be experts during a task inspired from the MATB-II (Multi-Attribute Task Battery developed by NASA, see Comstock Jr and Arnegard, 1992 for a review, and see Santiago-Espada, Myer, Latorella, and Comstock Jr, 2011 for a recent description). This task simulates the activities that aircraft crew-member perform in flight. We modified the MATB-II task to run it in Virtual Reality (VR) in order to respect the spatial positions of the flight instruments as in a cockpit. During the task performance, different warnings were implemented to alert the participants in case of abnormal situations, in particular when task parameters deviate too much (e.g., lack of response to a system monitoring task).The first objective was to create, thanks to eye-tracking, a database of the typical visual scanning behavior of experts during MATB-II as in FETA system (see chapter 5). Then, we could extract from this database some relevant metrics such as the average non-dwell times (periods of time during which an individual doesn't look at a parameter), the average dwell times (average time spend looking at a parameter), and the visual scanning patterns (such as entropy score, N-gram sequences, Focal-Ambient score and LZC). The second objective was to test the possibility of using these metrics in real time to improve untrained participants' performance. For that, participants unfamiliar with the MATB-II could be be recruited. this participants would be alerted when their visual scanning deviate too much from the typical visual scanning of the database and/or when the MATB-II flight parameters deviate markedly from that of the experts. Finally, the third objective was be to evaluate which sensory modality is the most effective in alerting participants (auditory, visual, haptic) under various conditions of difficulty (low workload vs high workload).

7.3 General hypotheses

- Firstly, we assumed that as in Shapiro and Raymond, 1989; Underwood, 2007 (see Gegenfurtner et al., 2011 for a review) experiments experts would develop more efficient visual scanning strategies after enough training sessions. In this way, the experts' eye movements can be recorded to serve as a standard database. It could be then possible to extract relevant measurements from this database and integrate them into a virtual assistant, with the aim of alerting participants when their visual scanning deviates too far from that of the experts.

- Secondly, we assumed that alerts based on MATB-II flight parametric data and eye-tracking data could be more effective (reduce the occurrence of deviations without degrading task performance and subjective feelings) than alerts based on MATB-II flight parameters only or eye tracking data only.

- Finally, we assumed that haptic alerts could be more effective than auditory and visual alerts, especially in high workload situations.

7.4 Method

7.4.1 Experiment 1: Train experts at MATB-II

For the first part of the study, 10 participants will be placed in VR environment using an HTC Vive headset with an integrated Tobii eye-tracker. They will have to perform the task inspired from the MATB-II (Multi-Attribute Task Battery II), basically a simplified piloting environment (see figure 7.1). Two conditions with differents worload levels will be proposed. To vary the workload, the number of deviations will be varied of the 6 parameters concerning the monitoring task, the difficulty of the tracking task (e.g. to maintain the tracker in the target at the center) and the flow rate of the pumps for the management task.

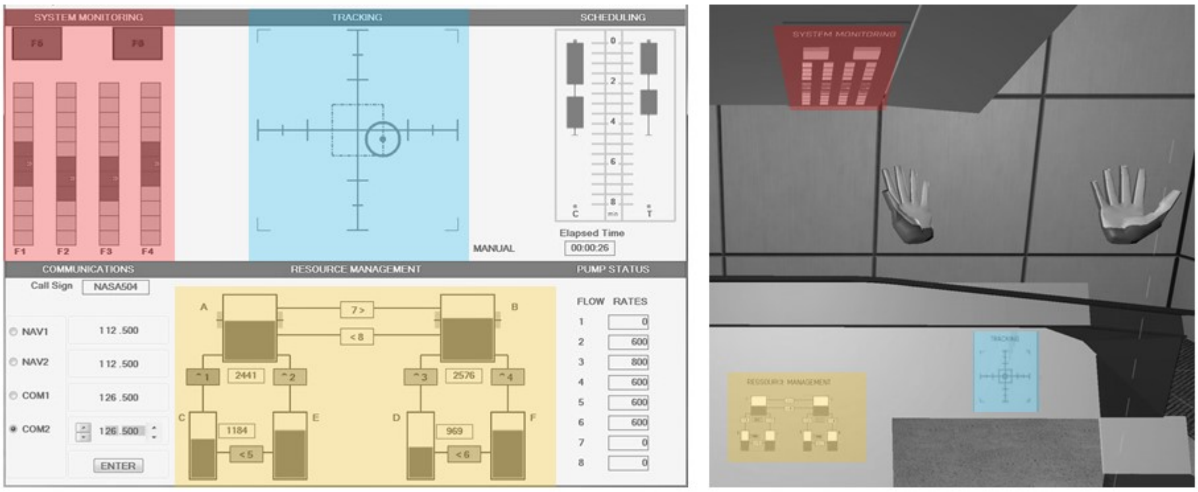


Figure 7.1: Left: The current MATB-II with their independent tasks. Right: The MATB-II In virtual reality containing 3D spatialization of the tasks. Rectangles with the same color represent the same tasks between both MATB.

The MATB-II is composed of a monitoring task of six parameters (see figure 1), a tracking task, where the participants will have to use a joystick to keep a target in the center, and finally, a management task, requiring the participants to manage the flow of several pumps (see appendix for more details). For the first experiment, the participants will be trained to be considered as experts. Based on Shapiro and Raymond, 1989 experiment, participants will be trained until their performance (task performances) remained stationary. Their eye movements will be recorded and analyzed and will allow to build the database of the visual scanning strategies.

7.4.2 Experiment 2: Choosing an input merging visual behavior and/or "Flight" parameters

For the second part, twenty participants are planned to be recruited in order to provide a sufficient amount of data to reach a significant effect. This value was obtained using data from a previous work using eye tracking during a cockpit monitoring task that was pretty analog to the current experiment (see Shapiro and Raymond, 1989). When the MATB-II parameter values deviate (from their nominal value) and/or when the participants' visual scanning deviate from the expert visual standard database, the participants will be alerted with auditory alarms. The tree tasks (monitoring, tracking and management) will be associated with a number 1, 2 or 3. The auditory alarms will be a number from 1 to 3 (e.g., corresponding to the different tasks: "1" corresponding to monitoring task, "2" corresponding to the tracking task, and "3" corresponding to the management task.) played in the headset to redirected the visual attention toward the relevant task. Two workloads condition will be tested, a low and a high workload condition. To vary the workload, the number of deviations will be varied of the 6 parameters concerning the monitoring task, the difficulty of the tracking task (e.g. to maintain the tracker in the target at the center) and the flow rate of the pumps for the management task. At the end of the experiment, the participants will be asked to complete three questionnaires a Workload Index questionnaire (NASA TLX), a Situation Awareness questionnaire (SART/SASHA), and a User experience questionnaire (AttrakDiff) (See all questionnaires in appendix). Participants will be equipped with an electrocardiogram (ECG) FAROS 4 with 5 electrodes sampled at 1000Hz to examine the effects of mental workload variations on heart rate and heart rate variability. Indeed, increase mental effort is known to provoke a shift of the balance of the autonomic nervous system towards a sympathetic dominance, increasing heart rate and reducing total HRV (e.g., Causse, Baracat, Pastor, and Dehais, 2011; Fairclough and Mulder, 2012). Measuring mental effort with ECG will have two important purposes: 1) validating our experimental variations of the mental effort via the number of deviations of the 6 parameters, and more importantly 2) investigating whether using participants' visual scanning as a trigger for alert will allow an overall reduction of the mental effort during the task in comparison to alerts triggered on the only basis of performance deviation.

7.4.3 Experiment 3: Testing countermeasure modalities to warn the pilots

The last part is similar to the second, 20 participants will be recruited. The warning system selected in the experiment 2 (e.g., warning with: MATB parameter deviation only, visual scanning behavior deviation only, or both MATB parameter deviation and visual scanning behavior deviation) will be used in this experiment. However, in this experiment, different sensory modalities will be tested to alert participants (e.g., auditory, visual or haptic) for any deviations. As in experiment 2, the monitoring task is associated with "1", the tracking task with "2" and the management task with "3". The auditory alarms are therefore "1", "2" and "3". Visual alarms are also "1", "2" and "3" and are displayed in the Graphical User Interface (GUI) in the HTC Vive headset. For the haptic alarm, the signal is a 250Hz signal with a duration of 800 ms. Studies have shown that this is the frequency and the duration at which sensitivity is highest (L. A. Jones and Sarter, 2008; Veitengruber, 1978; Verrillo, 1966; Wilska, 1954). The monitoring task is associated with a vibration of 800 ms, the tracking task with two vibrations of 350 ms and an inter-stimuli interval (ISI) of 100 ms and the management task with three vibrations of 200 ms with an ISI of 100 ms (this makes it possible to have the same stimulus duration). Concerning the location of the haptic alert, several studies have shown the effectiveness of the haptic modality on the thigh (Salzer et al., 2011), the torso (Cholewiak, Brill, and Schwab, 2004; Erp, Veen, Jansen, and Dobbins, 2005) or the arm (Sklar and Sarter, 1999 for a review see Myles and Binseel, 2007. The haptic device would therefore be placed either on the chest of the participant, as shown in figure 7.2. Between each participant, the order of presentation of the conditions (auditory, visual, haptic) is counterbalanced to avoid an order effect. As is experiment 2, two workloads conditions will be tested a low and a high workload. The participants will be asked to complete three questionnaires a Workload Index questionnaire (NASA TLX), a Situation Awareness questionnaire (SART/SASHA), and a User experience questionnaire (AttrakDiff) (see Appendix for the 3 questionnaires). Participants will be also equipped with an electrocardiogram (ECG) FAROS. As in experiment 2, it will allow validating our experimental variations of the mental effort via the number of deviations of the parameters.

Figure 7.2: Haptic Suit

7.5 Perspectives

The aim of this study was, in a first step, to develop a countermeasure system considering the eye recordings of ten experts at MATB-II. From these recordings, we wanted to extract relevant measurements, such as the period of time during which one of the parameters is not monitored (e.g., Average Non-Dwell Time) or visual scanning strategies (see chapter 6). These measures would have made it possible to build a database of optimal visual scanning strategies. We then wanted to assess the relevance of this database by alerting novice participants when their visual pathways deviated from the database standards and/or MATB-II parameters became abnormal. Finally, we also aimed at testing the effectiveness of different alerting modalities and their impact on performance, perceived workload and situational awareness (see figure 7.3).

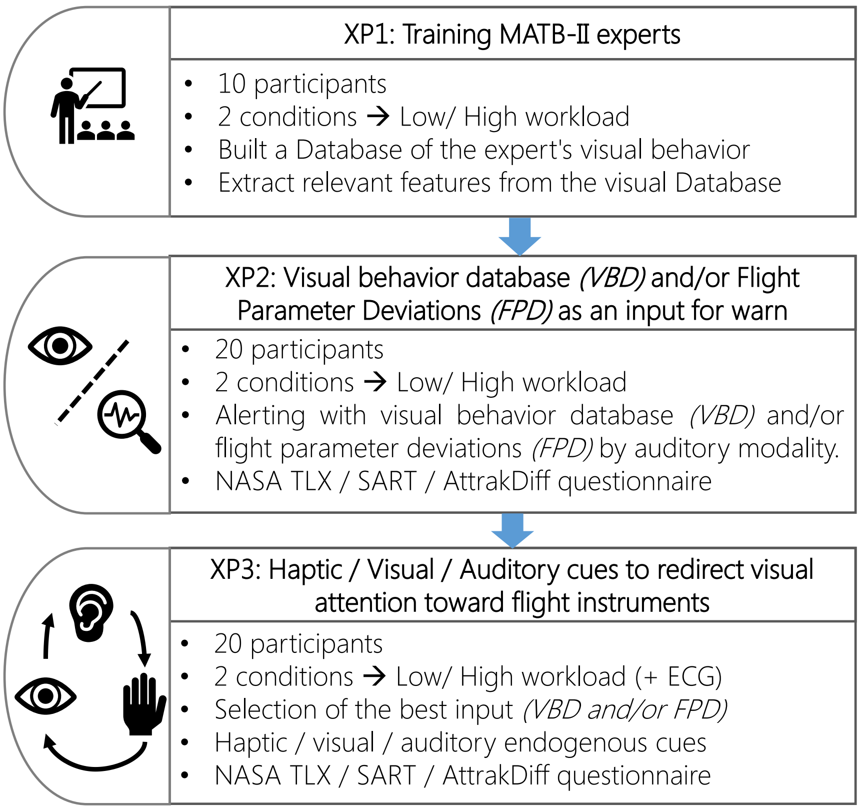


Figure 7.3: Overview of the three experiments

CHAPTER 8

DISCUSSION, PERSPECTIVES, AND CONCLUSION

Sommaire

8.1 Discussions

During this thesis work, we conducted experiments in ecological settings following the Neuroergonomics approach. All experiments involved professional airline pilots in order to deploy a user-centered approach. The overall purpose can be summed up following the question: **"which eye-tracking applications and metrics can help reducing monitoring issues in the commercial aviation field?"** This question raised several problematic investigated in this thesis. The first one was: **"Does eye-tracking techniques can redirect visual attention toward flight instruments in ecological context"**. The second one was: **"Does eye tracking metrics can distinguish over-focalization or under-focalization processing mode in the cockpit during flight phases"**. The third one was: **"how eye-tracking metrics can highlight visual scanning strategies built by the expertise during a manual landing task?"** The last one would have liked to answer two more questions: **"How to improve the FETA flight assistant based on flight parameter and/or visual scanning strategies?"** and **"What is the best warning sensory modality to use when the scanning strategies are no longer optimal?"**

8.1.1 Does eye-tracking techniques can redirect visual attention toward flight instruments in ecological context

The first experiment (Chapter 4) aimed to build a database of expert's visual behavior (concerning flight instruments) in the cockpit to design an eye-tracking rules-based flying assistant. This flying assistant emitted an auditory notification when pilots did not comply with the standard database. This first part of the experiment, which aimed at building the visual database, involved 16 certified pilots (ATPL/CPL) with more than 1,600 flight hours set up in a flight simulator. Then, a second part involving 5 certified pilots less experienced was conducted. We analyzed the visual behavior of these 5 pilots with and without enabling the FETA system in flight simulator scenarios performed according to three different levels of workload. Concerning the warning system assessment, no clear improvement in the maintenance of flight parameters (speed, vertical speed, and heading) was allowed by FETA. Subjective results showed that FETA was detrimental concerning situation awareness during scenarii with easy difficulty. Furthermore, we found that the FETA system helped to redirect visual attention toward the flight instruments which was the subject of notifications. These results confirm the hypothesis concerning the possibility of using a database of expert visual behavior to improve the monitoring of flight instruments. However, no significant results have been demonstrated regarding the improvement of flight parameters deviations, situation awareness, and subjective workload. The mixed results of this study were probably related to some shortcomings in the experiment that hampered the possibility to draw clear conclusions on the effectiveness of such an assistant, especially regarding the situational awareness aspects. First, a limitation concerns the access to a larger number of pilots in order to obtain clear-cut results regarding the FETA assessment. One of the limitations reported by the pilot group was the too frequent triggering of alarms, generating unnecessary alerts. This remark called the use of time thresholds based on average non-dwell time into question and raised the question of new metrics. One of the limitations was also the use of the auditory modality, already extremely employed in the cockpit (e.g. by the synthetic voice and ATCO) which pave the way for other alert modalities.

8.1.2 Does eye tracking metrics can distinguish over-focalization or under-focalization processing mode in the cockpit

This second experiment aimed to investigate new metrics to qualify the visual behaviour of pilots. Literature in eye-tracking proposes the K-coefficient to distinguish visual processing mode by subtracting the z-score of fixation durations to the z-score of saccade amplitudes. A positive value of K reflects a behavior with long fixations followed by short saccades (more "fixational" behavior). While a negative value reflects a behavior with short fixations followed by large saccade (more "saccadic" behavior). This index makes it possible to qualify as focal behavior when the value of K is positive and as ambient when the value of K is negative. The K coefficient is a proven measure for qualifying attentional modes (focal vs ambient) in tasks such as map viewing and/or artwork exploration. Involving 14 certified pilots with 11,500 flight hours in a full flight simulator, this experiment aimed at adapting the K coefficient based on flight instruments during flight phases. The results showed a sensitivity of the Modified K coefficient compared to the classical K coefficient during the different flight phases. This study verified the transposition of the K-coefficient AOI based, thus validating the use of such a metric in the cockpit to discern attentional modes (focal/ambient) considering the flight instruments. The k-coefficient can be used to discern in particular attentional tunneling phenomena where overfocus phenomena is associated with focal attentional modes. A recent study (Kortschot and Jamieson, 2020) has shown the possibility to classify attentional tunneling through behavioral indices. One of the problems encountered during this study concerns the characterization of attentional tunneling. The solution provided by this study is based on the time required to complete the task, whereas a visual metric such as the K coefficient could be an interesting approach. Other work should explore the use of the k-coefficient in situations offering different levels of automation as well as in different experimental conditions (clear weather, rain, engine failure, ...). Finally, an interesting idea would be the use of the K-coefficient on sliding time windows to detect dynamic changes of visual activity. This kind of metric would allow the on-going use of the K-Coefficient.

8.1.3 How eye-tracking metrics can highlight visual scanning strategies built by the expertise during a manual landing task

This third experiment aimed to investigate metrics for visual scanning strategies by using data from expert pilots when building the database (chapter 4) by comparing them with a group of novices. This study highlighted the differences between novices and expert pilots concerning visual scanning strategies and flight performances. Similar results to works have been found, with more dwells and shorter average dwell for pilots compared to novices. This has been interpreted in the literature as an optimization of the visual information processing allowing faster extraction of relevant information when consulting a flight instrument. The result confirmed that expertise exerts a top-down modulation on gaze behaviour. A wide variety of standard and advanced metrics were used to uncover the modification of the gaze behavior bring by the expertise. Expert pilots have a more efficient perception of the information, better dispersion of their attention, and more elaborate visual patterns. Expertise makes it possible, despite a dual-task costly in visuo-attentional resources, to maintain the type of visual patterns linked with the flying task (i.e. the irrelevant dual-task did not alter the nominal visual behavior). These results confirmed the hypothesis formulated by Reingold and Sheridan, 2011 concerning greater perceptual effectiveness of experts as "superior perceptual encoding of domain-related patterns". Overall, the eye metrics used in this research are relevant to finely assess the pilot's gaze behavior in the cockpit and can contribute to better characterize visual scanning in the cockpit, an important topic for safety C.-S. Yu et al., 2016. Future work should focus on the integration of these metrics in a flying assistant, such as FETA, to discern when current visual scanning strategies deviate from expert visual scanning strategies. Further research should consider more different expertise levels, from novice to expert, e.g., every 1000 hours, to finely examine the impact of the expertise on the visual strategies. This flying assistant should integrate the pilot's profile in its potential features.

8.2 Perspectives

8.2.1 How to improve the FETA flight assistant based on flight and/or eye parameters". and "What is the best way to warn a pilot when the scanning strategies are no longer optimal ?"

Unfortunately, this thesis cannot answer the following questions. Nevertheless, it provides a sound protocol for investigating these issues. Future work should address the issues of human-cockpit interaction.

8.2.2 From Human-Computer Interaction (HCI) toward Human-Cockpit Interaction (HCI)

One of the priority areas addressed in this thesis was to investigate AOI-based eye tracking in order to deal with the flight instrument deviations. If we assume the existence of a flight assistant who could on-going detect inappropriate visual behaviour or mental states. It will be essential to consider the interaction between the human and the cockpit and to warn pilots during these inappropriate behaviors. One of the areas for improvement concerning this interaction, which must take place without risk for the crew will be to design, in the best possible way, multimodal alerts that can warn the crew depending on the flight phase, the priority and/or urgency of the request, and the workload. Such a system should consider the occupied channel in order to select the best modalities or multimodalities to warn. The designs concerning the modalities used in the protocol presented in chapter 7 remain simplistic in the light of the possibility of interaction that is offered when using multi-modality. One of the perspectives would also be to design new ways to redirect visual attention such as visual cues in the vicinity of the gaze or to create visual salience for example. One of the perspectives would also be to design new ways to redirect visual attention such as visual cues in the vicinity of the gaze, or to create visual salience. For example, in case of the gaze would be too far from an appropriate Area Of Interest an auditory or haptic signal can be a trigger. Thus, various levels of interaction would be deployed.

8.2.3 FETA improvement and other research questions

One of the areas of improvement concerns the improvement of the flight assistant (FETA). The development of a wearable FETA interface co-designed with flight instructors would also make it possible to improve learning and better monitor the progress of pilots during their training. This interface could consider the metrics developed in the thesis in order to provide a wide spectrum of indicators allowing the instructor to check a learner's visual behaviour. One of the major problems concerns the data access and a promising avenue would be the integration of eye tracking in the wild by proposing to airline companies Eye tracking integration to obtain big data. This amount of data can be used next to improve FETA system. For example, by adapting at the pilot's profile, examine the effect of various automation levels or aircraft configuration on visual behavior, or investigate the differences in visual strategies between pilots carrying out different flights (short-haul, medium-haul, and long-haul). A simpler way to access the pilot's eye-tracking data would be to be able to interpret the pilots' gaze by analyzing the black boxes from aircraft accidents/incident. This would make it possible to replay accident scenarios / or scenario and test the effectiveness of the FETA system by confronting pilots with these flying scenarios in a flight simulator. On the other hand, analyzing the black boxes to estimate the gaze could be interesting for the investigation itself. A longitudinal study following the pilots from the flying school until their assignment in airlines will allow to compare their visual scanning strategies during their training. It would be an interesting avenue to check the consistency of the visual scanning strategies and patterns at the end of the training course or even after a few years of experience. An interesting avenue would also be to investigate these visual scanning strategies with aging. Finally, the literature (chapter 3) agrees that there are mental states that can be discerned using eye-tracking data. In order to follow this trend, numerous studies are still to be carried out by coupling other sensors such as the EEG and ECG to validate the metrics.

8.3 Conclusion

This thesis work has brought to light several contributions of eye tracking to improve safety and reduce human errors. We especially focused on the use of eye tracking in the cockpit to investigate monitoring issues during landing. Following Neuroergonomics approach, we deployed experiments in ecological settings involving certified-commercial pilots for user-centered approach. These experiments allowed to:

- Investigate the question of a flying assistant in the cockpit eye-tracking based;

- Distinguish from under- to over –focalization using ambient/focal mode analysis in the cockpit;
- Present a wide range of standard and advanced ocular metrics to compare visual scanning strategies;
- Present a protocol to investigate various interaction modalities in the cockpit.

Overall, the eye metrics used in this research thesis are relevant to finely assess pilot's gaze behavior in the cockpit and can contribute to better characterize pilots' visual scanning. Taken together, these findings could be used to enhance the research and development for eye tracking integration in the cockpit. Finally, we believe that the eye metrics employed in this study can be also useful for practitioners and researchers in other fields such as air traffic control and automotive.

CONTRIBUTIONS

Peer-Reviewed Conference Paper Article & Journal Paper

- **Christophe Lounis**, Vsevolod Peysakhovich, and Mickaël Causse. 2018. **Intelligent cockpit: eye tracking integration to enhance the pilot-aircraft interaction.** In Proceedings of the 2018 ACM Symposium on Eye Tracking Research and Applications (ETRA '18). Association for Computing Machinery, New York, NY, USA, Article 74, 1–3. DOI:https://doi.org/10.1145/3204493.3207420
- **Christophe Lounis**, Almoctar Hassoumi, Olivier Lefrancois, Vsevolod Peysakhovich, and Mickael Causse. 2020. **Detecting ambient/focal visual attention in professional airline pilots with a modified Coefficient K: a full flight simulator study.** In ACM Symposium on Eye Tracking Research and Applications (ETRA '20 Adjunct). Association for Computing Machinery, New York, NY, USA, Article 25, 1–6. DOI:https://doi.org/10.1145/3379157.3391412
- **Lounis C.**, Peysakhovich V., Causse M. (2020) Flight Eye Tracking Assistant (FETA): Proof of Concept. In: Stanton N. (eds) Advances in Human Factors of Transportation. AHFE 2019. Advances in Intelligent Systems and Computing, vol 964. Springer, Cham. https://doi.org/10.1007/978-3-030-20503-4_66
- **Christophe Lounis**, Vsevolod Peysakhovich, and Mickael Causse. (2020). **Lempel-Ziv Complexity of dwell sequences: visual scanning pattern differences between novice and expert aircraft pilots.** In 1st International Workshop on Eye-

Tracking in Aviation (ETAVI '20). In ETH Zurich Research Collection. ETH Library. https://doi.org/10.3929/ethz-b-000407653

- **C. Lounis**, V. Peysakovich, M. Causse. **Visual scanning strategies in the cockpit and pilot expertise: a flight simulator study.** PLOS ONE, In review.

Proceedings,Talk, and Poster

- **Christophe Lounis**, Vsevolod Peysakhovich, and Mickaël Causse. 2018. **Intelligent cockpit: eye tracking integration to enhance the pilot-aircraft interaction.** In Proceedings of the 2018 ACM Symposium on Eye Tracking Research and Applications (ETRA '18). Association for Computing Machinery, New York, NY, USA, Article 74, 1–3. DOI:https://doi.org/10.1145/3204493.3207420
- **Christophe Lounis**, Vsevolod Peysakhovich, and Mickaël Causse. 2018. **FETA : A flying assistant that looks deep into pilot's eyes**. 9th Young Researchers in Human-Computer Interaction Meeting (RJC IHM 2019). Association Francophone d'Interaction Homme-Machine (AFIHM)
- **C. Lounis**, V. Peysakovich, M. Causse. **Detecting pilots' expertise using transition matrix measures: a machine learning approach**. In International Conference on Cognitive Aircraft Systems (ICCAS '20).https://doi.org/10.34849/cfsb-t270

BIBLIOGRAPHY

Airplanes, B. C. (2016). Statistical summary of commercial jet airplane accidents: Worldwide operations 1959–2015. *Aviation Safety, Boeing Commercial Airplanes, Seattle, WA. http://www. boeing. com/resources/boeingdotcom/company/about_bca/pdf/statsum. pdf.*

Allsop, J., & Gray, R. (2014). 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.

Ancel, E., & Shih, A. (2012). The analysis of the contribution of human factors to the in-flight loss of control accidents. In *12th aiaa aviation technology, integration, and operations (atio) conference and 14th aiaa/issmo multidisciplinary analysis and optimization conference* (p. 5548).

Andrews, T. J., & Coppola, D. M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision research*, *39*(17), 2947–2953.

Antes, J. R. (1974). The time course of picture viewing. (pp. 62–70). doi:10.1037/h0036799

Aricò, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., & Babiloni, F. (2018). Passive bci beyond the lab: Current trends and future directions. *Physiological measurement*, *39*(8), 08TR02.

Arico, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., Colosimo, A., & Babiloni, F. (2017). Passive bci in operational environments: Insights, recent advances, and future trends. *IEEE Transactions on Biomedical Engineering*, *64*(7), 1431–1436.

Authority, C. A. (2013). Monitoring matters, guidance on the development of pilot monitoring skills. *CAA Paper*, *2*.

Bai, Y., Liang, Z., & Li, X. (2015). A permutation lempel-ziv complexity measure for eeg analysis. *Biomedical Signal Processing and Control*, *19*, 102–114.

Baldwin, C. L., Roberts, D. M., Barragan, D., Lee, J. D., Lerner, N., & Higgins, J. S. (2017). Detecting and quantifying mind wandering during simulated driving. *Frontiers in human neuroscience*, *11*, 406.

Baldwin, C. L., Spence, C., Bliss, J. P., Brill, J. C., Wogalter, M. S., Mayhorn, C. B., & Ferris, T. K. (2012). Multimodal cueing: The relative benefits of the auditory, visual, and tactile channels in complex environments. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 56, *1*, pp. 1431–1435). SAGE Publications Sage CA: Los Angeles, CA.

Ballard, J. C. (1996). Computerized assessment of sustained attention: A review of factors affecting vigilance performance. *Journal of clinical and experimental neuropsychology*, *18*(6), 843–863.

Barbé, J., Spaggiari, L., Clay, A., Bérard, P., Aissani, A., & Mollard, R. (2016). Why and how to study multimodal interaction in cockpit design. In *Proceedings of the 15th ergo'ia" ergonomie et informatique avancée" conference* (pp. 1–8).

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, space, and environmental medicine*.

Billings, C. E. (2018). *Aviation automation: The search for a human-centered approach*. CRC Press.

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2017). Visualization of eye tracking data: A taxonomy and survey. In *Computer graphics forum* (Vol. 36, *8*, pp. 260–284). Wiley Online Library.

Blignaut, P. J., Beelders, T. R., & So, C. (2008). The visual span of chess players. In *Proceedings of the 2008 symposium on eye tracking research & applications* (pp. 165–171).

Boy, G. A. (2014). Requirements for single pilot operations in commercial aviation: A first high-level cognitive function analysis. In *Csdm (posters)* (pp. 227–234). Citeseer.

Broadbent, D. E. (2013). *Perception and communication*. Elsevier.

Burgin, M. (2017). Inductive complexity and shannon entropy. In *Information and complexity* (pp. 16–32). World Scientific.

Caldwell, J. A. (2005). Fatigue in aviation. *Travel medicine and infectious disease*, *3*(2), 85–96.

Caldwell, J., & Caldwell, J. L. (2016). *Fatigue in aviation: A guide to staying awake at the stick*. Routledge.

Callan, D. E., & Dehais, F. (2019). Neuroergonomics for aviation, 55–58.

Casner, S. M., & Schooler, J. W. (2014). Thoughts in flight: Automation use and pilots' task-related and task-unrelated thought. *Human factors*, *56*(3), 433–442.

Causse, M., Baracat, B., Pastor, J., & Dehais, F. (2011). Reward and uncertainty favor risky decision-making in pilots: Evidence from cardiovascular and oculometric measurements. *Applied psychophysiology and biofeedback*, *36*(4), 231–242.

Causse, M., Dehais, F., Péran, P., Sabatini, U., & Pastor, J. (2013). The effects of emotion on pilot decision-making: A neuroergonomic approach to aviation safety. *Transportation research part C: emerging technologies*, *33*, 272–281.

Chang, W., Hwang, W., & Ji, Y. G. (2011). Haptic seat interfaces for driver information and warning systems. *International Journal of Human-Computer Interaction*, *27*(12), 1119–1132.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & cognition*, *29*(8), 1146–1152.

Cholewiak, R. W., Brill, J. C., & Schwab, A. (2004). Vibrotactile localization on the abdomen: Effects of place and space. *Perception & psychophysics*, *66*(6), 970–987.

Combrisson, E., & Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of neuroscience methods*, *250*, 126–136.

Comstock Jr, J. R., & Arnegard, R. J. (1992). The multi-attribute task battery for human operator workload and strategic behavior research.

Cowen, L., Ball, L. J., & Delin, J. (2002). An eye movement analysis of web page usability. In *People and computers xvi-memorable yet invisible* (pp. 317–335). Springer.

Crane, H. D. (1994). The purkinje image eyetracker, image stabilization, and related forms of stimulus manipulation. *Visual science and engineering: Models and applications*, 15–89.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 94.

Dahlstrom, N., & Nahlinder, S. (2009). Mental workload in aircraft and simulator during basic civil aviation training. *The International journal of aviation psychology*, *19*(4), 309–325.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. *Annual review of psychology*, *31*(1), 309–341.

De Valois, R. L., & De Valois, K. K. (1993). A multi-stage color model. *Vision research*, *33*(8), 1053–1065.

Dehais, F., & Callan, D. E. (2019). A neuroergonomics approach to human performance in aviation. *Improving Aviation Performance through Applying Engineering Psychology: Advances in Aviation Psychology*, *3*, 123.

Dehais, F., Causse, M., Régis, N., Menant, E., Labedan, P., Vachon, F., & Tremblay, S. (2012). Missing critical auditory alarms in aeronautics: Evidence for inattentional deafness? In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 56, *1*, pp. 1639–1643). Sage Publications Sage CA: Los Angeles, CA.

Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., & Tremblay, S. (2014). Failure to detect critical auditory alerts in the cockpit: Evidence for inattentional deafness. *Human factors*, *56*(4), 631–644.

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making*, *1*(3), 271–285.

Di Stasi, L. L., Catena, A., Canas, J. J., Macknik, S. L., & Martinez-Conde, S. (2013). Saccadic velocity as an arousal index in naturalistic tasks. *Neuroscience & Biobehavioral Reviews*, *37*(5), 968–975.

Di Stasi, L. L., & Diaz-Piedra, C. (2019). Eye movements research in aviation: Past, present, and future. *Improving Aviation Performance through Applying Engineering Psychology, Advances in Aviation Psychology*, *3*, 143–162.

Diaz-Piedra, C., Rieiro, H., Cherino, A., Fuentes, L. J., Catena, A., & Di Stasi, L. L. (2019). The effects of flight complexity on gaze entropy: An experimental study with fighter pilots. *Applied ergonomics*, *77*, 92–99.

Diaz-Piedra, C., Rieiro, H., Suárez, J., Rios-Tejada, F., Catena, A., & Di Stasi, L. L. (2016). Fatigue in the military: Towards a fatigue detection test based on the saccadic velocity. *Physiological measurement*, *37*(9), N62.

Dismukes, R., & Berman, B. (2010). Checklists and monitoring in the cockpit: Why crucial defenses sometimes fail.

Distler, C., & Hoffmann, K.-P. (2011). Visual pathway for the optokinetic reflex in infant macaque monkeys. *Journal of Neuroscience*, *31*(48), 17659–17668.

Duchowski, A., & Krejtz, K. (2017). Visualizing dynamic ambient/focal attention with coefficient k. (pp. 217–233). doi:10.1007/978-3-319-47024-5_13

Duchowski, A. T. (2007). Eye tracking methodology. *Theory and practice*, *328*(614), 2–3.

Duchowski, A. T., & Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice*. Springer.

Endsley, M. R. (1997). Supporting situation awareness in aviation systems. In *1997 ieee international conference on systems, man, and cybernetics. computational cybernetics and simulation* (Vol. 5, pp. 4177–4181). IEEE.

Ephrath, A., Tole, J., Stephens, A. T., & Young, L. (1980). Instrument scan—is it an indicator of the pilot's workload? In *Proceedings of the human factors society annual meeting* (Vol. 24, *1*, pp. 257–258). SAGE Publications Sage CA: Los Angeles, CA.

Erp, J. B. F. V., Veen, H. A. H. C. V., Jansen, C., & Dobbins, T. (2005). Waypoint navigation with a vibrotactile waist belt. *ACM Trans. Appl. Percept.* *2*(2), 106–117. doi:10.1145/1060581.1060585

et d'Analyses, B. d. et al. (2012). Final report on the accident on 1st june 2009 to the airbus a330-203 registered f-gzcp operated by air france flight af 447 rio de janeiro–paris. *Paris: BEA*.

Fairclough, S. H., & Mulder, L. (2012). Psychophysiological processes of mental effort investment.

Feit, A. M., Williams, S., Toledo, A., Paradiso, A., Kulkarni, H., Kane, S., & Morris, M. R. (2017). Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In *Proceedings of the 2017 chi conference on human factors in computing systems* (pp. 1118–1130). CHI '17. doi:10.1145/3025453.3025599

Fernández, A., López-Ibor, M.-I., Turrero, A., Santos, J.-M., Morón, M.-D., Hornero, R., ... López-Ibor, J. J. (2011). Lempel–ziv complexity in schizophrenia: A meg study. *Clinical Neurophysiology*, *122*(11), 2227–2235.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1949). *Eye fixations of aircraft pilots. iii. frequency, duration, and sequence fixations when flying air force ground-controlled approach system (gca).* AIR MATERIEL COMMAND WRIGHT-PATTERSON AFB OH.

Follet, B., Le Meur, O., & Baccino, T. (2011). New insights into ambient and focal visual fixations using an automatic classification algorithm. *i-Perception*, *2*(6), 592–610.

Fox, J., Merwin, D., Marsh, R., McConkie, G., & Kramer, A. (1996). Information extraction during instrument flight: An evaluation of the validity of the eye-mind hypothesis. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 40, *2*, pp. 77–81). SAGE Publications Sage CA: Los Angeles, CA.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*. Springer series in statistics New York.

Gaffary, Y., & Lécuyer, A. (2018). The use of haptic and tactile information in the car to improve driving safety: A review of current technologies. *Frontiers in ICT*, *5*, 5.

Ganesan, A., Alakhras, M., Brennan, P. C., & Mello-Thoms, C. (2018). A review of factors influencing radiologists' visual search behaviour. *Journal of medical imaging and radiation oncology*, *62*(6), 747–757.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, *23*(4), 523–552.

Glaholt, M. G. (2014). *Eye tracking in the cockpit: A review of the relationships between eye movements and the aviators cognitive state*. Defence Research and Development Toronto (Canada).

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International journal of industrial ergonomics*, *24*(6), 631–645.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: Design implications. In *Proceedings of the 2002 symposium on eye tracking research & applications* (pp. 51–58).

Group, A. P. M. W. et al. (2014). A practical guide for improving flight path monitoring. *Final Report of the Active Pilot Monitoring Working Group. Alexandria, VA: Flight Safety Foundation*.

Grunwald, P., & Vitanyi, P. M. B. (2004). Shannon information and kolmogorov complexity. *CoRR*, *1*.

Hammoud, R. I. (2008). *Passive eye monitoring: Algorithms, applications and experiments*. Springer Science & Business Media.

Hankins, T. C., & Wilson, G. F. (1998). A comparison of heart rate, eye activity, eeg and subjective measures of pilot mental workload during flight. *Aviation, space, and environmental medicine*, *69*(4), 360.

Hansen, D. W., & Ji, Q. (2009). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE transactions on pattern analysis and machine intelligence*, *32*(3), 478–500.

Harris Sr, R. L., Glover, B. J., & Spady Jr, A. A. (1986). Analytical techniques of pilot scanning behavior and their application.

Hayashi, M. (2004). *Hidden markov models for analysis of pilot instrument scanning and attention switching* (Doctoral dissertation, Massachusetts Institute of Technology).

Hayashi, M., Beutter, B., & McCann, R. S. (2005). Hidden markov model analysis for space shuttle crewmembers' scanning behavior. In *2005 ieee international conference on systems, man and cybernetics* (Vol. 2, pp. 1615–1622). IEEE.

Helmreich, R. L. (2000). On error management: Lessons from aviation. *Bmj*, *320*(7237), 781–785.

Henderson, J. M., & Luke, S. G. (2014). Stable individual differences in saccadic eye movements during reading, pseudoreading, scene viewing, and scene search. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1390.

Hendricks, D., Fell, J., & Freedman, M. (1999). *Identifying unsafe driver actions that lead to fatal car-truck crashes*. Summary Technical Report, DTNH22-94-C-05020, US Department of Transportation.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & psychophysics*, *57*(6), 787–795.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Hubel, D. H., & Wiesel, T. N. (1974). Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor. *Journal of Comparative Neurology*, *158*(3), 295–305.

Jarvis, S. R. (2017). Concurrent pilot instrument monitoring in the automated multi-crew airline cockpit. *Aerospace medicine and human performance*, *88*(12), 1100–1106.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, space, and environmental medicine*.

Jones, D. H. (1985). An error-dependent model of instrument-scanning behavior in commercial airline pilots. ph. d. thesis-may 1983.

Jones, L. A., & Sarter, N. B. (2008). Tactile displays: Guidance for their design and application. *Human factors*, *50*(1), 90–111.

Jones, R. E., Milton, J. L., & Fitts, P. M. (1949). Eye fixations of aircraft pilots, i. a review of prior eye-movement studies and a description of a technique for recording the frequency, duration and sequences of eye-fixations during instrument flight. *Wright Patterson AFB, OH, USAF Tech. Rep*, *5837*.

Kahneman, D. [Daniel]. (1973). *Attention and effort*. Citeseer.

Kang, Z., & Landry, S. J. (2014). Using scanpaths as a learning method for a conflict detection task of multiple target tracking. *Human factors*, *56*(6), 1150–1162.

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during vfr flight. In *Proceedings of the 11th international symposium on aviation psychology* (Vol. 6). Citeseer.

Kharoufah, H., Murray, J., Baxter, G., & Wild, G. (2018). A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016. *Progress in Aerospace Sciences*, *99*, 1–13.

Klein, R. M. (2004). On the control of visual orienting.

Kortschot, S. W., & Jamieson, G. A. (2020). Classification of attentional tunneling through behavioral indices. *Human factors*, *62*(6), 973–986.

Krejtz, K., Çöltekin, A., Duchowski, A. T., & Niedzielska, A. (2017). Using coefficient k to distinguish ambient/focal visual attention during map viewing. *Journal of eye movement research*, *10*(2), 1–13.

Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A., & Kopacz, A. (2016). Discerning ambient/focal attention with coefficient k. *ACM Trans. Appl. Percept. 13*(3). doi:10.1145/2896452

Krejtz, K., Duchowski, A., Szmidt, T., Krejtz, I., González Perilli, F., Pires, A., ... Villalobos, N. (2015). Gaze transition entropy. *ACM Transactions on Applied Perception (TAP)*, *13*(1), 4.

Lanzetta, T. M., Dember, W. N., Warm, J. S., & Berch, D. B. (1987). Effects of task type and stimulus heterogeneity on the event rate function in sustained attention. *Human Factors*, *29*(6), 625–633. doi:10.1177/001872088702900602

Laurens, J., Strauman, D., & Hess, B. J. (2011). Spinning versus wobbling: How the brain solves a geometry problem. *Journal of Neuroscience*, *31*(22), 8093–8101.

Lefrancois, O., Matton, N., Gourinat, Y., Peysakhovich, V., & Causse, M. (2016). The role of pilots' monitoring strategies in flight performance.

Lemercier, C., & Cellier, J.-M. (2008). Les défauts de l'attention en conduite automobile: Inattention, distraction et interférence. *Le travail humain*, *71*(3), 271–296.

Lenné, M. G., Roady, T., & Kuo, J. (2020). 11 driver state monitoring for decreased fitness to drive. *Handbook of Human Factors for Automated, Connected, and Intelligent Vehicles*.

Leung-Yan-Cheong, S., & Cover, T. (1978). Some equivalences between shannon entropy and kolmogorov complexity. *IEEE Transactions on Information Theory*, *24*(3), 331–338.

Li, W.-C., Chiu, F.-C., Kuo, Y.-s., & Wu, K.-J. (2013). The investigation of visual attention and workload by experts and novices in the cockpit. In *International conference on engineering psychology and cognitive ergonomics* (pp. 167–176). Springer.

Li, W.-C., Chiu, F.-C., & Wu, K.-J. (2012). The evaluation of pilots performance and mental workload by eye movement.

Lorenz, B., & Biella, M. (2006). Evaluation of onboard taxi guidance support on pilot performance in airport surface navigation. In *Proceedings of the human factors and ergonomics* *society annual meeting* (Vol. 50, *1*, pp. 111–115). SAGE Publications Sage CA: Los Angeles, CA.

Lounis, C., Peysakhovich, V., & Causse, M. (2018). Intelligent cockpit: Eye tracking integration to enhance the pilot-aircraft interaction. In *Proceedings of the 2018 acm symposium on eye tracking research & applications* (p. 74). ACM.

Lounis, C., Peysakhovich, V., & Causse, M. (2019). Flight eye tracking assistant (feta): Proof of concept. In *International conference on applied human factors and ergonomics* (pp. 739–751). Springer.

Marcus, J. H., & Rosekind, M. R. (2017). Fatigue in transportation: Ntsb investigations and safety recommendations. *Injury prevention*, *23*(4), 232–238.

Marey, E.-J. (1894). *Le mouvement*. G. Masson.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature reviews neuroscience*, *5*(3), 229–240.

Matthews, G., & Davies, D. R. (1998). Arousal and vigilance: The role of task demands.

Mele, M. L., & Federici, S. (2012). Gaze and eye-tracking solutions for psychological research. *Cognitive processing*, *13*(1), 261–265.

Melsted, P., & Pritchard, J. K. (2011). Efficient counting of k-mers in dna sequences using a bloom filter. *BMC bioinformatics*, *12*(1), 333.

Memmert, D., Simons, D. J., & Grimme, T. (2009). The relationship between visual attention and expertise in sports. *Psychology of Sport and Exercise*, *10*(1), 146–151.

Moray, N. (2013). *Mental workload: Its theory and measurement*. Springer Science & Business Media.

Mustafi, D., Engel, A. H., & Palczewski, K. (2009). Structure of cone photoreceptors. *Progress in retinal and eye research*, *28*(4), 289–302.

Myles, K., & Binseel, M. S. (2007). *The tactile modality: A review of tactile sensitivity and human tactile interfaces*. ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD HUMAN RESEARCH and ENGINEERING ...

Norris, J. R. (1998). *Markov chains*. Cambridge university press.

Noy, Y. I., Lemoine, T. L., Klachan, C., & Burns, P. C. (2004). Task interruptability and duration as measures of visual distraction. *Applied Ergonomics*, *35*(3), 207–213.

Ognjanovic, S., Thüring, M., Murphy, R. O., & Hölscher, C. (2019). Display clutter and its effects on visual attention distribution and financial risk judgment. *Applied ergonomics*, *80*, 168–174.

Oster Jr, C. V., Strong, J. S., & Zorn, C. K. (2013). Analyzing aviation safety: Problems, challenges, opportunities. *Research in transportation economics*, *43*(1), 148–164.

Ottati, W. L., Hickox, J. C., & Richter, J. (1999). Eye scan patterns of experienced and novice pilots during visual flight rules (vfr) navigation. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 43, *1*, pp. 66–70). SAGE Publications Sage CA: Los Angeles, CA.

Palmer, S., & Palmer, S. (1999). Visual selection: Eye movements and attention. *SE Palmer,(Vision Science: Photons to Phenomenology), Bradford Books/MIT Press, Cambridge, MA*, 519–571.

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A.-K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2*(2). doi:10.16910/jemr.2.2.4

Parasuraman, R. (1987). Human-computer monitoring. *Human Factors*, *29*(6), 695–706.

Parasuraman, R. (2003). Neuroergonomics: Research and practice. *Theoretical issues in ergonomics science*, *4*(1-2), 5–20.

Parasuraman, R., & Rizzo, M. (2008). *Neuroergonomics: The brain at work*. Oxford University Press.

Parasurannan, R., & Byrne, E. A. (2002). Automation and human performance in aviation. *Principles and practice of aviation psychology*, 311.

Peißl, S., Wickens, C. D., & Baruah, R. (2018). Eye-tracking measures in aviation: A selective literature review. *The International Journal of Aerospace Psychology*, *28*(3-4), 98–112.

Peters, R. A. (1964). *Dynamics of the vestibular system and their relation to motion perception, spatial disorientation, and illusions*. Systems Technology, Incorporated.

Peysakhovich, V. (2016). *Study of pupil diameter and eye movements to enhance flight safety* (Doctoral dissertation, Toulouse 3).

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.

Posner, M. I. (1982). Cumulative development of attentional theory. *American Psychologist*, *37*(2), 168.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160.

Previc, F. H., & Ercoline, W. R. (2004). *Spatial disorientation in aviation*. Aiaa.

Reani, M., Peek, N., & Jay, C. (2018). An investigation of the effects of n-gram length in scanpath analysis for eye-tracking research. In *Proceedings of the 2018 acm symposium on eye tracking research & applications* (pp. 1–8).

Regis, N., Dehais, F., Tessier, C., & Gagnon, J.-F. (2012). Ocular metrics for detecting attentional tunneling.

Reingold, E. M., & Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. In *Oxford handbook on eye movements* (Vol. 528). Oxford University Press Oxford, UK.

Russi-Vigoya, M. N., & Patterson, P. (2015). Analysis of pilot eye behavior during glass cockpit simulations. *Procedia Manufacturing*, *3*, 5028–5035.

Ryffel, C. P., Muehlethaler, C. M., Huber, S. M., & Elfering, A. (2019). Eye tracking as a debriefing tool in upset prevention and recovery training (uprt) for general aviation pilots. *Ergonomics*, *62*(2), 319–329.

Safety, A. (2018). Statistical summary of commercial jet airplane accidents worldwide operations: 1959–2017. *Boeing Commercial Airplanes*.

Salas, E., Maurino, D., & Curtis, M. (2010). Human factors in aviation: An overview. In *Human factors in aviation* (pp. 3–19). Elsevier.

Salzer, Y., Oron-Gilad, T., Ronen, A., & Parmet, Y. (2011). Vibrotactile "on-thigh" alerting system in the cockpit. *Human factors*, *53*(2), 118–131.

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock Jr, J. R. (2011). The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide.

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain research reviews*, *35*(2), 146–160.

Scannella, S., Peysakhovich, V., Ehrig, F., Lepron, E., & Dehais, F. (2018). Assessment of ocular and physiological metrics to discriminate flight phases in real light aircraft. *Human factors*, *60*(7), 922–935.

Senders, J. W. (1966). A re-analysis of the pilot eye-movement data. *IEEE Transactions on Human Factors in Electronics*, (2), 103–106.

Shannon, C. W., & Weaver, W. (1948). W.:(1949) the mathematical theory of communication. *Press UoI, editor*.

Shapiro, K. L., & Raymond, J. E. (1989). Training of efficient oculomotor strategies enhances skill acquisition. *Acta Psychologica*, *71*(1-3), 217–242.

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2007). Human error and commercial aviation accidents: An analysis using the human factors analysis and classification system. *Human factors*, *49*(2), 227–242.

Sharafi, Z., Shaffer, T., Sharif, B., & Gueheneuc, Y.-G. (2015). Eye-tracking metrics in software engineering. In *2015 asia-pacific software engineering conference (apsec)* (pp. 96–103). IEEE.

Shiferaw, B., Downey, L., & Crewther, D. (2019). A review of gaze entropy as a measure of visual scanning efficiency. *Neuroscience & Biobehavioral Reviews*, *96*, 353–366.

Simon, H., & Chase, W. (1988). Skill in chess. In *Computer chess compendium* (pp. 175–188). Springer.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in cognitive sciences*, *4*(4), 147–155.

Sklar, A. E., & Sarter, N. B. (1999). Good vibrations: Tactile feedback in support of attention allocation and human-automation coordination in event-driven domains. *Human factors*, *41*(4), 543–552.

Sommer, D., & Golz, M. (2010). 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.

Spelke, E., Hirst, W., & Neisser, U. (1976). Skills of divided attention. *Cognition*, *4*(3), 215–230.

Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: A possible measure of fatigue. *Human factors*, *36*(2), 285–297.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of experimental psychology: Applied*, *9*(1), 23.

Sumwalt, R. L., Thomas, R., & Dismukes, K. (2002). Enhancing flight-crew monitoring skills can increase flight safety. In *Annual international air safety seminar* (Vol. 55, pp. 175–206). Flight Safety Foundation; 1998.

Svensson, E., Angelborg-Thanderez, M., Sjöberg, L., & Olsson, S. (1997). Information complexity-mental workload and performance in combat aircraft. *Ergonomics*, *40*(3), 362–380.

Talebinejad, M., Chan, A. D., & Miri, A. (2011). A lempel–ziv complexity measure for muscle fatigue estimation. *Journal of Electromyography and Kinesiology*, *21*(2), 236–241.

Tattersall, A. J., & Foord, P. S. (1996). An experimental evaluation of instantaneous self-assessment as a measure of workload. *Ergonomics*, *39*(5), 740–748.

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.

Team, C. A. S. (2008). Safety enhancement 30, revision-5: Mode awareness and energy state management aspects of flight deck automation.

Thomas, L. C., & Wickens, C. D. (2004). Eye-tracking and individual differences in off-normal event detection when flying with a synthetic vision system display. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 48, *1*, pp. 223–227). Sage Publications Sage CA: Los Angeles, CA.

Tiffin, J., & Bromer, J. (1943). *Analysis of eye fixations and patterns of eye movement in landing a piper cub j-3 airplane*. Civil Aeronautics Administration.

Tole, J., Stephens, A. T., Harris, R., & Ephrath, A. R. (1982). Visual scanning behavior and mental workload in aircraft pilots. *Aviation, Space, and Environmental Medicine*.

Tsai, Y.-F., Viirre, E., Strychacz, C., Chase, B., & Jung, T.-P. (2007). Task performance and eye activity: Predicting behavior relating to cognitive workload. *Aviation, space, and environmental medicine*, *78*(5), B176–B185.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. doi:10.1126/science.7455683. eprint: https://science.sciencemag.org/content/211/4481/453.full.pdf

Tweed, D., & Vilis, T. (1990). Geometric relations of eye position and velocity vectors during saccades. *Vision research*, *30*(1), 111–127.

Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics*, *50*(8), 1235–1249.

Unema, P., Pannasch, S., Joos, M., & Velichkovsky, B. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*, 473–494. doi:10.1080/13506280444000409 van de Merwe, K., van Dijk, H., & Zon, R. (2012). Eye movements as an indicator of situation awareness in a flight simulator experiment. *The International Journal of Aviation Psychology*, *22*(1), 78–95.

Van Erp, J. B., Groen, E. L., Bos, J. E., & Van Veen, H. A. (2006). A tactile cockpit instrument supports the control of self-motion during spatial disorientation. *Human factors*, *48*(2), 219–228.

Veitengruber, J. (1978). Design criteria for aircraft warning, caution, and advisory alerting systems. *Journal of aircraft*, *15*(9), 574–581.

Velichkovsky, B., Joos, M., Helmert, J., & Pannasch, S. (2005). Two visual systems and their eye movements: Evidence from static and dynamic scene perception. *Proceedings of the XXVII Conference of the Cognitive Science Society*.

Verrillo, R. T. (1966). Vibrotactile sensitivity and the frequency response of the pacinian corpuscle. *Psychonomic Science*, *4*(1), 135–136.

Wade, N. J. (2010). Pioneers of eye movement research. *i-Perception*, *1*(2), 33–68.

Wang, Z., & Fu, S. (2014). An analysis of pilot's physiological reactions in different flight phases. In *International conference on engineering psychology and cognitive ergonomics* (pp. 94–103). Springer.

Warm, J. S. [Joel S], Matthews, G., & Finomore Jr, V. S. (2008). Vigilance, workload, and stress. *Performance under stress*, 115–41.

Warm, J. (1993). Vigilance and target detection. *Workload transition: Implications for individual and team performance*, *1*, 139–170.

Washington, N. (1978). Dc: National transportation safety board; 1994. *A Review of Flightcrew-Involved Major Accidents of US Air Carriers*.

Wickens, C. D. (1984). Engineering psychology and human performance. charles e. *Merrill, Columbus, OH*.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, *3*(2), 159–177.

Wickens, C. D. (2005). Attentional tunneling and task management. In *2005 international symposium on aviation psychology* (p. 812).

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

Wickens, C. D., & Alexander, A. L. (2009). Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology*, *19*(2), 182–199.

Wickens, C. D., & Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human factors*, *30*(5), 599–616.

Wiegmann, D. A., & Shappell, S. A. (2017). *A human error approach to aviation accident analysis: The human factors analysis and classification system*. Routledge.

Wilska, A. (1954). On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica*, *31*(2-3), 285–289.

Woods, N. B. S. D̃. D. (1991). *The International Journal of Aviation Psychology*, *1*(1), 45–57.

Wu, X., Wanyan, X., & Zhuang, D. (2015). Pilot's visual attention allocation modeling under fatigue. *Technology and health care*, *23*(s2), S373–S381.

Xiong, W., Wang, Y., Zhou, Q., Liu, Z., & Zhang, X. (2016). The research of eye movement behavior of expert and novice in flight simulation of landing. In *International conference on engineering psychology and cognitive ergonomics* (pp. 485–493). Springer.

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye movements and vision* (pp. 171–211). Springer.

Young, J. J., Tan, H. Z., & Gray, R. (2003). Validity of haptic cues and its effect on priming visual spatial attention. In *11th symposium on haptic interfaces for virtual environment and teleoperator systems, 2003. haptics 2003. proceedings.* (pp. 166–170). IEEE.

Yu, C.-s., Wang, E. M.-y., Li, W.-C., & Braithwaite, G. (2014). Pilots' visual scan patterns and situation awareness in flight operations. *Aviation, space, and environmental medicine*, *85*(7), 708–714.

Yu, C.-S., Wang, E. M.-y., Li, W.-C., Braithwaite, G., & Greaves, M. (2016). Pilots' visual scan patterns and attention distribution during the pursuit of a dynamic target. *Aerospace medicine and human performance*, *87*(1), 40–47.

Zhou, S., Zhang, Z., & Gu, J. (2011). Interpretation of coarse-graining of lempel-ziv complexity measure in ecg signal analysis. In *2011 annual international conference of the ieee engineering in medicine and biology society* (pp. 2716–2719). IEEE.

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

APPENDIX A

RESEARCH ETHICS COMMITTEE FORMS

Neuroergonomics and
Human Factors
Departement of Aerospace Vehicles Design and Control
ISAE-SUPAERO, Université de Toulouse

Preliminary Questionnaire

Participant ID :

Age : years

Sex : M / F

Handedness : Left-handed / Right-handed

Eye colour :

Flight hours : hours

Obtained licences (if any) :

Aircraft flown (if any): :

Education level :

Eyewear : Glasses / Contact lenses / None

Eye condition (myopia, hyperopia, astigmatism, visual correction, etc., if any):

Left Eye	Right Eye

For how many hours, in average, do you sleep every night?

For how many hours did you sleep last night?

How would you rate your fatigue at this moment?

1	2	3	4	5
Not at all tired				Very tired

Figure A.1: Informed Consent form 1/3

Neuroergonomics and
Human Factors
Departement of Aerospace Vehicles Design and Control
ISAE-SUPAERO, Université de Toulouse

Consent to Participate in a Research Study

Title of Study: Monitoring of Pilots' Eye Movements during Approach-Landing phases
Investigators:

Christophe Lounis, Ph.D Researcher.	Dept.: DCAS	Phone: +33 (0)5 61 33 87 58
Vsevolod Peysakhovich, Ph.D.	Dept.: DCAS	Phone: +33 (0)5 61 33 87 46
Mickaël Causse, Ph.D.	Dept.: DCAS	Phone: +33 (0)5 61 33 81 28

Please check all boxes

1. I confirm that I have read and understood the instruction for the experiment. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. ☐
3. I understand that during the experiment my eye movements will be tracked and recorded and this recording will be anonymous. I give permission for this recording to be analysed for the purpose of this experiment. ☐
4. I understand that I will not receive any remuneration for my participation in this experiment. ☐
5. I agree to take part in the experiment. ☐

Participant's Name: ____________________________________

Participant's Signature: ________________________ Date : _____________

Investigator's Signature: ______________________ Date : _____________

Figure A.2: Informed Consent form 2/3

Neuroergonomics and
Human Factors
Departement of Aerospace Vehicles Design and Control
ISAE-SUPAERO, Université de Toulouse

Instructions for Participants

You are about to participate in a research study consisting in monitoring pilots' eye movements during approach-landing phases. Please read this form and ask any questions that you may have before agreeing to participate in the study.

You are going to perform three manual approach-landing phases. Each simulation will begin at the same position and will have the same objective, which is *to land the aircraft on LFBO Runway 14R*. You must do the simulation as if you were flying a real aircraft, performing normal monitoring patterns and paying attention to potential unexpected events (e.g. alarms, traffic, etc.). During some landings, you will be asked to perform another monitoring task in parallel (*double-task*), during the others, you will be asked to perform a normal landing (*single-task*). The parallel task consists of reading loudly the remaining distance to the VOR (as displayed on the Navigation Display) with different frequencies. For example, when asked to report the distance each 0.5 or 0.2 nautical miles, you will orally announce "5 nm, 4.8 nm, 4.6 nm etc." without pronouncing "nautical miles". You will be instructed whether it is double or single task and at what frequency, the remaining distance is to be reported before each simulation by the investigator.

The experiment is performed using the PEGASE flight simulator of DCAS, ISAE-SUPAERO, which is equipped with an eye tracker (Smart Eye). During each simulation, your eye movements will be tracked by this system and the data will be logged. This data will be anonymous. No videos or images of your face will be recorded.

The total experiment duration will be 2 hours. Each landing will last for about 5 minutes. Before starting, you will have 15 minutes to practice and familiarize yourself with the simulator. During this experiment, you will perform one single-task landing and 2 double-task landings. You can ask all remaining questions during this experiment. Between each condition, you will have 1-minute break.

If you agree to participate in this study, please turn off your mobile phone, and sign the consent form. Good luck!

Figure A.3: Informed Consent form 3/3

Université Fédérale
Toulouse
Midi-Pyrénées

Toulouse, le jeudi 7 février 2019

A l'attention de CAUSSE Mickaël

CER : Comité d'Ethique sur les Recherches

Objet : Avis de la commission du 29/01/2019

Numéro d'enregistrement : 2019-131

Titre du projet soumis : **Utilisation de l'oculométrie en vue d'améliorer l'interaction pilote aéronef**

Porteur de projet : CAUSSE Mickaël, laboratoire DCAS, ISAE

Monsieur,

Compte tenu des éléments fournis dans votre demande, le Comité d'Ethique pour les Recherches de l'Université de Toulouse émet l'avis suivant : **Favorable avec recommandations.**

Nous rappelons, par ailleurs, qu'il relève de la responsabilité des chercheurs de se conformer à leurs obligations légales notamment en ce qui concerne les aspects d'homologation du lieu de recherche ou RGPD : Règlement Général sur la Protection des Données.

Nous restons à votre disposition pour toute question.

Les membres du bureau CER.

Pr Maria Teresa Munoz Sastre — Pr Jacques Py — Rémi Capa

CER - Université Fédérale Toulouse Midi-Pyrénées
Département Recherche, Doctorat et Valorisation
41, Allées Jules Guesde - CS 61321 - 31013 Toulouse CEDEX 6 - Tél. : 05 61 10 80 30
Courriel : bureau-cerni@univ-toulouse.fr

Figure A.4: FETA/Visual Scanning Strategies Experiment approval

Université Fédérale
Toulouse Midi-Pyrénées

A l'attention de
LEFRANCOIS Olivier

À Toulouse, le 04 mars 2020

Affaire suivie par :
Sophie ACHTE
CER–DRDV
Courriel : bureau-cerni@univ-toulouse.fr
Tél. : 05 61 10 80 30

Objet : Avis du bureau du 03/03/2020 pour le projet 2020-210

Titre du projet soumis : Étude de la rééducation du circuit visuel des pilotes à l'aide de l'Eye tracking

Porteur de projet : LEFRANCOIS Olivier, laboratoire DCAS, ISAE

Monsieur,

Compte tenu des éléments fournis dans votre demande, le Comité d'Ethique pour les Recherches de l'Université de Toulouse émet l'avis suivant : **Avis favorable.**

Nous délivrons le numéro IRB : N°: IRB00011835-2020-03-03-210 (Universite Federale de Toulouse IRB #1)

Nous restons à votre disposition pour toute question.

Les membres du bureau CER.
Pr Maria Teresa Munoz Sastre

Pr Jacques Py

Rémi Capa

M^a Teresa L

Université Fédérale Toulouse Midi-Pyrénées
41, allées Jules Guesde – CS 61321 – 31013 Toulouse CEDEX 6
www.univ-toulouse.fr

Figure A.5: Attentional modes Experiment approval

Université Fédérale
Toulouse Midi-Pyrénées

Toulouse le 19 juin 2020

A l'attention de CAUSSE Michaël

Objet : Avis de la commission du CER du 18 juin pour le projet 2020-265 : **Gaze-contingent multimodal warnings during a virtual reality piloting task**

Monsieur,

Compte tenu des éléments fournis dans votre demande d'examen Flash, le Comité d'Ethique pour les Recherches de l'Université de Toulouse émet l'avis suivant : **Avis favorable.**

Nous rappelons, par ailleurs, qu'il relève de la responsabilité des chercheurs de se conformer à leurs obligations légales notamment en ce qui concerne les aspects d'homologation du lieu de recherche ou RGPD : Règlement Général sur la Protection des Données.

Le numéro IRB attribué est : IRB00011835-2020-06-18-2020-265 (Universite Federale de Toulouse IRB #1) .

Nous restons à votre disposition pour toute question.

Les membres du bureau CER.

Pr Maria Teresa Munoz Sastre, Pr Jacques Py , Rémi Capa

Université Fédérale Toulouse Midi-Pyrénées
41, allées Jules Guesde – CS 61321 – 31013 Toulouse CEDEX 6
Tél. 05 61 14 80 10 - contact@univ-toulouse.fr
www.univ-toulouse.fr

Figure A.6: EyeVori Experiment approval

APPENDIX B

QUESTIONNAIRES AND SCALES

SITUATION AWARENESS RATING TECHNIQUE (SART; Taylor, 1990)

Instability of Situation
How changeable is the situation? Is the situation highly unstable and likely to change suddenly (High) or is it very stable and straightforward (Low)?

1 2 3 4 5 6 7

Complexity of Situation
How complicated is the situation? Is it complex with many interrelated components (High) or is it simple and straightforward (Low)?

1 2 3 4 5 6 7

Variability of Situation
How many variables are changing within the situation? Are there a large number of factors varying (High) or are there very few variables changing (Low)?

1 2 3 4 5 6 7

Arousal
How aroused are you in the situation? Are you alert and ready for activity (High) or do you have a low degree of alertness (Low)?

1 2 3 4 5 6 7

Concentration of Attention
How much are you concentrating on the situation? Are you concentrating on many aspects of the situation (High) or focussed on only one (Low)?

1 2 3 4 5 6 7

Division of Attention
How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (High) or focussed on only one (Low)?

1 2 3 4 5 6 7

Spare Mental Capacity
How much mental capacity do you have to spare in the situation? Do you have sufficient to attend to many variables (High) or nothing to spare at all (Low)?

1 2 3 4 5 6 7

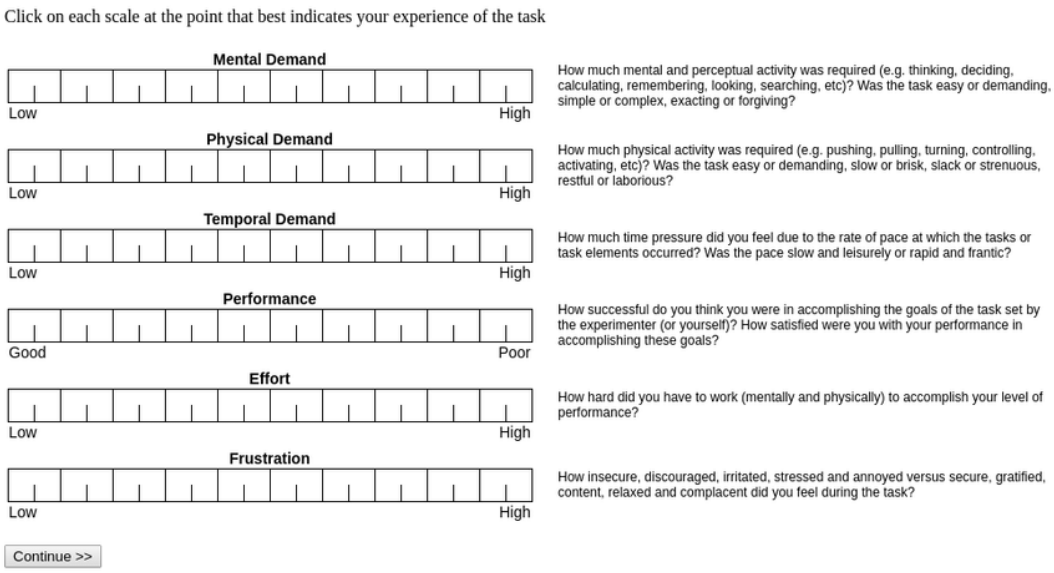
Information Quantity
How much information have you gained about the situation? Have you received and understood a great deal of knowledge (High) or very little (Low)?

1 2 3 4 5 6 7

Familiarity with Situation
How familiar are you with the situation? Do you have a great deal of relevant experience (High) or is it a new situation (Low)?

1 2 3 4 5 6 7

Figure B.1: Situation Awareness Rating Technique (Taylor, 2017)


Click on each scale at the point that best indicates your experience of the task

Mental Demand
Low High
How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Physical Demand
Low High
How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal Demand
Low High
How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Performance
Good Poor
How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Effort
Low High
How hard did you have to work (mentally and physically) to accomplish your level of performance?

Frustration
Low High
How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Continue >>

Figure B.2: NASA TLX workload questionnaire

Level	**Workload**	**Spare Capacity**	**Description**
5	Very High (Overload)	None	Behind on tasks; losing track of the full picture.
4	High (Fully Loaded)	Very Little	Non-essential tasks suffering. Could not work at this level very long.
3	Fair (Reasonable)	Some	All tasks well in hand. Busy but stimulating pace. Could keep going continuously at this level.
2	Low (Light Work)	Ample	More than enough time for all tasks. Active on ATC task less than 50% of the time.
1	Low (Light Work)	Very Much	Nothing to do. Rather boring.

Figure B.3: Workload Instaneous Self Assessment (ISA)(Tattersall and Foord, 1996)

AttrakDiff			
Sous-échelle	**Items (dans l'ordre de passation)**		
QP_1	Humain	○ ○ ○ ○ ○ ○ ○	Technique
QHI_1	M'isole	○ ○ ○ ○ ○ ○ ○	Me sociabilise
ATT_1	Plaisant	○ ○ ○ ○ ○ ○ ○	Déplaisant
QHS_1	Original	○ ○ ○ ○ ○ ○ ○	Conventionnel
QP_2	Simple	○ ○ ○ ○ ○ ○ ○	Compliqué
QHI_2	Professionnel	○ ○ ○ ○ ○ ○ ○	Amateur
ATT_2	Laid	○ ○ ○ ○ ○ ○ ○	Beau
QP_3	Pratique	○ ○ ○ ○ ○ ○ ○	Pas pratique
ATT_3	Agréable	○ ○ ○ ○ ○ ○ ○	Désagreable
QP_4	Fastidieux	○ ○ ○ ○ ○ ○ ○	Efficace
QHI_3	De bon goût	○ ○ ○ ○ ○ ○ ○	De mauvais goût
QP_5	Prévisible	○ ○ ○ ○ ○ ○ ○	Imprévisible
QHI_4	Bas de gamme	○ ○ ○ ○ ○ ○ ○	Haut de gamme
QHI_5	M'exclut	○ ○ ○ ○ ○ ○ ○	M'intègre
QHI_6	Me rapproche des autres	○ ○ ○ ○ ○ ○ ○	Me sépare des autres
QHI_7	Non présentable	○ ○ ○ ○ ○ ○ ○	Présentable
ATT_4	Rebutant	○ ○ ○ ○ ○ ○ ○	Attirant
QHS_2	Sans imagination	○ ○ ○ ○ ○ ○ ○	Créatif
ATT_5	Bon	○ ○ ○ ○ ○ ○ ○	Mauvais
QP_6	Confus	○ ○ ○ ○ ○ ○ ○	Clair
ATT_6	Repoussant	○ ○ ○ ○ ○ ○ ○	Attrayant
QHS_3	Audacieux	○ ○ ○ ○ ○ ○ ○	Prudent
QHS_4	Novateur	○ ○ ○ ○ ○ ○ ○	Conservateur
QHS_5	Ennuyeux	○ ○ ○ ○ ○ ○ ○	Captivant
QHS_6	Peu exigeant	○ ○ ○ ○ ○ ○ ○	Challenging
ATT_7	Motivant	○ ○ ○ ○ ○ ○ ○	Décourageant
QHS_7	Nouveau	○ ○ ○ ○ ○ ○ ○	Commun
QP_7	Incontrôlable	○ ○ ○ ○ ○ ○ ○	Maîtrisable

Note : Les items QP_1, ATT_1, QHS_1, QP_2, QHI_2, QP_3, ATT_3, QHI_3, QP_5, QHI_6, ATT_5, QHS_3, QHS_4, ATT_7 et QHS_7 sont inversés.

Figure B.4: Attrakdiff Questionnaire (Ux Experience)

VITA

July 3, 1993 .... Born - Marseille, France

July 2012 .... High School Diploma with major in Scientific, Gap, France

June 2015 .... B.S., Sport Science, University of Aix-Marseille, France

June 2016 .... M.D., Sport Science, University of Aix-Marseille, France

October 2017 .... M.S., Psychology, University of Aix-Marseille, France

November 2017 - Present .... PhD. Computer Science, ISAE-SUPAERO, Toulouse France

APPENDIX C

ABSTRACT

Résumé — Au cours d'un vol, les pilotes doivent surveiller de façon rigoureuse des instruments de vol spécifiques (e.g., indicateur d'attitude, vitesse, altimètre, les paramètres moteurs) ainsi que l'environnement extérieur (e.g., repérer des éléments du relief au sol notamment lors de conditions météorologiques dégagées et à basse altitude) dans le but de mettre à jour leur conscience de la situation. Cette activité de surveillance (monitoring en anglais), critique durant les phases de vols dites évolutives (e.g., décollage, phase d'approche, et atterrissage), tient compte de l'observation et de l'interprétation de la trajectoire, des modes d'automatisation sélectionnés, et des systèmes utilisés à bord. Cela suppose une comparaison en temps réel entre les données affichées aux instruments et les valeurs attendues lors des phases de vols. Une surveillance appropriée du cockpit permet de prendre des mesures correctives (e.g., ajuster la trajectoire de l'avion lors de la détection d'une déviation observable sur la zone d'attitude) en temps opportun lors de la déviation d'un paramètre, garantissant ainsi un niveau de sécurité optimal. Cette activité de surveillance est structurée en séquence d'engagement et de réorientation de l'attention visuelle du pilote d'un instrument vers un autre. Les rapports d'accidents ont démontré que bien souvent les erreurs de pilotage, tels que des trajectoires incorrectes ou bien une survitesse à l'atterrissage, étaient la résultante d'une surveillance défaillante et/ou inadéquate des instruments du cockpit. L'enjeu de ce travail de recherche est d''améliorer la sécurité des vols notamment grâce à l'intégration d'un oculomètre et/ou la recherche de solution pour améliorer l'entrainement des pilotes en vue de réduire les erreurs de surveillance à bord. Les mouvements des yeux sont une fenêtre sur l'état cognitif du pilote et permettent de révéler les chemins attentionnels empruntés par l'opérateur à travers son parcours visuel. En lien avec les problématiques de surveillance dans les cockpits, nous avons élaboré un assistant de vol (FETA : Flight Eye Tracking Assistant) basé sur des comportements visuels d'experts (e.g., 24 pilotes avec plus de 1600 heures de vols). Cet assistant prévient les pilotes, grâce à une alarme auditive, quand ces derniers ne consultent plus suffisamment un instrument de vol en comparaison avec la base de données des mouvements oculaires experts. Une évaluation facteurs humains de cet assistant a soulevé plusieurs problématiques et a ouvert la voie à de nouvelles recherches concernant notamment l'utilisation de métriques reflétant aux mieux les parcours oculaires dans le cockpit et permettant précisément de quantifier l'attention visuelle d'un pilote à bord. Une partie de ce travail de recherche s'appuie sur une comparaison entre novices et experts dans le but de quantifier la marque de l'expertise. Une méthode utilisant le K coefficient appliqué aux AOI a permis de qualifier l'attention visuelle des pilotes (focal vs ambient) au cours de scenario en simulateur de vols présentant différentes charges d'activité visuomoteur. Des méthodes d'apprentissage machine basée sur des matrices de transition ont permis de classifier l'expertise avec une précision de 91%. Enfin, deux méthodes ont été utilisés pour qualifier et quantifier les stratégies visuelles dans le cockpit. Une méthode utilisant la Complexité de Lempel-Ziv (LZC), un algorithme de compression des données, permettant de mettre en lumière la complexité des sequences de balayage dans le cockpit. Ainsi que le méthode N-gram, a l'origine issue de la recherche sur les séquences ADN, permettant de quantifier les patterns communs au groupe d'expert et la longueur des patterns utilisés. Ces contributions sont discutées à la lumière de l'amélioration d'un assistant basé sur des données oculométriques pour l'amélioration de l'apprentissage d'une part et pour éviter les problèmes de surveillances d'autre part. Finalement, l'évaluation du prototype FETA a soulevé des perspectives par rapport au choix de la modalité (e.g., auditive, visuelle, haptique) la plus pertinente concernant l'alerting.

Mots clés : Oculomètre, Facteur Humain, Strategies visuels, Interaction Homme-Machine, Neuroergonomie, Mouvements Oculaires.

Abstract — During a flight, pilots must rigorously monitor specific flight instruments (e.g., attitude indicator, airspeed, altimeter, engine parameters) as well as the external environment (e.g., locate terrain features on the ground, especially in clear weather conditions by low altitude) to update their situational awareness. This monitoring activity, which is critical during dynamic flight phases (e.g., takeoff, approach phase, and landing), consist in observing and interpreting the flight path, the selected automation modes, and the systems used onboard. This involves a real-time comparison between the data displayed on the instruments and the values expected during the flight phases. Appropriate monitoring of the cockpit enables to take corrective measures (e.g., adjust the aircraft's trajectory when a deviation is detected in the attitude zone) promptly when a parameter is deviated, thus guaranteeing an optimal level of safety. This monitoring activity is structured in a sequence of engagement and redirection of the operator's visual attention from one instrument to another. Moreover, accident reports have shown that piloting errors, such as incorrect trajectories or overspeed during landing, are often the result of inadequate monitoring of cockpit instruments. The purpose of this research work is to improve the flight safety thanks in particular to the integration of an eye-tracker. Eye movements are a window on the pilot's cognitive state and reveal the attentional paths taken by the operator through his visual path. In connection with cockpit monitoring issues, we have developed a Flight Eye Tracking Assistant (FETA) based on expert visual behaviors (e.g., 24 pilots with more than 1600 flight hours). This assistant warns the pilots, thanks to an audible alarm, when they no longer sufficiently consult a flight instrument in comparison with the expert eye movement database. A human factors evaluation of this assistant raised several issues with such an assistant and paved the way for further research including metrics that best reflect the eye paths in the cockpit and the need to find the right metric to quantify a pilot's visual attention onboard. Part of this research work is based on a comparison between novices and experts in order to quantify the mark of expertise. A method using the K coefficient applied to the AOIs allowed to qualify the visual attention of the pilots (focal vs ambient) during a flight simulator scenario with different loads of visuomotor activity. Machine learning methods based on transition matrices allowed to classify the expertise with an accuracy of 91%. Finally, two methods were used to qualify and quantify visual strategies in the cockpit. A method using Lempel-Ziv Complexity (LZC), a data compression algorithm, to highlight the complexity of the scanning sequences in the cockpit. Another called N-gram method, originally derived from DNA sequence research, which quantifies the patterns common to the expert group and the length of the patterns used. These contributions are discussed in the light of the improvement of a flying assistant based on eye tracking data for improving learning on the one hand and avoiding monitoring problems on the other. Finally, the evaluation of the FETA prototype raised perspectives on the choice of the most relevant modality (e.g. auditory, visual, haptic) for alerting.

Keywords: Eye-tracking, Human Factors, Visual Scanning, Human-Computer Interaction, Neuroergonomics, Eye movements.

Laboratoire de Neuroergonomie et de Facteur Humains, DCAS
Toulouse, France