Author(s): Hämäläinen, Raija; De Wever, Bram; Sipiläinen, Katriina; Heilala, Ville; Helovuo, Arto; Lehesvuori, Sami; Järvinen, Miitta; Helske, Jouni; Kärkkäinen, Tommi

Title: Using eye tracking to support professional learning in vision-intensive professions : a case of aviation pilots

Year: 2024

Version: Published version

Copyright: © 2024 the Authors

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Hämäläinen, R., De Wever, B., Sipiläinen, K., Heilala, V., Helovuo, A., Lehesvuori, S., Järvinen, M., Helske, J., & Kärkkäinen, T. (2024). Using eye tracking to support professional learning in vision-intensive professions : a case of aviation pilots. Education and Information Technologies, Early online. https://doi.org/10.1007/s10639-024-12814-9

Using eye tracking to support professional learning in vision-intensive professions: a case of aviation pilots

Raija Hämäläinen[1] · Bram De Wever[2] · Katriina Sipiläinen[1] · Ville Heilala[1] · Arto Helovuo[3] · Sami Lehesvuori[1] · Miitta Järvinen[1] · Jouni Helske[4] · Tommi Kärkkäinen[5]

Abstract

In an authentic flight simulator, the instructor is traditionally located behind the learner and is thus unable to observe the pilot's visual attention (i.e. gaze behaviour). The focus of this article is visual attention in relation to pilots' professional learning in an Airbus A320 Full Flight Simulator. For this purpose, we measured and analysed pilots' visual scanning behaviour during flight simulation-based training. Eye-tracking data were collected from the participants (N = 15 pilots in training) to objectively and non-intrusively study their visual attention behaviour. First, we derived and compared the visual scanning patterns. The descriptive statistics revealed the pilots' visual scanning paths and whether they followed the expected flight protocol. Second, we developed a procedure to automate the analysis. Specifically, a Hidden Markov model (HMM) was used to automatically capture the actual phases of pilots' visual scanning. The advantage of this technique is that it is not bound to manual assessment based on graphs or descriptive data. In addition, different scanning patterns can be revealed in authentic learning situations where gaze behaviour is not known in advance. Our results illustrate that HMM can provide a complementary approach to descriptive statistics. Implications for future research are discussed, including how artificial intelligence in education could benefit from the HMM approach.

Extended author information available on the last page of the article

1 Introduction

Simulations can enhance learning in multiple ways (Bergamo et al., 2022; Huang et al., 2022). In short, throughout the history of using computers to support education, simulations have been used to provide learners with a virtual environment in which to practice real-world scenarios, enabling them to rehearse their knowledge and skills in a safe setting (Kneebone, 2003). In aviation, the first ground-based flight simulator was introduced in 1909 (Krag, 2017). Currently, simulators are used to systematically train pilots in a wide range of scenarios (Salas et al., 1998), including critical moments of the flight, i.e. take-offs, landings, instrument approaches, and emergency procedures (Lekea et al., 2021; Villafaina et al., 2021). The primary benefit of aviation simulations is their ability to provide a safe training environment for making decisions, solving problems, and experimenting with different ways to operate (Harris et al., 2023). For example, pilots can practise complex manoeuvres and emergency procedures without the risk of damaging a real aircraft. Simulations are also used to replicate rare or uncommon scenarios, allowing pilots to gain experience (Myers et al., 2018). One advantage of simulations is that they provide immediate feedback to learners and allow them to see the consequences of their actions in real time (Jeffries, 2005). This feedback can be further used to help learners focus on areas where they need to learn new skills or improve their performance. Overall, simulations have been proven to be a valuable tool for enhancing professional learning and expertise development (see Lehtinen et al., 2020) by providing learners with a safe environment in which to practice real-world scenarios.

1.1 Professional learning in vision-intensive professions

Professional learning and expertise development refer to the development of the essential knowledge, skills, and attitudes for professionals to perform their jobs. In short, professional learning can be understood as learning processes that occur during work and learning that can be utilised for work (Gruber & Harteis, 2018; Lehtinen et al., 2020). In practice, it involves activities that help individuals learn and stay up to date within their fields (Vähäsantanen et al., 2017). Visual expertise is an essential part of professional learning and expertise development in vision-intensive professions, where professionals need to make decisions based on complex visual materials (e.g. Bellenkes et al., 1997; Jossberger, 2022; Lounis et al., 2021; Lu et al., 2020; Ziv, 2016). Visual expertise can be understood as a high level of proficiency and skill in a particular area of visual perception, such as recognising and interpreting visual patterns and/or identifying specific objects (Gegenfurtner et al., 2011; Lehtinen et al., 2020). For aviation pilots, visual expertise is crucial because one must be able to interpret and react to complex visual cues while flying. For instance, pilots must accurately navigate in changing weather conditions and make split-second decisions (Ahmadi et al., 2022). Without sufficient visual expertise, pilots risk making errors that could have disastrous consequences.

Pilots rely heavily on their visual senses to scan the flight path (subsequently referred to as the scan path), make decisions, and detect potential hazards in real time (see Ryffel et al., 2019). The specific areas of visual expertise are *visual focus of attention*, *attention allocation* (subsequently focus of attention), and *visual scanning* (e.g. Bellenkes et al., 1997; Chaudhuri et al., 2022; Lounis et al., 2021; Lu et al., 2020; Jin et al., 2021). The focus of attention (Chaudhuri et al., 2022) refers to the selective and conscious concentration of a pilot's mental resources on a particular stimulus, object, or task. This is a cognitive process that involves the ability to concentrate on what is most important at any given moment and filter out irrelevant information. The pilot's focus of attention is defined as the pilot's gaze on relevant targets (e.g. air traffic control and weather information) and/or certain gaze paths for processing information (see also Chaudhuri et al., 2022; Palazzi et al., 2018), and this can be measured by using eye tracking and gaze/fixation starting and ending times in the designated areas of interest (AOI).

Visual scanning involves shifting one's attention between AOIs via scan paths (see Kang et al., 2014). During visual scanning, pilots depend on various instruments and systems to assist them during flying. These include the aircraft's instruments, navigation aids, and communications systems as well as external sources, such as air traffic control and weather information. For example, during take-off and landing, pilots use visual scanning to ensure the aircraft is on the correct runway, aligned with the centreline, and at the correct altitude and speed. Thus, they are switching their attention between the outside environment (i.e. looking out of the window (OTW) and cockpit panels and indicators, such as the primary flight display (PFD, indicating artificial horizon, speed, altitude and flight path), navigation display (ND, indicating map and lateral flight path), flight control unit (FCU, including controls for speed, heading, altitude and autoflight system), and other cockpit panels. Overall, visual scanning is a critical skill for pilots that needs to be learned to ensure their ability to safely and efficiently operate an aircraft. Behrend and Dehais (2020) showed that the pilot role assignment (whether a pilot is flying or monitoring) plays an important role in visual scanning and decision-making behaviour.

When training pilots, it is essential to teach visual scanning, that is, how to look at the right place at the right time (correct scan paths). First, pilots need to learn to correctly focus their attention, be aware of their surroundings, and be able to scan, for example, the external environment for potential hazards. Second, they need to focus on the right instruments at the right time to ensure they are taking the correct actions and flying the plane safely. In relation to the above requirements, many studies highlight the importance of experience and the role of the level of practice on visual scanning behaviour (Gao & Wang, 2024; Haslbeck & Zhang, 2017). More experienced pilots can distribute their attention between instruments and exterior, whereas novices may focus more on the instruments (Lijing & Lin, 2016). The visual scanning patterns of more experienced pilots can be more flexible and adaptable, which can increase caution in unexpected flight scenarios as reflected, e.g., in go-around decision making (Dehais et al., 2017; Gao & Wang, 2024). Jin et al. (2021) found that experienced pilots were better at allocating their attention between relevant flight instruments and indicators. In addition to experience, changes in the

visual scanning patterns naturally reflect the state of flying and the complexity of the task at hand (Diaz-Piedra et al., 2019). Therefore, in practice, pilots' visual expertise is gained through training, experience, and exposure to different stimuli.

So far, the pedagogical problem has been that in an authentic flight simulator, the instructor is behind the learner and unable to observe the learner's gaze behaviour because of space constraints (Fig. 1, right). This makes it difficult or even impossible for the instructor to provide timely feedback on the learner's visual scanning. Our study seeks to investigate how advancements in eye-tracking technology could be applied to respond to this challenge (see also Lounis et al., 2021; Niehorster et al., 2020; Ryffel et al., 2019).

1.2 Eye tracking as a tool to study and support professional learning

Eye-tracking technology makes it possible to accurately and non-invasively track where a person is looking (or not looking) and for how long (Kok & Jarodzka, 2017; Van Gog & Jarodzka, 2013; Ziv, 2016). This information can reveal learners' information processing related to their attention, perception, and decision-making (Lai et al., 2013; Rappa et al., 2019). For example, in the context of pilot training, eye-tracking has been acknowledged as one of the pivotal technological methods to improve aviation safety by incorporating learning (Jin et al., 2021; Mengtao et al., 2023). Eye-tracking allows one to objectively trace the focus of attention processes (including scan paths) without interrupting the exercise. Tracking the movement of learners' eyes and measuring fixations and fixation durations provide insights into the learner's attention and how their gaze patterns may relate to decision-making (Holmqvist et al., 2011; Ziv, 2016). Despite this potential, studies have identified several concerns related to enhancing flying performance, such as a lack of training interventions (Ziv, 2016). For the future, the challenge is how gaze dynamics can be used to infer pilots' attention and how this information can be used a) *to develop pedagogical approaches* to support visual

Fig. 1 Flight academy's Airbus A320 Full Flight Simulator

scanning and b) *to personalise instructions* for different pilots (e.g. Rainieri et al., 2021; Sullivan et al., 2011), both of which can be accomplished either technologically or pedagogically, as further elaborated in the discussion section.

In terms of the pedagogical design, one of the most crucial issues is the balance between activities designed before (to do this we have to understand the learning needs of the pilots at the general level in advance) and during the learning process (to do this we have to capture the individual learning processes of each pilot). It has been argued that the combination of learning analytics and real-time eye tracking has the potential to meet these challenges in evaluating user models of subject content (e.g., Wang et al., 2021). For example, analysis of the visual scanning of professionals can provide an overview of the optimal performance and behaviour in a specific task, which can then be used as a baseline for assessing students' individual performance. Furthermore, combining the optimal baseline and students' individual performance could provide information supporting personalised learning. Capturing and automatically analysing gaze data could provide valuable insights into how learners utilise their visual scanning techniques and support the provision of objective, individual, and fine-grained feedback even during the simulation (see Ryffel et al., 2019; Rudi et al., 2020; Ahmadi et al., 2022). This kind of information about learners' behaviour and scanning processes can also increase trainees' self-awareness and help them evaluate their techniques and decision-making during the simulation training (see Muehlethaler & Knecht, 2016).

Currently, eye-tracking techniques are becoming more broadly available for learning settings, and there is a growing interest in capturing professional learners' information processing through eye tracking and relating it to their attention, perception, and decision-making. Relatedly, there is considerable hype regarding the possibilities of current technological advancements. Therefore, this development calls for a better understanding of many technological and pedagogical questions: How can learners' information processing be captured through their eyes in vision-intensive professions? How can information processing be visualised in pedagogically meaningful ways? What needs to be understood to design simulations that provide feedback for instructors and learners automatically, even in real time?

We use eye tracking and gaze modelling to study the visual scanning of pilots. Based on international aviation guidelines, there is an expected protocol regarding where pilots should be focusing and how they should shift their attention (subsequently referred to as the dynamic trace of attention) (Wickens et al., 2001; Ziv, 2016; EASA, 2020). Gathering and processing visual information in a dynamic aircraft environment is an essential aspect of pilots' competencies (e.g., EASA, 2020). In essence, the pilots are required to learn "when to look where to assure that the dynamic processes are under control" (Wickens et al., 2001, p. 1). First, we use descriptive statistics to examine pilots' gaze dynamics in terms of fixations and fixation durations and whether this behaviour complies with the protocols of different landing scenarios that the pilots should follow. Second, we aim to automate our analyses by using hidden Markov model (HMM) to identify the visual scanning behaviour of the pilots in order to provide that information to instructors later. The following two research questions guide our study:

RQ1: How do pilots visually scan the indicators and the window during landing scenarios in a flight simulator?

RQ2: How can we automatically detect whether pilots demonstrate the expected visual scanning behaviour during landing scenarios?

2 Materials and methods

2.1 Participants and experimental design

At the core of the experimental design and the data collection are flight scenarios performed in an Airbus A320 Full Flight Simulator environment (see Fig. 1). Simulator training is a mandatory part of the professional learning for all commercial pilots. The flight scenarios were authentic in the sense that they simulated real working scenarios that do not differ from scenarios practised in the simulator training or situations that pilots face in their work. In this context, understanding the dynamics of attention, especially during takeoff and landing, is crucial. In this study, we focus on the landing phase of each scenario, specifically, the three stages of three landing scenarios where pilots must follow a specific visual scanning strategy. A full-flight simulator realistically mimics cockpit instruments and provides a visual representation of the environment. Participants conducted a one-hour simulation, covering three scenarios requiring manual control of the aircraft during approach and landing based on visual cues, highlighting the importance of such scenarios in pilot training. The scenarios were: a calm wind approach, an approach with significant crosswind challenging the maintenance of the correct flight path, and an approach with a technical system failure. The final approach and landing, requiring active scanning of visual cues, particularly in poor visibility, were the study's focus. This phase includes a "transition to visual," where pilots shift their scanning from flight instruments to external cues, such as runway lights, at a determined decision altitude. Below this altitude, pilots are expected to change their scanning pattern and land using external cues, shifting their attention gradually to these cues. Hence, our analysis divides the final approach's visual scanning into three stages (see below), where the scanning strategy is anticipated to change.

The participants were (already graduated) Airbus A320 pilots, and the data consisted of eye-tracking data from 43 landing missions conducted by 15 participants. The time of landing scenarios ranged from 31.0 s to 40.7 s (M = 36.4 s, SD = 2.2 s). The stages times of the landing scenarios were as follows: The first stage ranged from 2.0 to 4.4 s (M = 3.1 s, SD = 0.6 s), the second stage ranged from 12.6 to 16.8 s (M = 14.8 s, SD = 1.0 s), and the third stage ranged from 14.1 s to 23.0 s (M = 18.6 s, SD = 2.1 s).

Ethical approval for the research was granted by The Human Sciences Ethics Committee of the University of Jyväskylä, and the research was conducted following the guidelines of the World Medical Association's (WMA) Declaration of Helsinki (each participant provided written consent to participate in the study). A data availability statement: Due The Human Sciences Ethics Committee of the University of Jyväskylä decision 755/13.00.04.00/2020 data can not be made available.

We used eye tracking to record how the pilot flying (PF) observed their working environment and how they gathered information from the instruments, such as the PFD and OTW view of the simulator. Eye-tracking data were collected from the participants during the scenarios using the Ergoneers Dikablis Glasses 3 (sampling rate 60 Hz). The main AOIs in this study relevant for the visual scanning during the approach were the PFD and the OTW views (cf. Faulhaber et al., 2022). PFD is the main source of information during the approach, and it shows information like the altitude, speed, and turn coordinator. In this study the PFD was analysed as a whole, i.e. as a one AOI. The OTW view is essential in the later stage of the approach and during landing, when the runway can be seen visually. We also captured possible glances to ND and FCU. Figure 3a–c depict the compiled multi-video and superimposed eye-tracking views of the experiment, analyzed AOIs, and different stages of the expected visual scanning pattern.

For the descriptive statistics (RQ1), we focused on visual scanning behaviours around predetermined auditory callouts that are made in the aircraft during a typical landing.

An ideal visual scanning pattern (see Fig. 2) includes three phases and should roughly match the three flight stages:

Stage 1: After the first automated callout 'Hundred above'(100 feet above the decision altitude) and before the actual landing procedure starts, following 'Approach lights' (verbal confirmation of visual contact with external cues), the pilot should still observe and monitor the PFD and other indicators (Figs. 2 and 3a). The 'Hundred above' callout is a safety measure and a reminder that the pilot is approaching the decision altitude at which the pilot must decide whether to continue the approach to landing or execute a missed approach.
Stage 2: Between the callouts 'Approach lights' (latest at the decision altitude) and at 'One hundred', the transition to primarily using external visual cues begins, and the pilot should start to switch their attention between the PFD and OTW (Figs. 2 and 3b). The callout 'One hundred' refers to the altitude being 100 feet above the ground during the final approach phase. This is a critical phase of flight, and the callout serves to alert the flight crew that they are nearing the runway and should be in the final stages of preparing for landing.
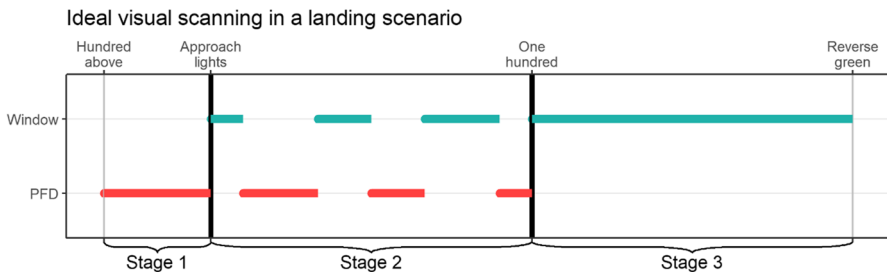

Fig. 2 Illustration of the three stages during a landing scenario, based on the four announcements made, and a representation of the ideal visual scanning behaviour that pilots are expected to follow

(a)

(b)

(c)

Fig. 3 **a** In the first stage, after the first callout 'Hundred above,' the PF should primarily monitor the PFD. **b** In the second, after the callout 'Approach lights,' the PF should switch attention between the PFD and OTW. **c** In the third stage, after the callout 'One hundred', the PF should only be looking at the OTW

Stage 3: In the last part between 'One hundred' and 'Reverse green', the pilots should only be looking at the OTW (Figs. 2 and 3c) in order to land the aircraft at the optimal touch down point on the runway. The callout 'Reverse green' indicates that the aircraft is on the ground, thrust reversers are active and working as expected. Thrust reversers are used to help slow the aircraft down after landing, reducing their reliance on wheel brakes and allowing for shorter landing distances.

For the second research question (RQ2), the same eye-tracking data were used, but the actual visual scanning phases of the pilots were modelled in an automatic way using an HMM. In this study, an HMM was used to automatically detect the different phases of the landing procedure. In line with the methodological terminology

related to HMMs, we will use the term 'states' to refer to the steps that can be automatically detected in the eye-tracking data. We will continue to use the term 'stages' for the pre-identified steps and 'phases ' when referring to the actual visual scanning behaviour of the pilots. In short, stages are expected to take place based on the predetermined announcements, and, in answering RQ1, we need to know precisely when those announcements are made in order to describe how well the pilots' visual scanning is matching these announcements. Regarding RQ2, the automatic HMM analysis is not restricted to the announcements. Rather, we identify the different states from the data independently from the announcements.

2.2 Data analysis

To answer RQ1 regarding how pilots visually scan the indicators and the window during landing scenarios in the flight simulator, the percentage of time focused on the PFD and OTW was calculated for each landing stage. In stage two, the number of times the pilot switched attention between the PFD and OTW was counted. With these descriptive statistics, we can see if the pilot is monitoring the landing scenario as we would expect based on the four announcements. However, this is a manual and time-consuming process for the flight instructor. Thus, we also present one potential method for automatically detecting the pilot's visual scanning accuracy and consistency, that is, when the pilot is and is not monitoring as expected (RQ2).

To answer RQ2 regarding how to automatically detect whether pilots demonstrate the expected visual scanning behaviour during landing scenarios, we used an HMM. The HMM approach provides an unsupervised, model based method for analysing unlabeled data, meaning that while we do know the exact time points when different flight stages start and end, we do not know when the pilots switch their visual scanning patterns to different phases. The fact we do not know the actual change points in the scanning patterns means that we cannot use supervised classification methods such as decision trees, and developing ad-hoc rules for the labelling is also not trivial. In addition and in contrast to ad-hoc methods, HMMs also provide automatic uncertainty estimates for the latent state probabilities, which allows coherent assessment of the labelling accuracy. HMMs are also more robust to outliers and measurement noise compared to methods such as observational level Markov models. For example, the number of observed states (e.g., observed AOIs) can be larger than the assumed true number of distinct gazing patterns, and occasional record of single AOI within a long sequence of other AOIs can be automatically "smoothed out" so that it does not necessarily impact the latent state trajectory.

Specifically, to analyse the scanning of the indicators and the window based on the input messages of the simulator and the aviation protocol the pilots are taught, we utilised an HMM to analyse eye-tracking data. In general, Markov models are used to model stochastic processes, where the probability of the next state depends on the previous states. In HMMs, this Markov process is not observed directly but rather through observations, which depend on this latent process (e.g. Helske et al., 2018; Rabiner, 1989). To date, HMMs have been used to analyse, for example, fixation sequences during visual inspection in a home appliance facility (Ulutas

et al., 2020), eye-tracking data of air traffic control officers (Muthumanickam et al., 2019), sustained attention in visual object tracking (Kim et al., 2020), complex life sequence data (Helske et al., 2018), and productivity in open-ended inquiry (Tissenbaum, 2020).

The sequences of fixations to AOIs (i.e. PFD, OTW) were extracted using Ergoneers D-Lab V. 3.55. From D-lab the AOIs name, its starting and ending time and its duration in milliseconds was extracted. The pre-processing of the sequences involved discretisation, which means that the landing scenarios of the pilots, which usually last around 30 to 40 s in total, were divided into time intervals of 150 ms. We selected this interval based on existing research on detecting fixations in a similar context (e.g. Allsop et al., 2016; Chen & Tsai, 2015; Li et al., 2014; Negi & Mitra, 2020; Udale et al., 2022; Wu et al., 2019). Each of the time intervals was coded with the AOI that the participant was looking at the most during the interval. One pilot quickly glanced at the ND during the third landing (2 times, 2.5% of the landing time). As ND is right next to the PFD and only one pilot looked at it quickly and as it is also acceptable to sometimes look at the ND during a landing, it was coded as a PFD in the analyses. This AOI sequence of a pilot was then used as an input sequence to the HMM. The HMMs were estimated via maximum likelihood in the R environment (v4.2.1; R Core Team, 2022) using the seqHMM package (Helske & Helske, 2019).

2.2.1 Hidden Markov model

We used a discrete-time HMM as a probabilistic method to gain information about the latent visual scanning patterns of the pilots. The HMM, a type of dynamic Bayesian network, consists of two state sequences: one observed and one so-called 'hidden', or latent, state sequence. With the HMM, it is possible to find latent states that might not be directly visible in the data. In the first-order Markov model, the next state's probability distribution is predicted by the preceding state. In an HMM, the observed state's outcomes are predicted by the hidden states, and the next hidden states are predicted by the previous hidden states. Thus, while we do not observe the hidden state sequence, we can estimate it based on the observed state sequence, which depends on the hidden states. We will present a brief overview of the HMM, but for a more comprehensive explanation, see, for example, Rabiner (1989).

Let $(O_t)_{t=1,...,T}$ denote an observed state sequence of length T and $(X_t)_{t=1,...,T}$ denote a hidden state sequence of the same length T. The observed state sequence can obtain an observed state from the set $V = \{v_1, ..., v_M\}$, which has M states. Similarly, the hidden state sequence can obtain a hidden state from the set $S = \{s_1, ..., s_N\}$ that has N states. To model the HMM, three more parameters are also needed: a transition matrix A, an emission matrix B, and an initial probability vector π.

Now the $N \times N$ transition matrix $A = \{a_{ij} \mid a_{ij} = P(X_t = s_j \mid X_{t-1} = s_i)\}$ is the probability of moving from hidden state s_i at time $t-1$ to hidden state s_j at time t. The NxM emission matrix $B = \{b_i(k) \mid b_i(k) = P(O_t = v_k \mid X_t = s_i)\}$ is the probability that the observed state $O_t = v_k$, given that the hidden state $X_t = s_i$ at time t. An initial probability vector $\pi = \{\pi_i \mid \pi_i = P(X_1 = s_i)\}$ is the probability that the hidden sequence starts with a hidden state s_i. We denote the HMM parameters as $\mu = (A, B, \pi)$.

To efficiently compute the probability of $P(O|\mu)$ (likelihood of the model), we used a forward–backward algorithm (Baum & Petrie, 1966; Rabiner, 1989). To find the parameter set μ that maximises $P(O|\mu)$, we used the Baum–Welch algorithm, a special case of the expectation–maximisation (EM) algorithm (Dempster et al., 1977; Rabiner, 1989). It is an iterative procedure that finds the best estimate of μ that locally maximises $P(O|\mu)$. When the HMM parameters have been estimated, the next step is often to find the optimal (most likely) hidden state sequence for the given observed sequence, which is found by using the Viterbi algorithm (Rabiner, 1989; Viterbi, 1967). Alternatively, output of the forward–backward algorithm can be used to compute the posterior probabilities of a specific state at a specific time point given the observed data and the model.

In a so-called fully connected HMM, the states can transition between themselves without restrictions. There is also a possibility to restrict the transitions. In this study, we used a so-called left-to-right HMM (LRHMM), in which the transitions between the states can either stay the same or move to a higher index (Rabiner, 1989). We restricted the transitions so that the states could only move one index higher or stay the same. Especially in this case, the posterior probabilities provide a straightforward way to assess the uncertainty in the timing of state transitions.

3 Results

In the following section, we describe where the pilots are looking and for how long during the three stages (RQ1) of the three landing scenarios. We describe if and how pilots visually scan the indicators and OTW in terms of temporal accuracy and consistency in relation to what we would expect based on the predetermined announcements of the flight simulator and the protocol. Then, in Section 3.2, we model the pilots' visual scanning behaviour using an HMM. We show that the HMM can provide complementary information to descriptive statistics. Specifically, the actual visual scanning behaviour (RQ2) of the pilots can be modelled in an automatic way through modelled states.

3.1 Descriptive statistics on visual scanning

3.1.1 Following the optimal scanning behaviour in three predefined stages of landing scenarios

In general, the predetermined auditory announcements define the behavioural prompts that should guide pilots' performance. It is essential for the pilots to start and end the visual scanning behaviour at a correct point in time based on the announcements (see Fig. 2). Overall the pilots looked at the PFD (34.5%), OTW (65.2%) and FCU (0.3%) during all the landings. Five pilots looked at the FCU in seven different landings. Using gaze behaviour data and the overall time pilots used to direct their visual attention to the PFD and OTW, we first examined the start and end of the visual scanning sequence in relation to the predetermined announcements.
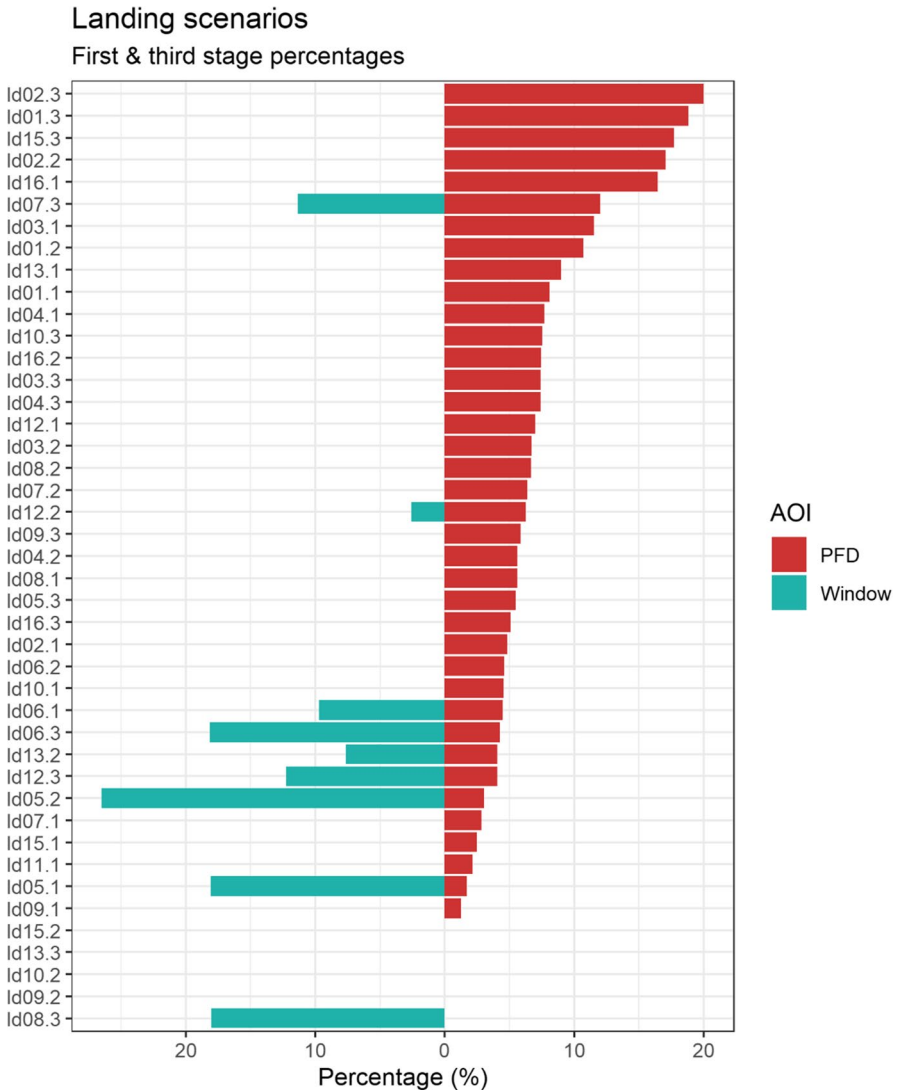
Fig. 4 Temporal precision of timing with respect to the predetermined announcements. Percentage of time (x-axis) that pilots (Indicated by Idxx number on the Y-axis) are looking at the OTW during the *first stage* (left side; the green horizontal bars; we expect 0% here, as the focus should be on the indicators) and at the indicators during the *third stage* (right side; the red horizontal bars; we also expect 0% here, as during this third stage the focus needs to be on the OTW) of the three (Idxx.1, Idxx.2, Idxx.3) landing scenarios. Note that the first stage times range from 2.0 s to 4.4 s (M=3.1 s, SD=0.6 s), and the third stage times range from 14.1 s to 23.0 s (M=18.6 s, SD=2.1 s)

Figure 4 presents the percentages of time that the pilots are looking through the windows during the first stage of the landing scenario. As we can see in Fig. 4, in nine (21%) landing scenarios pilots are looking at the window too early. For instance, in the second landing scenario of pilot ID05, the pilot is looking over 25% of the time

at the window during this first stage. Overall, six pilots show a premature gaze to the OTW before the second announcement. Further, Fig. 4 shows whether the sequence was started too early in the first stage (i.e. before the second announcement) and if it ended too late in the third stage (i.e. after the third announcement) (see description of the third stage in this article) by showing the share of overlap of the 'unexpected' AOI with the duration of the actual stage.

In the second stage, that is, in the middle of the landing scenario between the 'Approach lights' and 'One hundred' announcements (see, Fig. 2), pilots are supposed to divide their attention between the window and the indicator by looking back and forth. Figure 5 presents the percentages pilots are looking at the indicators during the second stage of the landing scenarios along with the number of transitions between the indicators and the window. Since the pilots are taught to look roughly half of the time at the indicators and the other half at the window, a dashed line at 50% was added to the figure. Vertical lines at 35% and 65% illustrate the variation in the pilots' scanning patterns which, based on discussions with the flight instructors, was still considered to be within acceptable limits for the second stage. In addition, pilots should shift their attention a couple of times, meaning that for the time looking at the specific AOI, we also considered the number of shifts between the two AOIs. In sixteen (37%) landings, the pilots are looking at the indicators too long, and in five (12%) landings, the pilots are looking at the window too long. In twenty-two (51%) landings, the pilots are monitoring as taught in the second stage: half of the time at the indicators and shifting fairly often between the AOIs. The shifting mean was 8.5 (SD=2.5).

During the last stage, the visual scanning is supposed to be the opposite of the first stage. Specifically, in the *third stage*, pilots are supposed to focus mainly on the outside environment, that is, to look through the window and not at the instrument panel. As we can observe in Fig. 4, which presents the percentages of time pilots are looking at the indicators, most pilots are still checking the indicators. The pilots are only looking purely out the window as taught in five (12%) of the landing scenarios. In the other landings, during which they are looking at indicators, they are looking at the indicators less than 5% of the time in 13 (30%) of the landing scenarios. In the other 25 (58%) cases, the pilots are looking at the indicators more than 5% of the time. For instance, pilot ID02 is looking at the indicators more than 20% of the time during the third stage of the third landing scenario (see Appendix 1). The pilots exhibited a general tendency to continue scanning the PFD after the final announcement.

3.1.2 Two descriptive examples

Next, we shall further illuminate the landing scenario with AOIs in two selected cases, pilot ID08 and ID10. Figure 6 shows the pilots' (ID08) AOIs in three different landing scenarios. On the colour dot, the pilot starts looking at the AOI, and the colour line continuing from it indicates how long the pilot keeps looking at the AOI. The first grey vertical line indicates the first announcement, 'Hundred above', when the landing scenario starts, and the second grey vertical line indicates the fourth (last) announcement, 'Reverse green', when the landing scenario ends. The
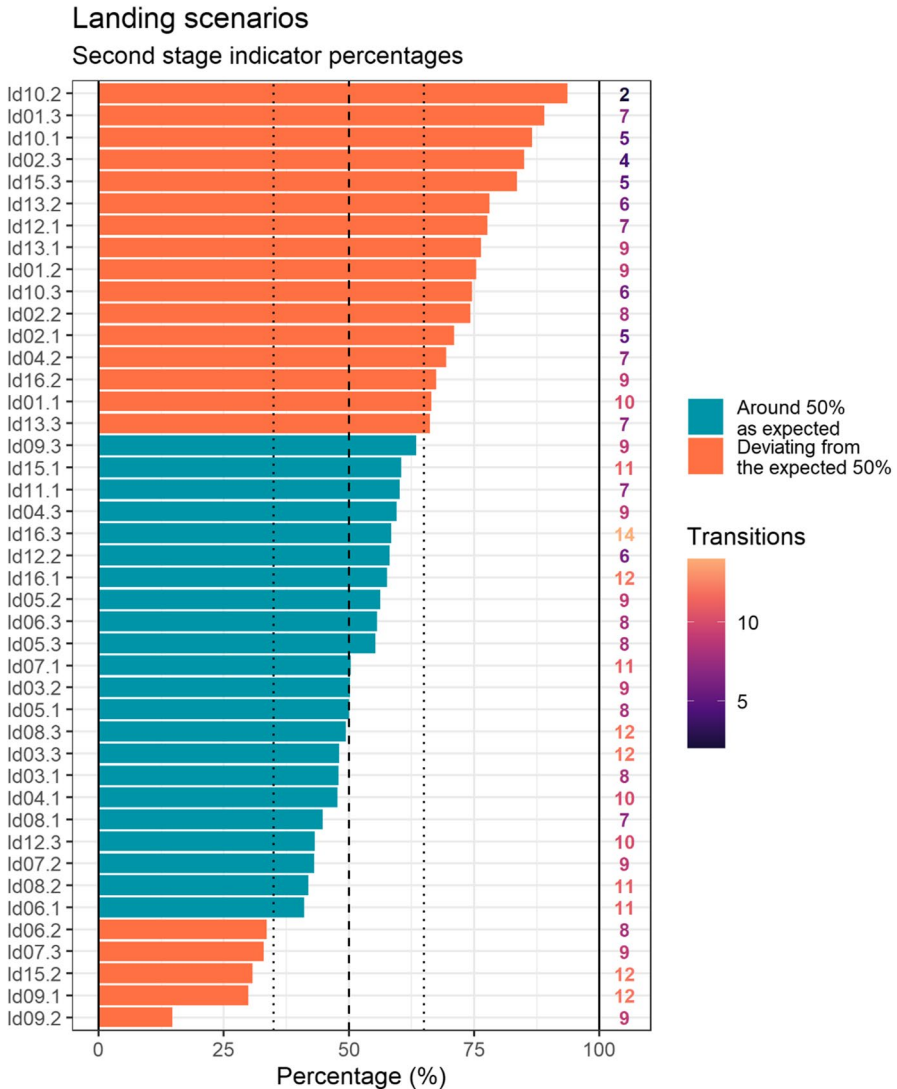
Fig. 5 Visual scanning behaviour dynamics based on gaze transitions and division of attention between the PFD and OTW. Percentage of time (x-axis) that pilots (Indicated by Idxx number on the Y-axis) are looking at the indicators during the *second stage* of the three (Idxx.1, Idxx.2, Idxx.3) landing scenarios. In this figure we expect the percentage to be around 50% (indicated by the dashed vertical line). All landing scenarios in which this 50% was approached (i.e. between 35 and 65%, indicated by the two vertical dotted lines) are coloured in green, while the other cases (i.e. deviating from the expected 50%) are coloured in orange. At the end of the bar, the number of switches (i.e. looking back and forth at the window and indicators) are listed

first black vertical line indicates the second announcement, 'Approach lights', when the first stage ends and the second stage starts, and the second black vertical line indicates the third announcement, 'One hundred', when the second stage ends and
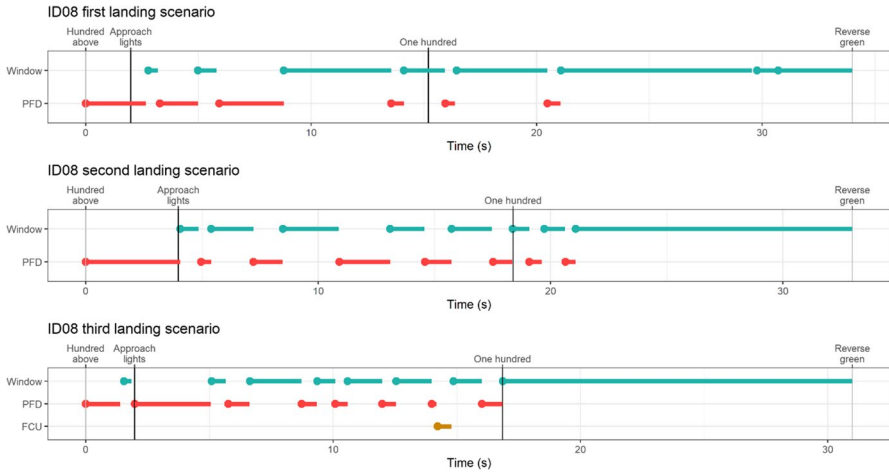
Fig. 6 Pilot ID08's three landing scenarios. The vertical axis shows the observed AOIs that the pilot is looking at during the landings, and the horizontal axis shows the time in seconds. At the coloured dot, the pilot starts looking at the AOI, and the line continuing from it indicates how long the pilot keeps looking at the AOI. The two grey vertical lines indicate the announcements 'Hundred above' and 'Reverse green', respectively, and the two black vertical lines indicate the announcements 'Approach lights' and 'One hundred', respectively

the third stage starts. We can see that in the first stage, the pilot mainly looks only at the PFD ('Hundred above' – 'Approach lights'), in the second stage, the pilot starts to shift between the indicators and the window and in the last stage, the pilot mainly looks only at the window.

Figure 7 shows the pilot's (ID10) AOIs in three different landing scenarios. The pilot looks only at the indicators in the first stage during all of the landings and
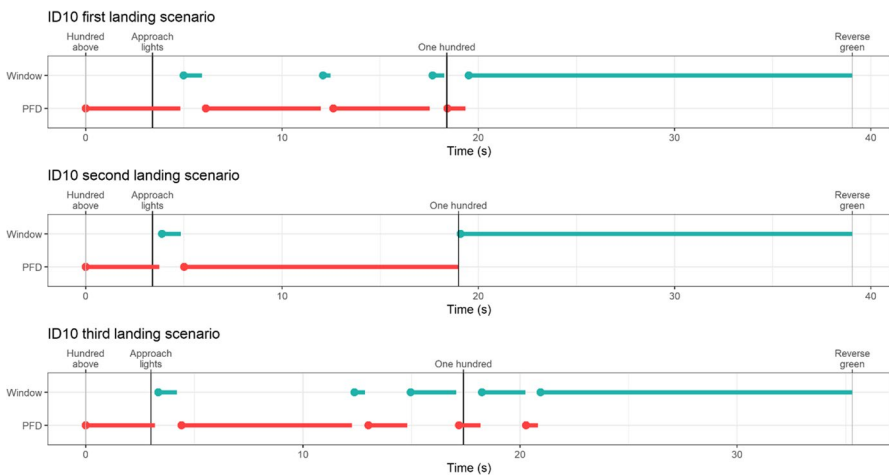

Fig. 7 Three landing scenarios of pilot ID10 (see Fig. 6 for details on the interpretation of this figure)

mainly outside in the third stage. However, in the second stage, pilot ID10 does not shift attention between the indicators and window very often. Especially in the second landing scenario, the pilot mainly looks at the PFD ($>90\%$) in the second stage and shifts attention only three times throughout the whole landing scenario. Additionally, in the third landing, the pilot starts to shift attention much later in the second stage; this shift occurs at the end of second stage and at the beginning of the third stage – too late relative to the taught visual monitoring.

Finally, when considering the individual differences of the pilots, the analysis revealed that, for example, pilots ID08 and ID03 showed an accurate start of the scanning sequence in most of the landings and relatively consistent visual scanning behaviour based on their attention distribution and number of gaze transitions (see Appendix 1 for a detailed overview of all pilots). Similarly, pilots ID05 and ID06 showed relatively consistent gaze behaviour but also a tendency to start the visual scanning slightly earlier than was expected based on the second announcement (Appendix 1). However, several landings showed inconsistent scanning behaviour and delayed ending capered what was expected (e.g. ID02, ID10).

In sum, the descriptive statistics show how the pilots are monitoring the instruments and the window and whether they are looking at the areas we expect them to look at, given the specific announcements in the flight simulator (see Fig. 2). However, the disadvantage of investigating the descriptive figures is that in an authentic learning situation, the flight instructor would have to look at each participant separately, which is a time-consuming manual process that is prone to errors. Furthermore, we need to know these stages in advance (e.g. based on predetermined announcements). Therefore, more automated options are needed to provide instructors with an overview of the visual scanning behaviours of all pilots, which will also allow identifying outliers, requesting details on their scanning behaviour, and investigating why they show different visual scanning behaviours. To respond to this need, we will next present a potential modelling approach for detecting accuracy and consistency in visual scanning behaviour.

3.2 Detecting visual scanning behaviour automatically

An HMM model was used to detect different states of visual scanning. In the following, we demonstrate that a modelling approach using an HMM can provide complementary information to descriptive statistics based on gaze behaviour. In short, the descriptives and figures of the 'raw' AOI data can contain considerable noise and mislabelling for various reasons (Muthumanickam et al., 2019). Therefore, with a large number of AOIs and time points, it can be challenging to interpret the results accurately. HMMs can help by providing 'smoothed' estimates that provide a clearer picture of what is happening. In our descriptive results presented above, the three stages are defined by the announcements. However, with HMM models, the model estimates when the three phases are actually taking place, based on the actual visual scanning of the pilots. For announcements, the points are fixed, and they can be used to check whether the pilot has monitored exactly as taught. But, with the help of an HMM, where the monitoring of this pilot starts and ends can be made visible. This

makes it possible to compare the actual monitoring with the most optimal monitoring based on the announcements. Thus, HMM models may help to pinpoint learners' monitoring patterns more easily.

First, we introduce the landing scenarios and the LRHMM that is used to estimate the three actual phases for all pilots individually. Second, examples of the individual pilots' results based on the HMM will be presented. In the future, this kind of information could be useful for personalising instructions for different pilots. Finally, we conclude our results with a heatmap (Fig. 10) of the pilots' visual scanning sequences, which illustrates the overall picture of all the landings. This information can be useful in developing pedagogical approaches to support professional learning.

3.2.1 Landing scenarios and LRHMM

We use the LRHMM to estimate the states of the pilots during the landing scenarios. With this model, we can estimate the three states for all pilots individually. We use the same model for all of the pilots in every landing scenario. As there are three stages taught to pilots during the landing scenarios, we give three states to the LRHMM as well. With the model and these three hidden states, our aim is to try to identify the actual phases corresponding to the pilot's observation. To find the second phase in the landing scenario, where the pilot switches between the indicators and window, we fix the second state's emission probabilities. Based on our a priori expectation that the pilots should monitor the indicators and window equally and the descriptive statistics from the observed data, which indicate that the pilots are looking at the PFD, on average, 58.2% of the time in the second stage, at the FCU 0.4% and at the window 41.4% of the time, we fix the second state probabilities to 59% (PFD), 40% (OTW) and 1% (FCU). See Appendix 3 for estimated transition and emission probabilities.

3.2.2 Individual-level results for all landing scenarios

Next, we briefly introduce all of the landing scenarios of all pilots. In Appendix 2, all landing scenarios of all pilots are visualised in graphs. These graphs provide information regarding where and how the pilots are looking throughout the landing scenarios and thus provide an in-depth view for further analysis and feedback from instructors. The data can be used to provide overall information regarding the pilots visual scanning, as illustrated above (see Figs. 2, 3, 4 and 5). However, the detailed graphs in Appendix 2 could also be presented using a type of dashboard to the flight instructor or even to the pilots, allowing for a more in-depth analysis of individual visual scanning paths. For instance, looking at the three landing scenarios of pilot ID08 (Figs. 6 and 8 and row 8 in Appendix 2), we can observe the visual scanning throughout all three phases of the three landing scenarios. On the colour dot, the pilot starts looking at the AOI, and the colour line continuing from it indicates how long the pilot continues to look at the AOI. The first black vertical line indicates the second announcement, 'Approach lights', and the second black vertical line indicates the third announcement, 'One hundred'. We can see that the pilot only looks at the
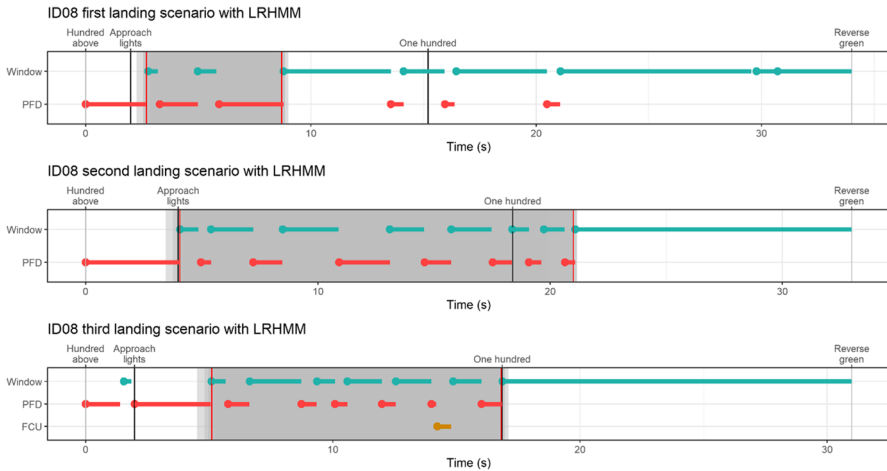
Fig. 8 Pilot ID08's three landing scenarios, as shown in Figure 6, but now with the LRHMM. The first white area indicates the first state, the grey area indicates the second state, and the last white area indicates the third state. The darker the grey area is, the higher the posterior probability is to be in the second state. The two red vertical lines indicate the most probable change point from state to another state

PFD in the first phase, starts to shift between the indicators and the window in the second phase, and looks mainly at the window in the final phase. It might be said that the third landing scenario of pilot ID08 is a text-book example of visual scanning during the landing: indicators during phase 1, switching during phase 2, and looking at the outside environment during phase 3.

Figure 8 shows the three landing scenarios of pilot ID08 based on the LRHMM. The first white area indicates the first state of the model, the grey area indicates the second state, and the last white area indicates the third state of the model. We can see that the model finds the three different phases based on the pilot's monitoring. In the second landing scenario, where the pilot is monitoring exactly as they are taught in training, the model also starts its second state (grey area) right at the second announcement and ends it a bit later following the third announcement. In the third landing scenario, the model finds the second state a little after the second announcement, indicating that the pilot starts to shift between the AOIs later than they are taught in the training but ends at the right time at the third announcement; that is, they change their monitoring at the right time. In the first landing scenario, the model's second state starts a little after the announcement. The pilot changes the monitoring marginally late but also ends long before the third announcement. As we can see from the raw data, the pilot does not perform systematic monitoring in the second phase of the first landing scenario. Rather, the pilot looks outside for a long time, and after that, only a few quick times at the indicator. According to the model, this pilot is looking too briefly at the indicators and thus locates the third state before the third announcement, indicating that the pilot is starting the final monitoring phase too early.

Figure 9 shows the three landing scenarios of pilot ID10 with the LRHMM. As can be seen from the recorded gaze transitions between the OTW and PFD and the
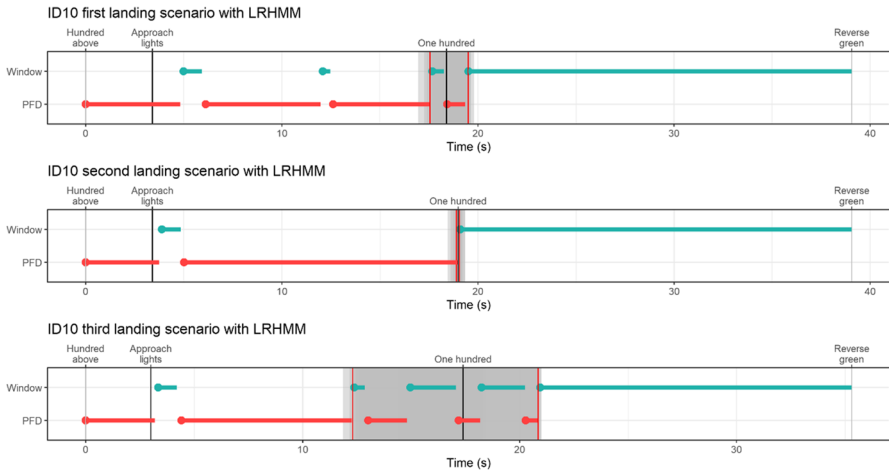
Fig. 9 Pilot ID10's three landing scenarios, as shown in Fig. 7, but now with the LRHMM (see Fig. 8 for details on the interpretation of this figure)

modelling results, pilot ID10 does not exhibit the consistent visual scanning behaviour instructed in the training. The clearest deviation from the expected visual scanning behaviour can be seen in the second landing scenario: there is only one glance to the OTW, and the LRHMM identifies the second phase as only a very brief moment before the final callout. Here the second state is not clearly visible, but the model is bound to find all three states, so it finds the second state right at the third announcement, when the pilot switches from the indicator to the window. In other words, the model indicates that the pilot monitors the first phase until the third announcement and remains in the second phase for less than a second, when the pilot changes AOIs and then starts monitoring the third phase at the correct time. The difference between the announcement and the start of a second state and the briefness of the second phase quickly alert the instructor that the pilot's monitoring is incorrect. In the first landing scenario, the model locates the second state in the middle of the third announcement. It regards the small glances to the window as too quick in the second stage. In the third landing scenario, the second state also starts late. The pilot looks outside quickly at the right time at the second announcement but does not systematically shift between the AOIs until the model's second state starts.

3.2.3 The overall representation of the HMM

Figure 10 depicts the overall results of the states detected by the HMM. It is a heatmap illustrating all pilots' monitoring states in all landing scenarios with the announcements. It summarises the pilots' overall monitoring behaviour and how their monitoring is related to the three stages based on the announcements (see Fig. 2). The darker the colour, the more pilots are engaging in the same visual behaviour at the same time in the second phase of the monitoring. As we can see
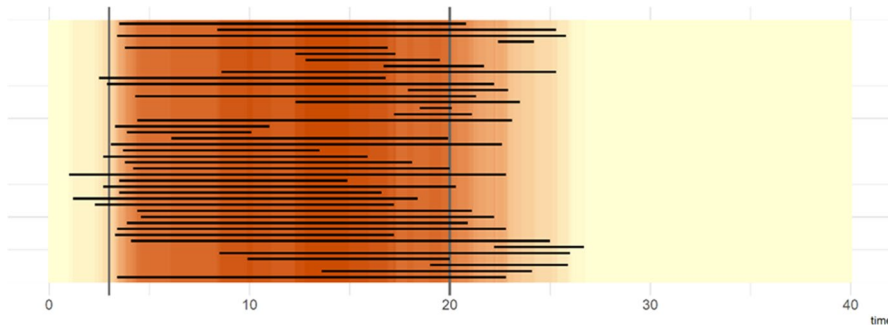
Fig. 10 Heatmap of all pilots' visual scanning in all landing scenarios. The second state is indicated in orange. The darker the colour, the more pilots switched their attention at that point, which we expect between the second and third announcements, indicated by the two black vertical lines (see also Fig. 2). Furthermore, the black horizontal lines represent all individual pilots' second states. Please note that as the time between the two announcements is not exactly the same for all pilots, we standardise this to ensure that the black vertical lines indicate the exact moment of the announcement

for a few of the pilots, the behaviour starts a bit earlier than expected (based on the expected stages in Fig. 2), and there is a tendency to continue the visual scanning after the last announcement.

While the heatmap is informative on a general level and shows the summary of the full group of pilots, the disadvantage is that relevant information regarding individuals is lost, and it does not show which pilots, in which landing scenarios, are showing (sub)optimal visual scanning. Therefore, it is also important to examine the information in more detail, and the black horizontal lines in Fig. 10 shows the visual scanning behaviour/second state of HMM of all landing scenarios for all pilots, allowing for a quick overview of every landing scenario for every pilot.

4 Discussion

In vision-intensive professions, it is critical that the professionals' visual attention behaviour matches predefined scenarios. This ensures the safe completion of tasks and efficient monitoring of operations. In this study, we investigated pilots' gaze behavior during three landing scenarios, and our findings offer insights into the pilots' visual scanning processes (RQ1). Eye tracking facilitated objective and non-intrusive observations of how accurately pilots looked at the right places at the right times (see also Peißl et al., 2018). This information could be valuable for refining pilot education and training practices in the future. In teaching landings, particular attention must be paid to three aspects. First, pilots tend to look out the window too early during challenging landing scenarios. Second, during landing, pilots have a

tendency to look at the indicators for too long. And third, during practice, there is a general tendency to continue scanning the PFD after the final announcement.

We also aimed to take a step towards more automated analysis by seeking an automatic method to determine whether pilots exhibit the expected visual scanning behavior during landing scenarios (RQ2). The results showed that an HMM was capable of modeling the various states of pilots' gaze behavior. The novel contribution of our study is that we combined existing techniques and procedures, such as eye tracking and HMM, to build a specific case for the field of pilot training. Through this study, we sought to contribute to the empirical literature on how technological tools (such as eye-tracking glasses) and analysis procedures (such as HMM) can be integrated with a focus on enhancing the learning and training environment for professionals in light of their further professional development. Namely, we were able to automatically detect actual phases of visual scanning within the different landing scenarios of the pilots through HMM. Thus, our study is one empirical example in which the advantage of HMM technique is that we are not bound to manual decisions based on graphs or descriptive data. In the future, it might be possible to identify patterns also in less clear learning situations, for instance, when no specific announcements are made and the number of hidden states is not known in advance. In such cases, models with varying numbers of states can be estimated and compared using common statistical model selection methods. Future research, including more complex task representations, should further help to uncover the opportunities and practical applicability in the field.

Our findings are in line with previous reports that eye tracking seems to have the potential to capture the temporal features of gaze behaviour (e.g. Niehorster et al., 2020; Rudi et al., 2020). However, gaze behaviour patterns are not yet systematically used in pilot training, and when they are used, they are often analysed after the training, resulting in delayed feedback. Although the importance of immediate feedback has been recognised (Lefrançois et al., 2021), steps remain to be taken to enable more immediate feedback to learners in a simulation environment. Furthermore, the manual analysis of eye-tracking data is resource-intensive. Our study is one step towards integrating eye tracking with analytics technology in the future, in order to enhance learning and support the development of adaptive learning systems (Tsai et al., 2022). In particular, there is an increasing trend in educational Artificial Intelligence (AI) to identify resources and pedagogical approaches that are considered appropriate for learners' needs and can predict potential outcomes and recommend the next steps of the learning process for them (e.g. Zhang & Aslan, 2021).

Despite acknowledged challenges (Mengtao et al., 2023), HMMs will contribute to real-time analysis (e.g. as in speech recognition) or very close to real-time analysis. The results of this study can be used to provide immediate feedback on gaze behaviour to learners immediately after they finish their landing scenarios. However, the results can also be used to support future developments in pilot training.

In this respect, combining real-time eye-tracking data with HMMs and AI-based estimations of the number of states to be expected might allow the provision of real-time on-the-fly information regarding how pilots are scanning the visual information. This could inform not only instructors by providing just-in-time feedback but also the flight simulator software itself, enabling it to provide more personalised adapted feedback, which may lead to more effective training approaches. In practice, this would mean that the model would first be trained on previous experiments (this gives us the estimates for the model parameters, i.e. the transition and emission probabilities), and then we could feed data on new individuals into the HMM and get so-called online predictions of the current state (as well as correct our previous predictions, if necessary).

From a pedagogical perspective, pilots' visual scanning is an integral part of their expertise development (Lu et al., 2020; Ziv, 2016), thus making it crucial for pilot training. This study presented the results in easily comprehensible graphs that could be displayed on dashboard-type instruments in the future, making the information useful for human (e.g. instructors or pilots) interpretation. These types of graphs could be built into a dashboard to inform instructors about the overall and individual performance of their trainees. A practical implication is that we could automatically provide instructors with information about the performance of their trainees by providing an overall view of the full group (see e.g. the heatmap in Fig. 10) as well as more detailed information (see e.g. individuals in Appendix Fig. 11). More generally, our results could support professional learning and expertise development in visual-intensive professions, in which decisions are based on visual information indicators (Gegenfurtner et al., 2022; Jossberger, 2022).

It seems clear that eye tracking could be useful in all professional learning and expertise development settings in which there is a need to ensure that learners are focusing on the right information at the right time. Despite this potential, pilot education has been primarily based on a minimum number of hours of flying. At the same time, there is increasing interest in how eye tracking can be applied in pilot education (see Lefrançois et al., 2021; Niehorster et al., 2020). For example, Ziv (2016) illustrated that experienced pilots frequently looked at flight instruments but did not focus on each instrument for too long. Sullivan et al. (2011) found that gaze parameters and scan management skills were associated with pilots' total flight hours and that experienced pilots shifted their gaze between the map display and OTW more often than novices. Regarding visual expertise, we would like to highlight that one specific area of interest has been overlooked – what is excluded (i.e. what is not looked at) can have just as much of an impact on professional performance as what is looked at. For example, in the case of pilots' professional learning and expertise development, analysing eye movements may be particularly useful for detecting when learners do not pay attention to the right place at the right time. Eye-tracking technology can support the development of more personalised training

programmes and enable pilot trainers to adjust their teaching approaches by identifying areas in which learners struggle or lack attention (see also Niehorster et al., 2020).

4.1 Limitations and strengths

First, all limitations of the case-study approach should be kept in mind, as the research is exploratory and limited by the sample size. Namely, a small sample size may not fully encompass all the different factors that impact pilots' visual attention during training, such as their level of experience, the type of aircraft they are flying, or the environmental conditions. Second, in educational settings, many open challenges still exist related to data, analytical methods, and educational practices (Wang et al., 2021). Third, regarding automatic analysis, the model was not always completely right. For example, the first landing of pilot ID08 showed accurate and consistent visual scanning behaviour, but the HMM indicated that visual scanning happened only partially, complementing the descriptive results. Fourth, we used HMMs in an offline way (after data collection and AOI extraction). While HMMs, in general, can be applied for online (real-time) prediction problems as well, future research is needed to examine how HMMs and other machine learning methods could be used efficiently for processing gaze data in real time. Fifth, statistical methods assessing the accuracy and uncertainty of the output of HMMs should be further studied, especially in the case of real-time automatic analysis. This is challenging because we do not know the ground truth (when the pilot switch between the scanning stages), but the robustness and stability of the HMM parameter estimates with respect to different but comparable datasets could be studied with additional data sources. Sixth, in authentic situations, two pilots are flying an aircraft, as flying involves numerous tasks, including navigation, monitoring and managing avionics and systems, and making critical decisions. Please note that in this study, we focused only on one of the pilots. In the future, the approach presented in this research needs to be extended to examine the interaction and collaborative sensing of the pilot-in-flying and co-pilot using dual eye tracking.

Despite the acknowledged limitations, our study has several strengths. The sample is unique, including professional pilots flying in an authentic full-flight simulator. The study is located at the frontline of empirical studies in professional learning focusing on if and how pilots are visually scanning the indicators during challenging landing scenarios. Information about visual expertise is also critical in other vision-intensive professional learning situations where actions need to be taken based on the visual information obtained. Several recent studies have indicated this need, for example, surgeons need to perform surgeries accurately based on visual information, teachers need to capture visual information about the classroom during scaffolding etc. (Haataja et. al., 2019; Jarodzka et al., 2021). Our results regarding the

HMM may be helpful for including AI in teaching–learning processes. Finally, this study was conducted in a multidisciplinary context, involving researchers and experts from different fields. Whereas educational researchers (Authors 1, 2, 6, and 7) specialise in evaluating learning processes with multimodal data in technology-enhanced learning settings, the researcher's expertise in aviation (Author 5) brought in the necessary information about the specific context. The role of the two researchers specialising in educational technology (Authors 4, 9) focused on the simulation environment itself and how to collect and analyse multimodal data on learning patterns and behaviours. Finally, the statistics researchers (Authors 3, 8) had specific methodological expertise, especially with regard to HMM models. Thus, although the data may seem limited in terms of the number of participants, data collected through eye-tracking classes can be efficiently and reliably analysed with statistical methods.

4.2 Conclusions

Simulations play a critical role in modern aviation training, ensuring that pilots are well prepared to handle a wide range of scenarios safely and effectively. So far, in authentic flight simulators, instructors have been located behind the learner and thus unable to see what the pilot is looking at and focusing attention on. Using eye tracking allowed us to analyse gaze paths to determine how the pilots were observing their working environment. This study demonstrated that pilots' visual scanning behaviour during landing can be successfully detected using eye tracking and modelled utilising HMMs. Automated analysis of gaze behaviour is a prerequisite for adaptive learning systems capable of providing individualised learning support. It could be particularly useful for the future development of flight simulator software. Automatic detection of whether professionals are looking at the right place at the right time could provide immediate feedback to instructors and pilots during flights. Then, based on this information, flight simulators would be able to detect suboptimal visual scanning behaviour, indicate essential attention points during the landing scenario, and instruct pilots on where to look and how to proceed. Overall, descriptive statistics and HMMs appear to complement each other. In the future, the HMM approach could be used as a basis for identifying the best scanning strategies and developing tools that provide automated information and feedback on specific pilot competencies to cadets and flying instructors. Providing individual feedback in real time could allow pilots to assess and correct their operations during the learning process.
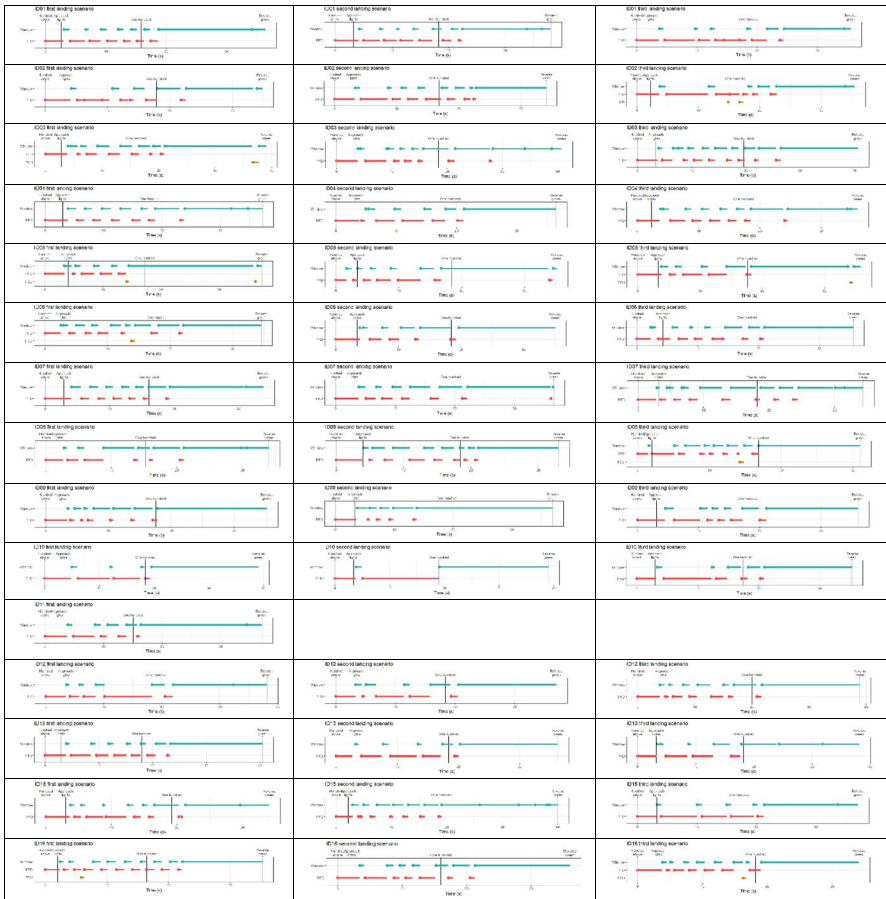
Appendix 1

Fig. 11 Figures of the landings without HMM

Appendix 2

Appendix 2. represents the individual results for each pilot based of the HMM. The grey areas indicate temporally where the model defines the scanning sequence to happen. In general, HMM results are aligned with the descriptive statistics and indicate accuracy and consistency of visual scanning.
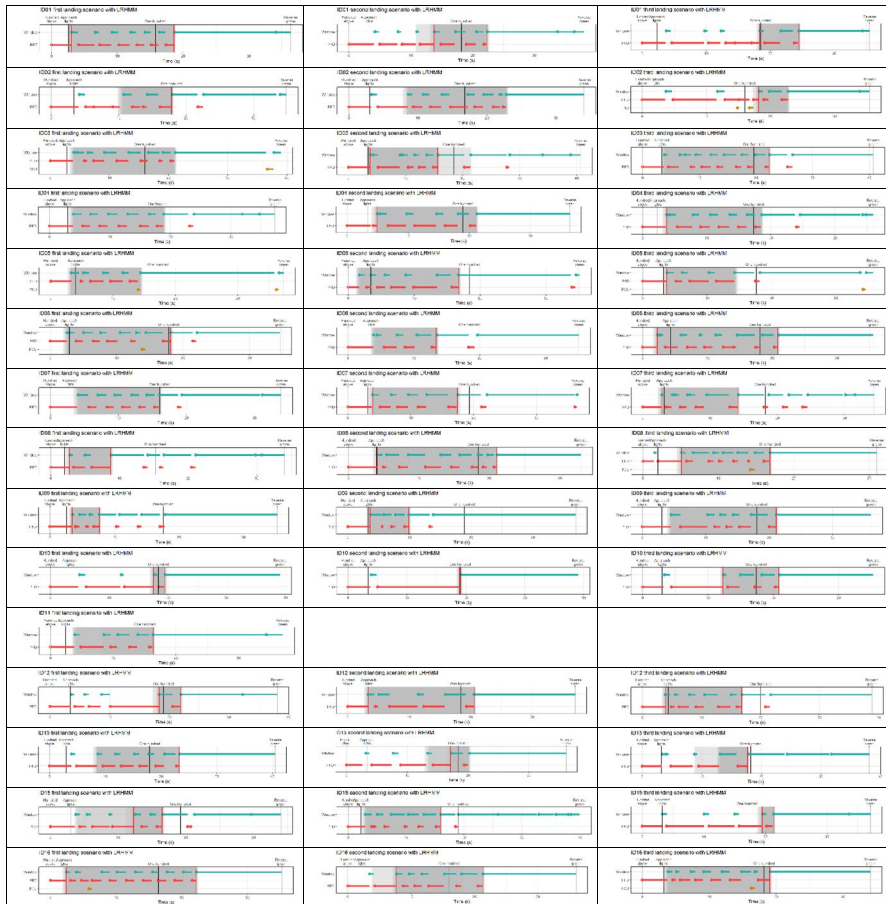

Fig. 12 Figures of the landings with HMM

Appendix 3

Table 1 Estimated transition and emission probabilities. Emission probabilities in state 2 are fixed

	State 1	State 2	State 3
State 1	0.979	0.021	0.000
State 2	0.000	0.987	0.013
State 3	0.000	0.000	1.00
	PFD	**OTW**	**FCU**
State 1	0.915	0.085	0.000
State 2	**0.590**	**0.400**	**0.010**
State 3	0.023	0.974	0.003

Declarations

Conflict of interest None.

References

Ahmadi, N., Romoser, M., & Salmon, C. (2022). Improving the tactical scanning of student pilots: A gaze-based training intervention for transition from visual flight into instrument meteorological conditions. *Applied Ergonomics, 100*, 103642.

Allsop, J., Gray, R., Bulthoff, H. H., & Chuang, L. (2016). Effects of anxiety and cognitive load on instrument scanning behavior in a flight simulation. In Proceedings of the 2016 IEEE Second Workshop on Eye Tracking and Visualization (ETVIS) (pp. 55–59). IEEE.

Baum, L. E., & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The Annals of Mathematical Statistics, 67*(6), 1554–1563. https://doi.org/10.1214/aoms/1177699147

Behrend, J., & Dehais, F. (2020). How role assignment impacts decision-making in high-risk environments: Evidence from eye-tracking in aviation. *Safety Science, 127*, 104738.

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, 68*(7), 569–579.

Bergamo, P. A. D. S., Streng, E. S., de Carvalho, M. A., Rosenkranz, J., & Ghorbani, Y. (2022). Simulation-based training and learning: A review on technology-enhanced education for the minerals industry. *Minerals Engineering, 175*, 107272.

Chaudhuri, S., Muhonen, H., Pakarinen, E., & Lerkkanen, M. K. (2022). Teachers' visual focus of attention in relation to students' basic academic skills and teachers' individual support for students: An eye-tracking study. *Learning and Individual Differences, 98*, 102179.

Chen, Y., & Tsai, M.-J. (2015). Eye-hand coordination strategies during active video game playing: An eye-tracking study. *Computers in Human Behavior, 51*, 8–14.

Dehais, F., Behrend, J., Peysakhovich, V., Causse, M., & Wickens, C. D. (2017). Pilot flying and pilot monitoring's aircraft state awareness during go-around execution in aviation: A behavioral and eye tracking study. *The International Journal of Aerospace Psychology, 27*(1–2), 15–28.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological), 39*(1), 1–38. http://www.jstor.org/stable/2984875.

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.

European Union Aviation Safety Agency. (2020). Appendix to opinion No 08/2019 (A) (RMT.0599). https://www.easa.europa.eu/sites/default/files/dfu/Appendix%20to%20Opinion%20No%2008-2019%20%28A%29%20%28RMT.0599%29.pdf. Accessed 13 Dec 2023.

Faulhaber, A. K., Friedrich, M., & Kapol, T. (2022). Absence of pilot monitoring affects scanning behavior of pilot flying: implications for the design of single-pilot cockpits. *Human Factors, 64*(2), 278–290.

Gao, S., & Wang, L. (2024). How flight experience impacts pilots' decision-making and visual scanning pattern in low-visibility approaches: Preliminary evidence from eye tracking. *Ergonomics*, 1–17. https://doi.org/10.1080/00140139.2023.2298992

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 Peview, 23*, 523–552.

Gegenfurtner, A., Gruber, H., Holzberger, D., Keskin, Ö., Lehtinen, E., Seidel, T., Stürmer, K., & Säljö, R. (2022). Towards a cognitive theory of visual expertise: Methods of inquiry. In C. Damşa, A. Rajala, G. Ritella, & J. Brouwer (Eds.), *Re-theorizing learning and research methods in learning research* (pp. 142–158). Routledge.

Gruber, H., & Harteis, C. (2018). Individual and social influences on professional learning. Supporting the acquisition and maintenance of expertise. Springer. https://doi.org/10.1007/978-3-319-97041-7

Haataja, E., Moreno-Esteva, E. G., Salonen, V., Laine, A., Toivanen, M., & Hannula, M. S. (2019). Teacher's visual attention when scaffolding collaborative mathematical problem solving. *Teaching and Teacher Education, 86*, 102877.

Harris, D. J., Arthur, T., de Burgh, T., Duxbury, M., Lockett-Kirk, R., McBarnett, W., & Vine, S. J. (2023). Assessing expertise using eye tracking in a virtual reality flight simulation. *The International Journal of Aerospace Psychology, 33*(3), 153–173.

Haslbeck, A., & Zhang, B. (2017). I spy with my little eye: Analysis of airline pilots' gaze patterns in a manual instrument flight scenario. *Applied Ergonomics, 63*, 62–71.

Helske, S., & Helske, J. (2019). Mixture hidden Markov models for sequence data: The seqHMM package in R. *Journal of Statistical Software, 88*(3), 1–32.

Helske, S., Helske, J., & Eerola, M. (2018). Combining sequence analysis and hidden Markov models in the analysis of complex life sequence data. In G. Ritschard & M. Studer (Eds.), *Sequence Analysis and Related Approaches: Innovative Methods and Applications. Life Course Research and Social Policies* (Vol. 10, pp. 185–200). SpringerOpen.

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.

Huang, Y. M., Silitonga, L. M., & Wu, T. T. (2022). Applying a business simulation game in a flipped classroom to enhance engagement, learning achievement, and higher-order thinking skills. *Computers & Education, 183*, 104494.

Jarodzka, H., Skuballa, I. T., & Gruber, H. (2021). Eye-tracking in educational practice: Investigating visual perception underlying teaching and learning in the classroom. *Educational Psychology Review, 33*, 1–10.

Jeffries, P. R. (2005). A framework for designing, implementing, and evaluating: Simulations used as teaching strategies in nursing. *Nursing Education Perspectives, 26*(2), 96–103.

Jin, H., Hu, Z., Li, K., Chu, M., Zou, G., Yu, G., & Zhang, J. (2021). Study on how expert and novice pilots can distribute their visual attention to improve flight performance. *IEEE Access: Practical Innovations, Open Solutions, 9*, 44757–44769. https://doi.org/10.1109/access.2021.3066880

Jossberger, H. (2022). Eye tracking in professional learning and development: Uncovering expertise development among residents in radiology. In Methods for researching professional learning and

development: Challenges, applications and empirical illustrations (pp. 467–489). Springer International Publishing.

Kang, Z., Bass, E. J., & Lee, D. W. (2014). Air traffic controllers' visual scanning, aircraft selection, and comparison strategies in support of conflict detection. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 58, No. 1, pp. 77–81). SAGE Publications.

Kim, J., Singh, S., Thiessen, E. D., & Fisher, A. V. (2020). A hidden Markov model for analyzing eye-tracking of moving objects: Case study in a sustained attention paradigm. *Behavior Research Methods, 52*(3), 1225–1243.

Kneebone, R. (2003). Simulation in surgical training: Educational issues and practical implications. *Medical Education, 37*(3), 267–277.

Kok, E. M., & Jarodzka, H. (2017). Before your very eyes: The value and limitations of eye tracking in medical education. *Medical Education, 51*(1), 114–122.

Krag, B. (2017). Modeling and simulation—Basics and benefits. In P. G. Hamel (Ed.), *In-Flight Simulators and Fly-by-Wire/Light Demonstrators: A Historical Account of International Aeronautical Research* (pp. 19–25). Springer International Publishing.

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., Lee, M. H., Chiou, G. L., Liang, J. C., & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review, 10*, 90–115.

Lefrançois, O., Matton, N., & Causse, M. (2021). Improving airline pilots' visual scanning and manual flight performance through training on skilled eye gaze strategies. *Safety, 7*(4), 70. https://doi.org/10.3390/safety7040070

Lehtinen, E., Gegenfurtner, A., Helle, L., & Säljö, R. (2020). Conceptual change in the development of visual expertise. *International Journal of Educational Research, 100*, 101545.

Lekea, I. K., Stamatelos, D. G., & Raptis, P. (2021). Learning how to escape the unthinkable with virtual reality: The case of pilots' training on emergency procedures. In IOP conference series: Materials science and engineering (vol. 1024, No. 1, p. 012098). IOP Publishing.

Li, W.-C., Braithwaite, G., & Yu, C.-S. (2014). The investigation of pilots' eye scan patterns on the flight deck during an air-to-surface task. In Engineering Psychology and Cognitive Ergonomics (pp. 325–334). Springer International Publishing.

Lijing, W., & Lin, C. (2016). Eye movement comparison of professional and novice pilots in key actions of takeoff phase. *IEEE International Conference on Aircraft Utility Systems (AUS), 2016*, 703–706.

Lounis, C., Peysakhovich, V., & Causse, M. (2021). Visual scanning strategies in the cockpit are modulated by pilots' expertise: A flight simulator study. *PLoS ONE, 16*(2), e0247061.

Lu, Y., Zheng, Y., Wang, Z., & Fu, S. (2020). Pilots' visual scanning behaviors during an instrument landing system approach. *Aerospace Medicine and Human Performance, 91*(6), 511–517.

Mengtao, L., Fan, L., Gangyan, X., & Su, H. (2023). Leveraging eye-tracking technologies to promote aviation safety-a review of key aspects, challenges, and future perspectives. *Safety Science, 168*, 106295.

Muehlethaler, C. M., & Knecht, C. P. (2016). Situation awareness training for general aviation pilots using eye tracking. *IFAC-PapersOnLine, 49*(19), 66–71. https://doi.org/10.1016/j.ifacol.2016.10.463

Muthumanickam, P., Helske, J., Vitoria, A., Johansson, J., & Cooper, M. (2019). Comparison of attention behaviour across user sets through automatic identification of common areas of interest In Proceedings of the 53rd Annual Hawaii International Conference on System Sciences 2020 (HICSS), 1360—1369. HICSS. https://doi.org/10.24251/HICSS.2020.167

Myers, P. L., III., Starr, A. W., & Mullins, K. (2018). Flight simulator fidelity, training transfer, and the role of instructors in optimizing learning. *International Journal of Aviation, Aeronautics, and Aerospace, 5*(1), 6.

Negi, S., & Mitra, R. (2020). Fixation duration and the learning process: An eye tracking study with subtitled videos. *Journal of Eye Movement Research*, 13(6). https://doi.org/10.16910/jemr.13.6.1

Niehorster, D. C., Hildebrandt, M., Smoker, A., Jarodzka, H., & Dahlströhm, N. (2020). Towards eye tracking as a support tool for pilot training and assessment. In Eye-tracking in aviation: Proceedings of the 1st international workshop (ETAVI 2020) (pp. 17–28). ISAE-SUPAERO, Institute of Cartography and Geoinformation (IKG), ETH Zurich.

Palazzi, A., Abati, D., Solera, F., & Cucchiara, R. (2018). Predicting the driver's focus of attention: The DR (eye) VE Project. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 41*(7), 1720–1733.

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.

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/. Accessed 13 Dec 2023.

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77*, 257–286. https://doi.org/10.1109/5.18626

Rainieri, G., Fraboni, F., Russo, G., Tul, M., Pingitore, A., Tessari, A., & Pietrantoni, L. (2021). Visual scanning techniques and mental workload of helicopter pilots during simulated flight. *Aerospace Medicine and Human Performance, 92*(1), 11–19.

Rappa, N. A., Ledger, S., Teo, T., Wai Wong, K., Power, B., & Hilliard, B. (2019). The use of eye tracking technology to explore learning and performance within virtual reality and mixed reality settings: A scoping review. *Interactive Learning Environments, 30*(7), 1338–1350.

Rudi, D., Kiefer, P., & Raubal, M. (2020). The instructor assistant system (iASSYST) – Utilizing eye tracking for commercial aviation training purposes. *Ergonomics, 63*(1), 61–79. https://doi.org/10.1080/00140139.2019.1685132

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. https://doi.org/10.1080/00140139.2018.1501093

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The International Journal of Aviation Psychology, 8*(3), 197–208.

Sullivan, J., Yang, J. H., Day, M., & Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation, Space, and Environmental Medicine, 82*(9), 871–878.

Tissenbaum, M. (2020). I see what you did there! Divergent collaboration and learner transitions from unproductive to productive states in open-ended inquiry. *Computers & Education, 145*, 103739.

Tsai, M.-J., Wu, A.-H., Bråten, I., & Wang, C.-Y. (2022). What do critical reading strategies look like? Eye-tracking and lag sequential analysis reveal attention to data and reasoning when reading conflicting information. *Computers & Education, 187*, 104544.

Udale, R., Tran, M. T., Manohar, S., & Husain, M. (2022). Dynamic in-flight shifts of working memory resources across saccades. *Journal of Experimental Psychology, Human Perception and Performance, 48*(1), 21–36.

Ulutas, B. H., Özkan, N. F., & Michalski, R. (2020). Application of hidden Markov models to eye tracking data analysis of visual quality inspection operations. *Central European Journal of Operations Research, 28*(2), 761–777.

Vähäsantanen, K., Hökkä, P., Paloniemi, S., Herranen, S., & Eteläpelto, A. (2017). Professional learning and agency in an identity coaching programme. *Professional Development in Education, 43*(4), 514–536.

Van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance cognitive and metacognitive processes in computer-based learning environments. In International handbook of metacognition and learning technologies (pp. 143–156). Springer New York.

Villafaina, S., Fuentes-García, J. P., Gusi, N., Tornero-Aguilera, J. F., & Clemente-Suárez, V. J. (2021). Psychophysiological response of military pilots in different combat flight maneuvers in a flight simulator. *Physiology & Behavior, 238*, 113483.

Viterbi, A. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory, 13*(2), 260–269. https://doi.org/10.1109/TIT.1967.1054010

Wang, Y., Lu, S., & Harter, D. (2021). Towards collaborative and intelligent learning environments based on eye tracking data and learning analytics: A survey. *IEEE Access, 9*, 137991–138002.

Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, W. J. (2001). *Pilot task management: Testing an attentional expected value model of visual scanning* (ARL-01–14/NASA-01–7). Savoy, IL, UIUC Institute of Aviation Technical Report.

Wu, D., Huang, H., Liu, N., & Miao, D. (2019). Information processing under high and low distractions using eye tracking. *Cognitive Processing, 20*(1), 11–18.

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence, 2*, 100025.

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.

Authors and Affiliations

Raija Hämäläinen[1] · **Bram De Wever[2]** · **Katriina Sipiläinen[1]** · **Ville Heilala[1]** · **Arto Helovuo[3]** · **Sami Lehesvuori[1]** · **Miitta Järvinen[1]** · **Jouni Helske[4]** · **Tommi Kärkkäinen[5]**

✉ Raija Hämäläinen
raija.h.hamalainen@jyu.fi

[1] Department of Education, University of Jyväskylä, P.O. Box 35, 40014 Jyväskylä, Finland

[2] Department of Educational Studies, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium

[3] Finnair, Helsinki-Uusimaa, Finland

[4] Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35, 40014 Jyväskylä, Finland

[5] Faculty of Information Technology, University of Jyväskylä, P.O. Box 35, 40014 Jyväskylä, Finland