Visual Scanning Techniques and Mental Workload of Helicopter Pilots During Simulated Flight

Giuseppe Rainieri; Federico Fraboni; Gabriele Russo; Martin Tušl; Andrea Pingitore; Alessia Tessari; Luca Pietrantoni

INTRODUCTION: The visual scanning techniques used by helicopter pilots are a critical skill to accomplish safe and correct landing. According to the human information processing theory, visual scanning techniques can be analyzed as a function of fixation location, number, and duration of fixations.

METHODS: This study assessed these techniques in expert and novice pilots during an open sea flight simulation in a low-workload condition, consisting of a daylight and good weather simulation, and in a high-workload condition of night-time, low visibility, and adverse weather conditions. Taking part in the study were 12 helicopter pilots. Mental workload was assessed through psychological measures (NASA-TLX). The pilots' performance was assessed and eye movements were recorded using an eye-tracker during four phases of the flight simulations.

RESULTS: Overall, pilots made more fixations out of the window (OTW; 22.54) than inside the cockpit (ITC; 11.08), Fixations were longer OTW (830.17 ms) than ITC (647.97 ms) and they were shorter in the low-demand condition (626.27 ms). Further, pilots reported higher mental workload (NASA-TLX) in the high-demand condition compared to the low-demand condition, regardless of their expertise, and expert pilots reported a lower mental workload compared to novice pilots.

DISCUSSION: Pilots' performance and perceived mental workload varied as a function of expertise and flight conditions. Pilots rely on instrument support during the cruise phase and external visual cues during the landing phase. The implications for a new visual landing system design are discussed.

KEYWORDS: mental workload, helicopter pilots, visual scanning techniques.

Rainieri G, Fraboni F, Russo G, Tušl M, Pingitore A, Tessari A, Pietrantoni L. *Visual scanning techniques and mental workload of helicopter pilots during simulated flight*. Aerosp Med Hum Perform. 2021; 92(1):11–19.

Maneuvering a helicopter is a cognitively complex task that implies high mental workload situations and involves various environmental and human factors aspects that affect pilots' performance and safety. Operations over the open sea and, in particular, approaching and landing on a ship deck are some of the most demanding tasks for helicopter pilots.[3,16,21] Lack of visual cues,[2,23] restricted and unstable landing areas,[9] the dynamic environment around the ship,[14] and sea spray[10] are some of the challenges that pilots need to handle during takeoff, approach, and landing. Moreover, scanning of the flight instruments inside the cockpit requires frequent head-down movements that further increase the pilot's mental workload and may compromise his or her situation awareness.[18] As a result, extreme levels of mental workload reduce the pilots' ability to react to incoming information and increase the likelihood of human error.[24] Effective visual scanning techniques help pilots in maintaining a high level of situation awareness to effectively collect and integrate relevant information at the right moment.[5] Research has shown that efficient visual scanning is a skill that pilots need to develop, and there are significant interindividual differences among pilots.[29]

Eye tracking has proven to be a valuable method for detecting visual scanning techniques.[17] Taking in to account the principles of human information processing, a pilot's visual attention should be understood as an endogenously controlled process which enables the acquisition of relevant information.[28] Therefore, the visual scanning techniques partly consist of gaze

From the Alma Mater Studiorum, University of Bologna, Bologna, Italy.

This manuscript was received for review in May 2020. It was accepted for publication in October 2020.

Address correspondence to: Luca Pietrantoni, M.S., Ph.D., Alma Mater Studiorum, Department of Psychology, University of Bologna, Viale Berti-Pichat, 5, Bologna 40126, Italy; luca.pietrantoni@unibo.it.

DOI: https://doi.org/10.3357/AMHP.5681.2021

concentrations, so-called fixation location (FL) and regular gaze, on the sources of information.[18] Studies based on ocular motion analysis show many benefits; for instance, eye movements are insensitive to limb movements, subjects do not need any special training to use eye-tracking devices, and the data allow understanding where the attention is focused during the task.

Several studies have used eye-tracking in a simulated helicopter flight to assess pilots' visual scanning techniques and mental workload. Some researchers studied pilots' scanning in different phases of a simulated flight with different levels of workload/job demands.[3] The authors identified takeoff and landing as the phases with the highest mental workload. Their results also showed that the higher the mental workload, the more random or untargeted the fixation locations were.

Previous research suggests differences in visual scanning strategies adopted by novice and expert helicopter pilots. Expert pilots showed more complex scanning patterns, in terms of higher distribution of the gazes, when compared to novices.[15] Another investigation found that the gaze duration of experts was shorter and fixations on instruments were more frequent in experts in comparison to novices, allowing them to react more flexibly to mission demands.[1] A study also showed that experienced pilots had significantly shorter gaze duration, overall more fixations, and more fixations on relevant points or instruments than novices.[11] In the same line, a research study revealed that more experienced pilots had shorter gaze duration and more frequent saccades between the cockpit and outside-of-the-window (OTW), suggesting faster and more accurate visual scanning.[20]

Most studies have shown that experienced and novice helicopter pilots differ in the frequency of scanning OTW.[20,25,26] Whether expert or novice pilots have more or fewer fixations OTW is not always clear and it seems to depend also on the mission demands and the phase of the flight.[17] Interaction between flight experience and mission demands has been investigated.[18] The authors tested multivariate effects of flight experience and mission demands on FL, mental workload, and performance in military helicopter pilots. They also explored the deviation between objective measures and subjective assessment of scanning techniques. Landing on a frigate was considered a high demand situation, whereas landing on a pinnacle was ranked as a low demand situation. The study revealed differences in OTW gazes between novice and expert pilots. The results suggest that 54% of the variance could be explained by a combination of pilots' competence and mission demands. Expert pilots had more OTW gazes in low demand situations, whereas the opposite result was found for high demand situations. The study also found that pilots overestimated the amount of OTW gazes and underestimated their instrument checks. This deviation was more pronounced in student pilots. The authors concluded that there are significant differences in visual scanning techniques of experienced and novice pilots when facing different mission demands.

Additionally, the NASA-TLX questionnaire was adopted to assess the mental workload in the two missions. They noticed greater mental workload for student pilots than for flight instructors, which was not true for the total workload. Finally, a portion of the study explored mental workload of helicopter pilots in two flight situations: 'standard demands' and 'high demands'.[6] Pilots verbally rated their workload level every 1.5 min using a rating scale from 1 (very low) to 4 (very high). The expert subjects showed medium level of workload, reporting positive emotions with low emotional intensity. The less-experienced pilots showed increasing physiological activation (measured in terms of skin conductance) as the perceived workload increased, and their emotional state (evaluated by the Izard differential emotions scale) referred to both positive and negative emotions. The authors argue that the high interindividual variability of the results highlights the complex link between physiological and psychological parameters with workload.

Following the experimental design of Robinski and Stein,[18] the present study explored the visual scanning techniques of pilots during different phases of flight and the effect of flight experience and mission demands. We collected both objective and subjective quantitative data from the pilots using an eye tracker and quantitative scales.

Because of the ongoing debate about these issues, our research focused on the replicability of the results. Additionally, Robinski and Stein analyzed the pilots' scanning techniques during a flight on a frigate and a pinnacle.[18] Our experimental setting was a simulated flight on a frigate in two weather conditions. Moreover, compared to the previous research,[4] the present study explored the hypotheses on a larger sample of Italian Navy pilots using a helicopter flight simulator.

The present study investigates how pilots' expertise (expert vs. novice), different flight conditions (high or low-demand task conditions), and flight phases (i.e., takeoff, cruise, approach, landing) affect the pilots' mental effort, performance, and fixations. More precisely, we considered pilots' scanning technique as a function of brief glimpses and intentional fixations, according to the holistic model of imagine perception.[30] Compared to novices, expert pilots are able to extract more information during the first glimpse and are then able to fixate the relevant areas.

Previous studies suggest that experienced pilots use more efficient visual scanning techniques; however, these techniques may differ in relation to the task demands.[6,18] Moreover, the variation between the objective number of fixations and those self-reported could be beneficial when interview data are considered.

Therefore, we tested our hypotheses in a two-way approach, specifically for the two flight conditions and for each of the simulation's phases with different levels of task demands. Variables' interactions were tested using a multivariate approach. We hypothesized that:

- H1: expert pilots will experience lower mental workload than novice pilots and higher task demands will increase the amount of mental workload experienced by pilots;
- H2: expert pilots will receive higher scores in the evaluation of their performance compared to novice pilots, and higher

task demands will lower the performance evaluation of the pilots compared to the low-demand task condition;

- H3: expert pilots will have more frequent ITC and OTW fixations than novice pilots;
- H4: flight experience, task demands, phases of flight, and the fixation locations will affect the amount and duration of fixations by helicopter pilots;
- H4A: expert pilots will have more and shorter fixations than novice pilots;
- H4B: in the high-demand task condition pilots will have more and shorter fixations than in the low task demand condition;
- H4C: pilots will have more frequent and shorter fixations ITC than OTW;
- H4D: pilots' amount and duration of fixations will vary according to the specific phase of flight;
- H5: pilots will tend to overestimate the amount of OTW fixations and underestimate the amount of ITC fixations; and
- H5A: the subjective and objective evaluation of pilots' scanning techniques will vary according to task demands and phases of flight.

METHODS

Subjects

Voluntarily taking part in the study were 12 male helicopter pilots recruited from the Italian Navy. The sample included six novice and six expert pilots. Together with a flight instructor, a cutoff for flight hours was defined to differentiate between expert and novice pilots: novice $<$ 1500 and expert $>$ 1500 flight hours. On average, the experience of the novice pilots was 759 (SD = 442.25) flight hours, while the expert pilots had 3300 (SD = 1800.74) flight hours. Three pilots were trained to pilot the EH101 helicopter, whereas the other nine pilots were trained to pilot the SH90 helicopter. Pilots had normal or corrected-to-normal vision with contact lenses. Pilots 9 and 11 were removed from the eye-tracking analysis due to the low quality of the registrations. Each subject provided written informed consent before participating. Institutional Review Board approval was obtained. The study was reviewed by the Ethics Committee of the University of Bologna and by the Ministry of Foreign Affairs (nUCLEON project, industrial track, prot. MAE 01060522018-06-14, 14 June 2018).

Equipment

Flight simulator. All simulations were conducted in the Leonardo Helicopters EH-101 flight simulator. The simulator screens (compass and altimeter) could be set according to the subjects' preferences, therefore cockpit design, operation, and flight dynamics were comparable.

Eye-tracking. We used the Pupil Labs[12] (Berlin, Germany) head-mounted eye-tracker to collect objective data about the pilots' visual scanning. The maximum dispersion was set at 1.0°. The minimum fixation duration was set at 200 ms, while the maximum was set at 2000 ms. We measured two parameters of visual scanning behaviors: FLs and duration of fixations. Based on previous research,[18,22] we divided the areas of interests for FLs into three main categories: OTW, ITC, and Null if the fixation position was not detected correctly.

Subjective workload. We used a smartphone app version of the NASA-TLX[8] for the pilots' subjective evaluation of their mental workload during the tasks. Pilots were asked to evaluate their overall effort after each simulation.

Subjective evaluation of visual scanning techniques. After each simulation, we asked pilots to subjectively assess their visual scanning techniques during the different phases of flight with four items. Pilots were asked to rate the percentage of time they spent monitoring OTW during the respective four phases of flight (i.e., takeoff, cruise, approach, landing) from 0 to 100%. The items were administrated after each simulation.

Performance. An experienced flight instructor who participated in all simulations as a copilot rated the pilots' performance from 1 (very poor execution) to 10 (excellent execution). A score above 6 was considered sufficient. The evaluation considered different performance indicators such as quality of the communication with the copilot, time spent to accomplish the task, the accuracy of approach and landing, and the overall safety of the maneuvering.

Procedure

Before the simulation, each pilot provided information about the number of flight hours and the type of helicopter flown. The Pupil Labs eye-tracker was then mounted on the pilot's head and calibrated. Subsequently, both simulated missions were performed. After each simulation, the pilot self-assessed his workload, scanning techniques, and an expert evaluated the pilot's overall performance during the simulation flight.

The flight simulation consisted of two trials. The first simulation was a flight in standard conditions with daylight and good weather (i.e., a low demand situation). The second simulation was a flight at night with low visibility, unstable ship's deck, and difficult weather conditions (i.e., a high demand situation). These specific scenarios were chosen in cooperation with the flight instructor to account for a low demand and a high demand situation. The simulation started with the helicopter on a ship's deck and the task consisted of a takeoff, cruise, approach, and landing on the same ship. The pilot was supported by an experienced copilot and a Flight Deck Officer who was in the control room and communicated with the pilot via radio. The position of the pilots was always the same, with the copilot in the left seat and the pilot in the right seat. We divided the simulation into four phases: takeoff, cruise, approach, and landing. Takeoff started at the beginning of the simulation and ended when the helicopter was completely off the ship's deck. Cruise ended when the pilot aligned the helicopter along the ship's direction and initiated the descent. Approach ended when the helicopter reached the deck's edge, and the Landing phase ended when the pilot accomplished the touchdown. The flight simulation lasted an average of 4 min 49 s; each of the phases had a different average time: Takeoff 24 s, Cruise 2 min 24 s, Approach 1 min 25 s, and Landing 29 s.

Statistical Analysis

Data analysis was performed with R Studio software (v. 1.2.1335; R Core Team; https://www.R-project.org/). Linear mixed regression model (lme4 package) analyses were computed to analyze the data. Subjective evaluation of mental workload and pilots' performance were analyzed as a function of Expertise (experts vs. novices) and Flight Condition (low demand vs. high demand). Eye-tracking data (Number of Fixations and Duration of Fixations) were analyzed as a function of Expertise (experts vs. novices), Flight Condition (low demand vs. high demand), Phase of Flight (takeoff, cruise, approach, and landing) and Fixation Locations (OTW vs. ITC). The subjective evaluation of visual scanning was analyzed as a function of Expertise (experts vs. novices), Flight Condition (low demand vs. high demand), and Phase of Flight (takeoff, cruise, approach, and landing).

For multiple comparisons, Bonferroni post hoc analyses were performed. Moreover, stepwise regression was involved in modeling the linear mixed regression model in order to have the best model for each dependent variable. Furthermore, data collected through the eye tracker and pilots' subjective evaluation of visual scanning techniques were compared for the explorative analysis. The objective percentage of fixation OTW was calculated: objective percentage of fixations OTW = number of fixations OTW × (1000 / total number of fixations). Deviation variable was calculated: deviation = subjective percentage of fixation OTW − objective percentage of fixation OTW. Negative values indicated an underestimation of the pilot's own workload/visual scanning OTW; positive values indicated an overestimation of the pilot's own workload/visual scanning OTW (e.g., the percentage of visual scanning OTW measured via eye tracker was lower than when it was subjectively estimated).

RESULTS

Mental Workload

The analysis of the mental workload through the NASA-TLX revealed a trend to significance for the single factor Expertise [$F(1, 19) = 4.76$, $\eta_p^2 = 0.307$, $P = 0.055$] and significant differences for Flight Condition [$F(1, 19) = 7.33$, $\eta_p^2 = 0.40$, $P < 0.024$]. Specifically, expert pilots (M = 51.86; SD = 10.27) reported a lower mental workload compared to novice pilots (M = 63.94; SD = 9.91). All pilots reported higher mental workload in the high demand task condition (M = 62.00; SD = 11.22) compared to the low demand task condition (M = 54.83; SD = 13.57).

The interaction Expertise × Flight Condition [$F(1, 19) = 5.61$, $\eta_p^2 = 0.34$, $P < 0.042$] was significant. Post hoc analysis showed that in the low demand task condition expert pilots reported a lower mental workload compared to novices [$t(12.92) = 2.97$, $P = 0.047$]. Moreover, expert pilots reported a higher mental workload in the high demand task condition compared to the low demand task condition [$t(9.02) = 3.74$, $P = 0.020$]. In the high demand task condition both the groups reported similar mental workload [$t(13.75) = 1.073$, $P = 0.71$]. **Fig. 1** shows the means and standard errors for the NASA-TLX scores according to Expertise and Flight Condition.

Performance Evaluation

The evaluation of the performance assessed by the flight instructor revealed significant differences for the factor Flight Condition [$F(1, 10) = 13.75$, $\eta_p^2 = 0.55$, $P = 0.004$] since subjects performed worse in the high demand task condition (M = 5.81; SD = 1.60) compared to the low demand one (M = 6.82; SD = 1.17). Neither the factor Expertise [$F(1, 9) = 1.92$, $\eta_p^2 = 0.15$, $P = 0.20$] nor the interaction Expertise × Flight Condition were significant ($P > 0.05$).

Visual Search Data

Analysis of number of fixations showed significant differences for Fixation Location [$F(1, 109) = 15.97$, $\eta_p^2 = 0.04$, $P > 0.0000$] and Phase [$F(3, 109) = 26.28$, $\eta_p^2 = 0.2$, $P < 0.000$] (see **Table I**). The factors Expertise [$F(1, 109) = 0.89$, $\eta_p^2 = 0.002$, $P = 0.35$] and Flight Condition [$F(1, 109) = 0{,}00$, $\eta_p^2 = 0.000$, $P = 9$] were not significant. Expert pilots (M = 17.46; SD = 15.71) showed a similar number of fixations compared to novices (M = 21.18; SD = 17.78) and the pilots fixated more OTW than ITC overall. According to the phase of flight, the number of fixations ITC or OTW changed. In the takeoff phase, no differences between the fixations ITC [$t(109) = 1.50$, $P = 0.80$] were found. In the cruise phase, pilots made more fixations ITC compared to OTW [$t(101) = 6.014$, $P < 0.0001$]. In contrast, pilots fixated more OTW compared to ITC in the

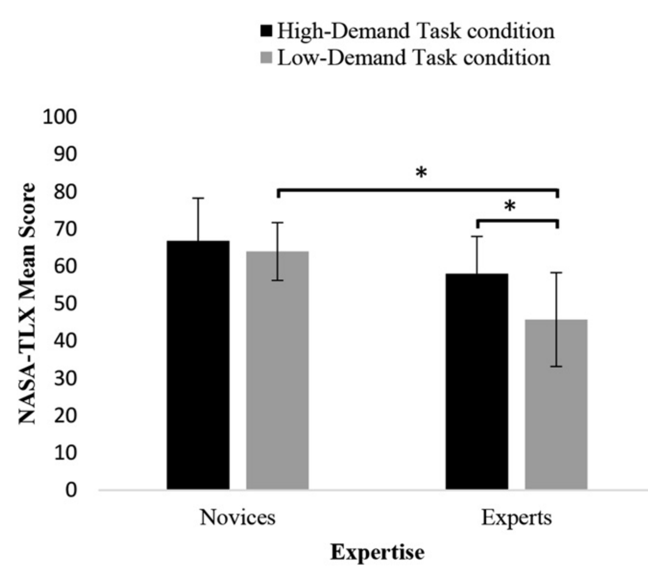

Fig. 1. Means and standard errors of NASA-TLX scores according to Flight Condition and Expertise. *$P < 0.05$.

Table I. Means and Standard Deviations of Number of Fixations and Duration of Fixations According to Location and Phase.

	NUMBER OF FIXATIONS	DURATION OF FIXATIONS (ms)
Fixation Location		
OTW	22.54 (16.46)	830.17 (589.65)
ITC	11.08 (20.41)	647.97 (496.13)
Phase		
Takeoff	12.42 (7.13)	683.07 (521.19)
Cruise	67.17 (20.69)	656.18 (509.10)
Approach	43.58 (17.01)	939.24 (605.06)
Landing	11.33 (7.25)	814.91 (588.96)

OTW: out the window; ITC: inside the cockpit.

approach phase [$t(108) = 10.99$, $P < 0.0001$], while no difference during landing [$t(108) = 1.95$, $P = 0.52$] was found. The means and standard deviations are shown in Table I.

The interactions Phase × Fixation Location [$F(1, 109) = 48.95$, $\eta_p^2 = 0.40$, $P = 0.001$] was significant. Post hoc analysis revealed differences between OTW and ITC fixation in the Cruise and Approach phases [$t(101) = 6.01$, $P < 0.0001$; $t(101) = 11.00$, $P < 0.0001$, respectively]. No differences emerged in the Takeoff and Landing phases [$t(109) = 1.50$, $P = 1$; $t(108) = 1.95$, $P = 0.87$, respectively].

Triple interaction of Fixation Location × Phase × Expertise was significant [$F(3, 109) = 3.85$, $\eta_p^2 = 0.031$, $P = 0.012$]. Post hoc analysis revealed that the number of ITC fixations of expert (M = 32.83; SD = 19.46) and novice pilots (M = 48.50; SD = 22.01) in the Cruise phase (t = 3.33, $P = 0.084$) were different. No differences emerged in the other phases [Takeoff: $t(103) = 0.13$, $P = 1$; Approach: $t(99.7) = 0.64$, $P = 1$; Landing: $t(108.9) = 3.40$, $P = 1$)] or between the groups for the number of fixations made OTW [Takeoff: $t(97.5) = 0.66$, $P = 1$; Cruise: $t(97.5) = 1.85$, $P = 1$; Approach: $t(97.5) = 0.80$, $P = 1$; Landing: $t(97.5) = 1.47$, $P = 1$) The means and standard deviations are presented in **Table II**.

Across the phases, differences in the expert pilots' number of fixations ITC were found. In the Takeoff phase subjects fixated less ITC than in the Cruise phase ($t = 4.65$, $P = 0.0004$) and the number of fixations ITC was higher in Cruise than the Approach ($t = 6.56$, $P < 0.0001$) and Landing phases ($t = 3.66$, $P = 0.031$).

Results were very similar for novice pilots. The number of fixations ITC was higher in Cruise when compared to the Takeoff ($t = 5.31$, $P = 0.0001$), Approach ($t = 7.66$, $P < 0.0001$), and Landing phases ($t = 4.24$, $P = 0.005$).

Table II. Means and Standard Deviations of Number of Fixations During the Phases of Flight According to the Fixation Locations and the Pilots' Expertise.

FL	EXPERTISE	PHASE: TAKEOFF	CRUISE	APPROACH	LANDING
OTW	--	9.90 (4.38)	19.25 (9.56)	39.05 (7.15)	11.55 (4.95)
ITC	--	0.30 (6.57)	35.40 (14.91)	4.25 (2.95)	1.33 (2.31)
OTW	Novice	11.63 (8.69)	14.38 (14.62)	41.13 (16.85)	15.38 (9.70)
OTW	Experts	8.75 (3.35)	22.50 (10.75)	37.66 (6.02)	9.00 (2.40)
ITC	Novice	2.50	44.25 (22.01)	6.38 (5.74)	0
ITC	Experts	4.75 (7.50)	29.5 (13.76)	2.83 (1.89)	2.00 (2.82)

FL: fixation location; OTW: out the window; ITC: inside the cockpit.

The results on OTW number of fixations in expert pilots revealed more fixation during the Approach than the Takeoff [$t(101) = 7.29$, $P < 0.0001$], Cruise [$t(101) = 3.83$, $P = 0.019$], and Landing phases [$t(101) = 7.23$, $P < 0.0001$]. Moreover, fixations recorded OTW in experts were less in the Takeoff phase when compared to the Cruise phase [$t(101) = 3.54$, $P = 0.024$] as well as in the Landing phase compared to Cruise [$t(101) = 3.47$, $P = 0.030$]. No difference between the Takeoff and Landing phases emerged [$t(101) = 0.063$, $P = 1$].

In novice pilots, the results were slightly different. Fixations OTW were more in the Approach compared to Takeoff [$t(101) = 6.21$, $P < 0.0001$], Cruise [$t(101) = 5.63$, $P < 0.0001$], and Landing [$t(101) = 5.42$, $P = 0.0001$]. However, between the Takeoff and Cruise [$t(101) = 0.57$, $P = 1$], Takeoff and Landing [$t(101) = 0.772$, $P = 1$], and Cruise and Landing [$t(101) = 0.21$, $P = 1$] phases no differences emerged.

Interaction of Fixation Location × Flight Condition was significant [$F(1, 109) = 6.43$, $\eta_p^2 = 0.017$, $P = 0.013$]. Post hoc analysis revealed that according to task demand, the number of fixations OTW and ITC changed. In particular, the number of fixations ITC (M = 15.35; SD = 17.46) in the low demand task condition was smaller than OTW (M = 22.18; SD = 15.25) [$t(104) = -4.51$, $P < 0.001$]. The number of fixations ITC in the low demand task condition was smaller than OTW in the high demand task (M = 17.70; SD = 14.29) [$t(104) = -3.01$, $P = 0.02$]. The number of fixations ITC in the high demand task (M = 18.96; SD = 21.03) was smaller than OTW in the low demand task [$t(106) = -3.23$, $P = 0.008$]. The other comparisons did not reveal any significant difference ($P > 0.05$).

Data analysis on Fixation Duration showed that both Phase [$F(3, 104.63) = 4.62$, $\eta_p^2 = 0.115$, $P = 0.004$] and Fixation Location [$F(1, 105.15) = 6.76$, $\eta_p^2 = 0.060$, $P = 0.011$] factors were significant. Specifically, fixations during the Approach phase were longer than fixations made in the Takeoff [$t(101) = 2.97$, $P = 0.019$] and in the cruise [$t(101) = 3.18$, $P = 0.010$] phases. No difference between the approach and landing phases was found ($P > 0.05$). Regardless the phases, pilots' fixations were longer OTW then ITC [$t(2185.6) = -8.65$, $P < 0.001$]. Averages and standard deviations are reported in Table I.

The single factors of Expertise [$F(1, 15.51) = 0.30$, $\eta_p^2 = 0.003$, $P = 0.59$] and Flight Condition [$F(1, 101.73) = 0.72$, $\eta_p^2 = 0.005$, $P = 0.40$] were not significant. The interaction of Phase × Fixation Location [$F(3, 104.65) = 3.89$, $\eta_p^2 = 0.074$, $P = 0.011$] revealed differences in term of fixation's duration between the OTW (M = 980.07; SD = 602.55) and the ICT fixations (M = 574.23; SD = 497.55) made during Approach [$t(102) = 4.50$, $P = 0.0003$]. The comparisons are reported in **Fig. 2**. Although fixations OTW made during Approach were longer than fixations OTW made during Takeoff [M = 684.48; SD = 520.30; $t(101) = 4.88$, $P = 0.0001$] and Cruise [M = 653.57; SD = 531.77; $t(101) = 4.61$, $P = 0.0002$], no differences in the other phases were found ($P > 0.05$). Furthermore, no difference emerged between the Approach and Landing phases in OTW fixations [$t(101) = 2.059$, $P = 0.67$].

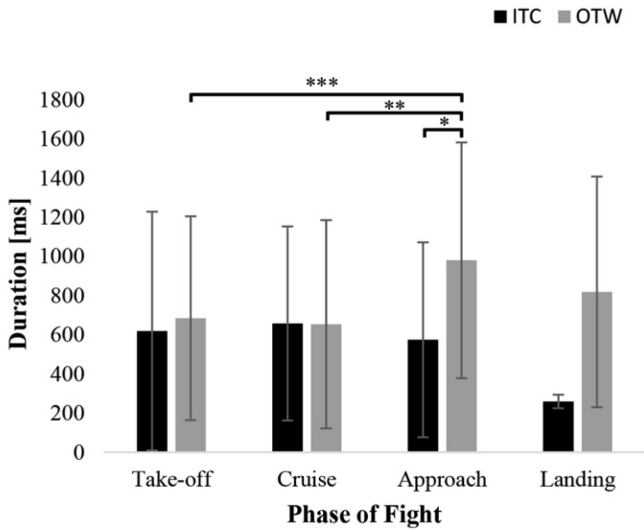

Fig. 2. Means and standard deviations of Fixation Duration ITC and OTW across the Phases of Flight. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

The interaction of Flight Condition × Fixation Location [$F(1, 101.73) = 5.017$, $\eta_p^2 = 0.032$, $P = 0.027$] showed longer OTW fixations (M = 860.11; SD = 600.72) compared to ICT fixations (M = 626.27; SD = 483.34) in the low demand task condition. No differences in the high demand task condition were found ($P > 0.05$). Moreover, no differences emerged from the analysis of the OTW and ICT fixation duration ($P > 0.05$). **Fig. 3** shows the averages of ITC and OTW fixation durations in different conditions.

The triple interaction Expertise × Fixation Location × Phase [$F(3, 104.69) = 2.78$, $\eta_p^2 = 0.073$, $P = 0.045$] was significant. Post hoc analysis revealed that OTW fixations of expert pilots were longer in compared the fixations made during Takeoff, Cruise, and Approach [$t(101) = 3.53$, $P = 0.025$; $t(101 = 3.58$, $P = 0.021$, respectively]. No differences between the Takeoff and Landing phase [$t(101) = 1.00$, $P = 1$], Cruise and Landing phase [$t(101) = 1.05$, $P = 1$], and Approach and Landing phase [$t(101) = 2.53$, $P = 0.51$) were found. OTW fixations of novice pilots made in the Approach phase were longer than fixation made in the Takeoff phase [$t(101) = 3.42$, $P = 0.036$]. No differences were found between Takeoff and Cruise [$t(101) = 0.39$, $P = 1$], Takeoff and Landing [$t(101) = 2.82$, $P = 0.22$], Approach and Cruise [$t(101) = 0.12$], Cruise and Landing [$t(101) = 2.44$, $P = 0.65$], and Approach and Landing [$t(101) = 0.59$, $P = 1$].

Data analysis on subjective evaluation of visual scanning percentage OTW revealed a statistically significant effect of Flight Condition [$F(1, 77) = 6.00$, $P = 0.02$] and Phase of Flight [$F(3, 77) = 22.44$, $P < 0.001$]. The pilots' evaluation of the time they spent gazing OTW was higher in the low demand task condition than in the high demand task condition. Additionally, pilots estimated the percentage of time they spent looking OTW as higher during Landing than in Takeoff [$t(77) = -3.260$, $P = 0.008$], in Cruise [$t(77) = -8.15$, $P < 0.001$], and in Approach [$t(77) = -3.667$, $P = 0.002$]. The estimations were different between Cruise and Approach [$t(77) = -4.482$, $P < 0.001$] and Cruise and Takeoff [$t(77) = 4.890$, $P < 0.001$] (see **Table III**). The two- and threefold interactions did not reveal any significant differences ($P > 0.05$).

The analysis of variance revealed statistical differences between the subjective (M = 70.42; SD = 19.89) and objective (M = 83.16; SD = 26.79) percentages [$F(1, 87.38) = 45.40$, $P < 0.001$]. The analysis of deviation between the subjective and objective percentage of fixation OTW revealed significant differences for the single factor Phase [$F(1, 80) = 18.07$, $P < 0.001$]. Pilots underestimated the time spent fixating OTW in Takeoff, Approach, and Landing, whereas they overestimated it in the Cruise phase. The single factors Expertise [$F(1, 80) = 0.24$, $P = 0.62$] and Flight Condition [$F(1, 80) = 1.41$, $P = 0.23$], and the twofold interactions were not significant ($P > 0.05$). The discrepancy values are reported in Table III.

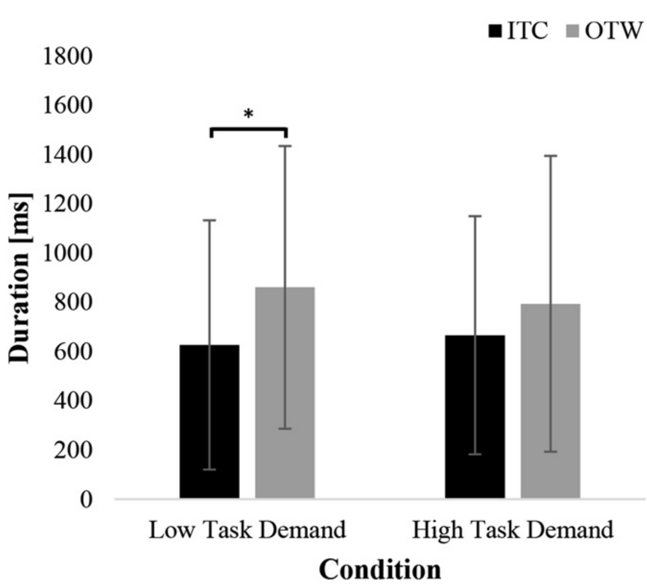

Fig. 3. Means and standard deviations of Fixation Duration ITC and OTW in high and low-demand task conditions. *$P < 0.05$.

DISCUSSION

The present research examined mental workload and visual scanning techniques of expert and novice helicopter pilots in

Table III. Means Percentage and Standard Deviations of Subjective Perception of Time Spent Scanning OTW and Deviation in Visual Scanning During Different Phases of Flight.

	SUBJECTIVE EVALUATION (%)	DISCREPANCY (%)
Condition		
Low-demand task condition	73.96 (17.23)	--
High-demand task condition	66.88 (21.85)	--
Phase		
Takeoff	72.50 (18.24)	−26.00 (19.63)
Cruise	52.50 (18.47)	10.05 (22.30)
Approach	70.83 (14.72)	−21.50 (16.28)
Landing	85.83 (12.48)	−13.51 (12.51)

OTW: out the window.

two experimental conditions with different levels of difficulty/complexity, as well as assessing differences in pilots' performance. Moreover, it analyzed differences in various phases of the flight simulation.

Task demands were manipulated in the flight simulator. Pilots had to takeoff from a ship and they had to land on the same ship after a flight of around 5 min. Similar simulations were involved in training the pilots. The low demand task flight was in good weather conditions and during the day, while the difficult flight consisted of a scenario where pilots had to maneuver the helicopter with rough seas and during the night. Subjective (i.e., mental workload) measures were collected during the experiment and an eye-tracker device was used to assess visual scanning behavior, following the recommendations of previous research.[18]

We hypothesized that experts would report a lower mental workload compared with novices, while higher mental workload scores would be reported by both groups in the difficult compared to the easy condition. The assumption of an interaction between Expertise and Condition has been supported.

Our results also indicate that low visibility conditions and rough seas affect the pilots' perceived mental workload and that expertise plays an important role in mitigating such effect, similar to what has been found.[6] The results show that the perceived mental workload is a function of both the amount of personal resources (expertise levels) and the evaluation of the environmental constraints.

The negative effect of the flying condition on pilots' performance is also supported by our results, showing that all pilots performed worse when the task demand was higher. It is worth discussing that higher mental workload and worse performance could critically affect the safety of the pilots and their crew, not to mention that they could compromise the results of crucial missions in such a sensitive domain as military operations. This represents both a threat and an opportunity. Modern helicopters and ships are equipped with a plethora of devices and advanced systems that are meant to aid pilots in high mental workload situations. However, the devices might have a detrimental effect on pilots' performance themselves since pilots are prompted to check many systems in a short time, moving their attention to different positions. A new system that gathers all the crucial information needed by pilots in such situations and presents them in a clear nondistractive way could be beneficial for increasing safety.

Results on visual search were partially in line with previous research. No differences in the number of fixations between expert and novice pilots were found, and the duration of fixations was similar. The number of fixations was overall higher OTW than ITC. That was also true in each of the phases, except for the cruise, in which pilots made significantly more fixations ITC compared to OTW.

Moreover, no significant differences were found toward the number and duration of fixations between the approach and landing phases. This data is in line with previous observations.[21] Approaching the ship, the pilot switches to a predominantly external visual flight and the copilot provides assistance monitoring air speed and closure rate. Additionally, the simulation presented a series of visual landing aids such as glide slope indicator, horizon reference bar, and deck reference lights. Therefore, the pilots' visual search behaviors were similar in the approach and landing phases. Pilots can adopt visual flying rules, but still rely on the visual landing aids on the ship.

These data support previous research stating the importance of navigation instruments during the cruise phase. Pilots, when cruising without meaningful visual cues on the exterior, often focus their attention on instruments. Notably, a study found that the experienced pilots better maintained a constant altitude above the ground, which in turn is associated with more fixations ITC than OTW.[13] In this regard, our results support that expert pilots make more fixations ITC during the cruise phase compared with less experienced pilots. The condition is reversed in the following phase, the approach, when pilots need to redirect their visual attention toward the ship, align with its rear, and set the descending angle based on the Glide Slope Indicator (which is a visual reference lighting system that provides the pilot with a visual cue for the right angle of the descent toward the ship). Our results also suggest that this is true for both novices and experts. The situation remains vastly similar even in the landing phase when pilots focus on external clues.

Considering the task demand, our results are in line with previous research. We found that the condition affects pilots' performance and the duration of fixations.[19] Additionally, task difficulties do not affect the amount of fixation.[27]

Furthermore, results on fixation duration showed that OTW fixations are significantly longer than ITC fixations. This result suggests that the most crucial info for pilots is gathered through looking outside of the cockpit. Critical information ITC (e.g., altitude) is still needed for the pilot, but s/he has to move attention to the instrumentation and then again OTW. This could entail the onset of spatial disorientation phenomena, which are deemed a relevant risk.[7] Generally, the pilots' skills in monitoring the environment and the helicopter's instruments are essential factors that affect both the decision-making and the safety of the flight. Designing innovative systems, such as head-up display, that allow pilots to gather such information while still looking out of the window, reducing switches of visual attention, is desirable.

Considering the discrepancy between the objective and subjective estimation of fixations OTW, the findings support the belief that subjects are not always able to correctly assess their scanning techniques. Metacognition about visual acquisition patterns should be beneficial for decision-making and situation awareness; the pilots' underestimation of the environmental clues and the visual landing aids should be taken into consideration when performance is subjectively assessed. Furthermore, our findings suggest a more critical consideration of subjects' self-reported data, e.g. by interviews.[16,21] Thus, the combination of both subjective and objective measurements is recommended to analyze the users' needs and the system's requirements.

The present study has some limitations. Although past research used fewer subjects to study pilots' eye fixations,[4] our sample consisted of only 12 male pilots. Nevertheless, our findings are similar to those found in another study, which included more than 30 subjects.[18] Regarding technical aspects, it is important to mention that the eye tracker did not perform well in the low light (high task demand) condition as in the daylight (low task demand) condition. Additionally, although the cutoff of 1500 flight hours was agreed to by the instructor, a more distinct criterion would be beneficial to better understand the differences in expert and novice pilots. Despite the study limitations we believe our findings provide a valuable insight into the pilots' experience and represent an important contribution to this still under-researched field of study.

Regarding future studies, there is a need to investigate the relationship between mental workload and physiological data in expert and novice pilots. Also, a study by Sullivan et al. revealed that more experienced pilots had shorter gaze duration and more frequent saccades between the ITC and OTW.[20] Future studies could focus on saccade measurements in order to understand differences in attention switching between expert and novice pilots.

The present study contributes to deepening the knowledge regarding mental workload and gaze behavior of novice and expert helicopter pilots landing on ships. Implications are relevant for organizations involved in developing systems and interfaces to reduce pilots' mental workload, improving their visual scanning behavior, and ultimately increasing the pilot's safety and operations success rate.

ACKNOWLEDGMENTS

Financial Disclosure Statement: The authors have no competing interests to declare.

Authors and affiliations: Giuseppe Rainieri, M.S., Federico Fraboni, M.S., Martin Tušl, M.S., Alessia Tessari, Ph.D, and Luca Pietrantoni, Ph.D, Alma Mater Studiorum, Department of Psychology, University of Bologna, Bologna, Italy; Gabriele Russo, M.S., Alma Mater Studiorum, Department for Life Quality Studies, University of Bologna, Rimini, Italy; and Andrea Pingitore, M.S., Italian Navy, Ministry of Defence, Maristaeli Luni, Sarzana, Italy.

REFERENCES

1. Bellenkes AH, Wickens CD, Kramer AF. Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. Aviat Space Environ Med. 1997; 68(7):569–579.
2. Carico D, Ferrier B. Evaluating landing aids to support helicopter/ship testing and operations. IEEE Aerospace Conference; 2006. New York: IEEE; 2006.
3. Di Nocera F, Camilli M, Terenzi M. A random glance at the flight deck: pilots' scanning strategies and the real-time assessment of mental workload. J Cogn Eng Decis Mak. 2007; 1(3):271–285.
4. Di Nocera F, Ranvaud R, Pasquali V. Spatial pattern of eye fixations and evidence of ultradian rhythms in aircraft pilots. Aerosp Med Hum Perform. 2015; 86(7):647–651.
5. European Aviation Safety Agency (EASA). Annual safety recommendations review. 2010; [Accessed Aug. 2020]. Available from https://www.easa.europa.eu/document-library/general-publications/annual-safety-recommendations-review-2010.
6. Gaetan S, Dousset E, Marqueste T, Bringoux L, Bourdin C, et al. Cognitive workload and psychophysiological parameters during multitask activity in helicopter pilots. Aerosp Med Hum Perform. 2015; 86(12):1052–1057.
7. Gibb R, Ercoline B, Scharff L. Spatial disorientation: decades of pilot fatalities. Aviat Space Environ Med. 2011; 82(7):717–724.
8. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Advances in psychology, 52. Human mental workload. Amsterdam: North-Holland; 1988:139–183.
9. Hodge SJ, Forrest JS, Padfield GD, Owen I. Simulating the environment at the helicopter-ship dynamic interface: research, development and application. The Aeronautical Journal. 2012; 116(1185):1155–1184.
10. Hoencamp A, Holten T, Prasad VR. Relevant aspects of helicopter-ship operations. 34th European Rotorcraft Forum; 16–19 Sept. 2008; Liverpool, UK. London (UK): Royal Aeronautical Society; 2008.
11. Kasarskis P, Stehwien J, Hickox J, Aretz A, Wickens C. Comparison of expert and novice scan behaviors during VFR flight. In: Proceedings of the 11th International Symposium on Aviation Psychology, vol. 6; March 5–8, 2001; Columbus, OH, USA. Columbus (OH): Department of Aviation and Aeronautical Engineering, Columbus State University; 2001.
12. Kassner M, Patera W, Bulling A. Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. New York: ACM; 2014:1151–1160.
13. Kirby CE, Kennedy Q, Yang JH. Helicopter pilot scan techniques during low-altitude high-speed flight. Aviat Space Environ Med. 2014; 85(7):740–744.
14. Lee D, Horn JF. Simulation of pilot workload for a helicopter operating in a turbulent ship airwake. Proceedings of the Institution of Mechanical Engineers, Part G. J Aerosp Eng. 2005; 219(5):445–458.
15. Lounis C, Peysakhovich V, Causse M. 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; March 17, 2020; Toulouse, France. Zurich (Switzerland): geoGAZElab at ETH Zurich; 2020.
16. Minotra D, Feigh K. Studying pilot cognition in ship-based helicopter landing maneuvers. Proceedings of the American Helicopter Society International Forum 74; May 14–17, 2018; Phoenix, AZ, USA. Fairfax (VA): Vertical Flight Society; 2018.
17. Peißl S, Wickens CD, Baruah R. Eye-tracking measures in aviation: a selective literature review. Int J Aerosp Psychol. 2018; 28(3–4): 98–112.
18. Robinski M, Stein M. Tracking visual scanning techniques in training simulation for helicopter landing. J Eye Mov Res. 2013; 6(2):1–17.
19. Schriver AT, Morrow DG, Wickens CD, Talleur DA. Expertise differences in attentional strategies related to pilot decision making. Hum Factors. 2008; 50(6):864–878.
20. Sullivan J, Yang JH, Day M, Kennedy Q. Training simulation for helicopter navigation by characterizing visual scan patterns. Aviat Space Environ Med. 2011; 82(9):871–878.
21. Tušl M, Pietrantoni L, Fraboni F, De Angelis M, Rainieri G. Helicopter pilots' tasks, cognitive workload, and the role of external visual cues during shipboard landing. J Cogn Eng Decis Mak. 2020; 14(3): 242–257.
22. van de Merwe K, van Dijk H, Zon R. Eye movements as an indicator of situation awareness in a flight simulator experiment. Int J Aviat Psychol. 2012; 22(1):78–95.
23. Wang Y, White M, Owen I, Hodge S, Barakos G. Effects of visual and motion cues in flight simulation of ship-borne helicopter operations. CEAS Aeronautical Journal. 2013; 4(4):385–396.
24. Wickens CD. Situation awareness and workload in aviation. Curr Dir Psychol Sci. 2002; 11(4):128–133.
25. Yang JH, Kennedy Q, Sullivan J, Fricker RD. Pilot performance: assessing how scan patterns & navigational assessments vary by flight expertise. Aviat Space Environ Med. 2013; 84(2):116–124.

26. Yu CS, Wang EMY, Li WC, Braithwaite G, Greaves M. Pilots' visual scan patterns and attention distribution during the pursuit of a dynamic target. Aerosp Med Hum Perform. 2016; 87(1):40–47.
27. Zhang X, Xue H, Qu X, Li T. Can fixation frequency be used to assess pilots' mental workload during taxiing? In: Harris D, editor. Engineering psychology and cognitive ergonomics: performance, emotion and situation awareness. EPCE; July 9–14, 2017; Vancouver, BC, Canada. Cham (Switzerland): Springer International Publishing AG; 2017. https://doi.org/10.1007/978-3-319-58472-0_7
28. Zimmer A, Stein M. Information systems in transportation. In: Information Ergonomics. Berlin (Germany): Springer; 2012:1–22. https://doi.org/10.1007/978-3-642-25841-1_1
29. Ziv G. Gaze behavior and visual attention: a review of eye tracking studies in aviation. Int J Aviat Psychol. 2016; 26(3–4):75–104.
30. Ziv G. The need for eye tracking studies in helicopter pilots: a position stand. In: 1st International Workshop on Eye-Tracking in Aviation; March 17, 2020; Toulouse, France. Zurich (Switzerland): geoGAZElab at ETH Zurich; 2020.