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

Published Version:
Visual Scanning Techniques and Mental Workload of Helicopter Pilots During Simulated Flight / Rainieri, Giuseppe; Fraboni, Federico; Russo, Gabriele; Tul, Martin; Pingitore, Andrea; Tessari, Alessia; Pietrantoni, Luca. - In: AEROSPACE MEDICINE AND HUMAN PERFORMANCE. - ISSN 2375-6314. - ELETTRONICO. - 92:1(2021), pp. 11-19. [10.3357/AMHP.5681.2021]

This version is available at: https://hdl.handle.net/11585/788331.2 since: 2021-01-12

Published:

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

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

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.

The final published version is available online at:

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

Title Page

Manuscript title: Visual Scanning Techniques and Mental Workload of Helicopter Pilots during Simulated Flight

Authors/Co-author names and affiliations

- Giuseppe Rainieri – University of Bologna, Department of Psychology, via Filippo Re 10, 40126, Bologna, Italy
- Federico Fraboni – University of Bologna, Department of Psychology, via Filippo Re 10, 40126, Bologna, Italy
- Gabriele Russo – University of Bologna, Department for Life Quality Studies, C.so D'Augusto 237, 47921, Rimini, Italy
- Martin Tušl – University of Bologna, Department of Psychology, via Filippo Re 10, 40126, Bologna, Italy
- Andrea Pingitore - Centro Sperimentale Aeromarittimo c/o Maristaeli Luni, Stato Maggiore Marina, Italian Navy, via Ghiarettolo 101, Sarzana (SP), Sarzana, It 19038
- Alessia Tessari – University of Bologna, Department of Psychology, via Filippo Re 10, 40126, Bologna, Italy
- Luca Pietrantoni – University of Bologna, Department of Psychology, via Filippo Re 10, 40126, Bologna, Italy

ABSTRACT

Introduction: The visual scanning techniques (VST) used by helicopter pilots are a critical skill to accomplish safe and correct land. According to the human information processing theory, the VST can be analysed 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 low-workload condition, consisted in daylight and good weather simulation, and in high-workload condition, night-time, low-visibility and adverse weather conditions. Twelve helicopter pilots took part in the study.

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 the shorter in 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 instruments 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

INTRODUCTION

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 in 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 area,[10] dynamic environment around the ship,[14] and sea spray[9] are some of the challenges that pilots need to handle during take-off, approaching 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 (VST) 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 inter-individual 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, the pilot '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 concentrations, so-called Fixations 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 take-off 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 compare to novices.[15] Another investigation found that 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 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 helicopter pilots and novice 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 objectively measures and subjectively 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 demands situations, whereas the opposite result was found for high demands 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 assessing the mental workload in the two missions. They noticed greater mental workload for student pilots than for flight instructors, that 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 shown medium level of workload, reporting positive emotions with low emotional intensity. The less-experienced pilots shown 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 inter-individual 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 analysed the pilots 'scanning techniques during a flight on a frigate and a pinnacle.[18] Our experimental setting foresees to 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. Take-off, 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 interviews data are considered.

Therefore, we test our hypothesis 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 hypothesise 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 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 high-demand task condition pilots will have more and shorter fixations than in low task demands;

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.

H5a: the subjective and objective evaluation of pilots' scanning techniques will vary according to task demands and phases of flight.

METHODS

Subjects

Twelve male helicopter pilots recruited from the Italian Navy voluntarily took part in the study. The sample included six novice and six expert pilots. Together with a flight instructor, a cut-off 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. Each subject provided written informed consent before participating. Pilots 9 and 11 were removed from eye-tracking analysis due to low quality of the registrations.

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, flight dynamics were comparable.

Eye-tracking. We used the Pupil Labs[12] head-mounted eye-tracker to collect objective data about the pilots´ visual scanning. The maximum dispersion was set at 1.0 degree. The minimum fixation duration was set at 200 ms, while the maximum was set at 2000 ms. We measured two parameters of visual scanning behaviours: fixation locations (FL) and duration of fixations (DF). Based on previous research,[18,22] we divided the areas of interests (AOI) for FLs into three main categories:

out-of-the-window (OTW), inside on the cockpit instruments (ITC), and Null if the fixation position was not detected correctly.

Subjective workload. We used a smartphone app version of 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., take-off, cruise, approach, landing) from 0 to 100%. The items were administrated after each simulation.

Performance. An experienced flight instructor that participated in all simulations as a co-pilot 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 co-pilot, 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 he flies. 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., low demands situation). The second simulation was a flight at night with low-visibility, unstable ship's deck and difficult weather conditions (i.e., high demands situation). These specific scenarios were chosen in cooperation with the flight instructor to account for low demands and high demands situation. The simulation started with the helicopter on a ship's deck, and the task consisted in a take-off, cruise, approach, and landing on the same ship. The pilot was supported by an experienced co-pilot and a Flight Deck Officer (FDO), that was in the control room and communicated with the pilot via radio. The position of the pilots was always the same, the co-pilot in the left seat and the pilot in the right seat. We divided the simulation into four phases: take-off, cruise, approach and landing. Take-off started at the beginning of the simulation, and it ended when the helicopter completely abandoned 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 on average of 4 min. 49 sec. each of the phases with a different average time: Take-off 24 sec, Cruise 2 minutes 24 sec, Approach 1 min 25 sec, and Landing 29 sec.

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 models (lme4 package) analyses were computed to analyse the data.

Subjective evaluation of mental workload and Pilots' performance were analysed as a function of Expertise (experts vs. novices) and Flight condition (low-demands vs. high-demands).

Eye-tracking data (Number of Fixations and Duration of Fixations) were analysed as a function of Expertise (experts vs. novices), Flight Condition (low-demands vs. high-demands), Phase of the Flight (take-off, cruise, approach and landing) and Fixation Locations (OTW vs. ITC).

The subjective evaluation of visual scanning was analysed as a function of Expertise (experts vs. novices), Flight condition (low-demands vs. high-demands), and Phase of the flight (take-off, cruise, approach and landing).

For multiple comparisons, Bonferroni post-hoc analyses were performed. Moreover, stepwise regression was involved in modelling the linear missed regression model in order to have the best model for each dependent variable.

Furthermore, data collected through 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 x (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 = .055) and significant differences for Flight Condition (F(1, 19) = 7.33, $\eta_p^2 = 0.40$, p < .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 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 X Flight Condition (F (1, 19) = 5.61, $\eta_p^2 = 0.34$, p < .042) was significant. Post-hoc analysis showed that in low-demand task condition expert pilots reported a lower mental workload compared to novices (t (12.92) = 2.97, p = .047). Moreover, expert pilots reported a higher mental workload in high-demand task condition compared to low-demand task condition (t (9.02) = 3.74, p = .020). In high-demand task condition both the groups reported similar mental workload (t (13.75) = 1.073, p = .71). Figure 1 shows the means and standard errors for the score of NASA-TLX according to Expertise and Flight Condition.

[Fig. 1 Here]

Performance evaluation

The evaluation of the performance assessed by the flight instructor revealed significant differences for the factor Flight Condition ($F (1, 10) = 13.75$, $\eta_p^2 = 0.55$, $p = .004$) 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 ($F (1, 9) =1.92$, $\eta_p^2 = 0.15$, $p = .20$) nor the interaction Expertise X Flight Condition were significant ($p > .05$).

Visual search data

Analysis of number of fixations showed significant differences for Fixations Location ($F (1, 109) = 15.97$, $\eta_p^2 = 0.04$, $p > .0000$) and Phase ($F (3, 109) =26.28$, $\eta_p^2 = 0.2$, $p < .000$). See Table I. Factors Expertise ($F (1, 109) =0.89$, $\eta_p^2 = 0.002$, $p = .35$) 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 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 take-off phase, no differences between the fixations ITC ($t (109) = 1.50$, $p = .80$) were found. In the cruise phase, pilots made more fixations ITC compared to OTW ($t (101) = 6.014$, $p < .0001$). In contrast, pilots fixated more OTW compared to ITC in the approach phase ($t (108) = 10.99$, $p < .0001$), while no difference during the landing ($t (108) = 1.95$, $p = .52$) was found. The means and standard deviations are shown in Table I.

[Table I Here]

The interactions Phase X Fixation Location (F (1, 109) = 48.95, η_p^2 = 0.40, p = .001) was significant. Post-hoc analysis revealed differences between OTW and ITC fixation in Cruise and Approach phases (t (101) = 6.01, p < .0001; t (101) = 11.00, p < .0001, respectively). No differences emerged in Take-off and Landing phases (t (109) = 1.50, p = 1; t (108) = 1.95, p = .87, respectively).

[Table II Here]

Triple interaction Fixation Location X Phase X Expertise was significant (F (3, 109) = 3.85, η_p^2 = 0.031, p = .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 = .084) was different. No differences emerged in the other phases (Take-off: t (103 = 0.13, p =1; Approach: t (99.7) = 0.64, p = 1; Landing: t (108.9) = 3.40, p 1) as well as between the groups for the number of fixations made OTW (Take-off: 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 Take-off phase subjects fixated less ITC than in Cruise phase (t = 4.65, p = .0004) and the number of fixations ITC was higher in Cruise than Approach (t = 6.56, p <.0001) and Landing phases (t = 3.66, p = .031).

Results were almost similar for novice pilots. The number of fixations ITC was higher in cruise when compared to Take-off ($t = 5.31$, $p = .0001$), Approach ($t = 7.66$, $p < .0001$) and Landing phases ($t = 4.24$, $p = .005$)

The results on OTW number of fixations in expert pilots revealed more fixation during the approach than the take-off ($t (101) = 7.29$, $p < .0001$), cruise ($t (101) = 3.83$, $p = .019$) and landing phases ($t (101) = 7.23$, $p < .0001$). Moreover, fixations recorded OTW in experts were less in the Take-off phase when compared to the cruise phase ($t (101) = 3.54$, $p =.024$) as well as in landing phase compared to cruise one ($t (101) = 3.47$, $p = .030$). No difference between take-off and landing phase emerged ($t (101) = 0.063$, $p = 1$).

In novice pilots, the results were slightly different. Fixations OTW were more in the approach compared to take-off ($t (101) = 6.21$, $p < .0001$), cruise ($t (101) = 5.63$, $p < .0001$) and landing ($t (101) = 5.42$, $p = .0001$) phases. While between take-off and cruise ($t (101) = 0.57$, $p = 1$), take-off and landing (t 101) = 0.772, $p = 1$) and cruise and landing ($t (101) = 0.21$, $p = 1$) phases no differences emerged.

Interaction Fixation Location X Flight Condition was significant ($F (1, 109) = 6.43$, $\eta_p^2 = 0.017$, $p = .013$). 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 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 low-demand task condition was smaller than OTW in 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 low-demand task ($t (106) = -3.23$, $p = 0.008$). The other comparisons did not reveal any significant difference ($p > .05$).

Data analysis on Fixation Duration showed that both Phase (F (3, 104.63) = 4.62, $\eta_p^2 = 0.115$, p = .004) and Fixations Location (F (1, 105.15) = 6.76, $\eta_p^2 = 0.060$, p = .011) factors were significant. Specifically, fixations during the approach phase were longer than fixations made in the take-off (t (101) = 2.97, p = .019) and in the cruise (t (101) = 3.18, p =.010) phases. No difference between the approach and the landing phase was found (p > .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.

Single factors Expertise (F (1, 15.51) = 0.30, $\eta_p^2 = 0.003$, p = .59) and Flight condition (F(1, 101.73) = 0.72, $\eta_p^2 = 0.005$, p = .40) were not significant.

The interaction Phase X Fixation Location (F (3, 104.65) = 3.89, $\eta_p^2 = 0.074$, p = .011) revealed differences in term of fixation's duration between the OWT (M = 980.07; SD = 602.55) and the ICT fixations (M = 574.23; SD = 497.55) made in approach (t (102) = 4.50, p = .0003). The comparisons are reported in Figure 2. Although fixations OTW made during approach were longer than fixations OTW made during take-off (M = 684.48; SD = 520.30; t (101) = 4.88, p = .0001) and cruise (M = 653.57; SD = 531.77; t (101) = 4.61, p = .0002), no differences in the other phases were found (p > .05). Furthermore, no difference emerged between approach and landing phase in OTW fixations (t (101) = 2.059, p = .67).

[Fig. 2 Here]

The interaction Flight Condition X Fixation Location (F (1, 101.73) = 5.017, η_p^2 = 0.032, p = .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 high-demand task condition were found ($p > .05$). Moreover, no differences emerged from the analysis of the OTW and the ICT fixations duration ($p > .05$). Figure 3 shows the averages of ITC and OTW fixations 'durations in different conditions.

[Fig. 3 Here]

The triple interaction Expertise X Fixation Location X Phase (F (3, 104.69) = 2.78, η_p^2 = 0.073, p = .045). was significant.

Post-hoc analysis revealed that OTW fixations of expert pilots were longer in compared the fixations made in take-off and cruise, approach (t (101) = 3.53, p = .025; t (101 = 3.58, p = .021, respectively) phases. No differences between take-off 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 = .51) were found. OTW fixations of novice pilots made in Approach phase were longer than fixation made in take-off phase (t (101) = 3.42, p = .036). Non differences between take-off and cruise phases (t (101) = 0.39, p = 1), take-off and landing phases (t (101) = 2.82, p = .22), Approach and cruise phases (t (101) = 0.12), cruise and landing phases (t (101) = 2.44, p = .65) and approach and landing phases (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 evaluated that 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 stare OTW was higher during landing, than in take-off ($t(77) = -3.260$, $p = 0.008$), 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 take-off ($t(77) = 4.890$, $p < 0.001$). See Table III. The two- and three-fold interactions did not reveal any significant difference ($p > .05$).

[Table III Here]

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 subjective and objective percentage of fixation OTW revealed significant differences for the single factor Phase ($F(1, 80) = 18.07$, $p < .001$). Pilots underestimated the time spent fixating OTW in take-off, approach and landing, whereas overestimated it in cruise phase. Single factor Expertise ($F(1, 80) = 0.24$, $p = .62$) and Flight condition ($F(1, 80) = 1.41$, $p = .23$), and the two-fold interactions were not significant ($p > .05$). The discrepancy values are reported in table III.

DISCUSSION

The present research aimed to examine mental workload and visual scanning techniques of expert and novice helicopter pilots in two experimental conditions with different levels of difficulty/complexity, as long as assessing differences in pilots 'performance. Moreover, it analysed differences in various phases of the flight simulation.

Task demands were manipulated in the flight simulator. Pilots had to take-off from a ship, and they had to land on the same ship after a flight of around 5 minutes. Similar simulations were involved in training the pilots. The low-demand task flight was in good weather condition and during the day, while the difficult flight consisted in a scenario where pilots had to maneuver the helicopter with rough sea 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 behaviour, following the recommendation previous research.[18]

We hypothesised 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 interaction between Expertise and Condition has been supported.

Our results also indicate that low visibility conditions and rough sea 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 confer 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 to discuss 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 are. 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 non-distractive 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 is overall higher OTW than ITC. That is true also considering the phases, except for the cruise, in which pilots made significantly more fixations ITC compared to OTW.

Moreover, no significant differences were found towards the number and duration of fixations between approach and landing phases. This data is in line with previous observation.[21] Approaching the ship, the pilot switches to a predominantly external visual flight, and the co-pilot 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, deck reference lights. Therefore, the pilots’ visual search behaviour were similar in approach and landing phases. Pilots can adopt the visual flying rules (VFR) but still relying on the VLAs on the ship.

This data supports previous research stating the importance of navigation instruments during the cruise phase. Pilots, when cruising, often remain without meaningful visual cues on the exterior, thus they 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 confronted with less experienced pilots. The condition is reverted in the following phase, the approach, when pilots need to redirect their visual attention towards the ship, align with its rear, and set the descending angle basing 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 towards 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 researches. 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, as head-up display (HUD), that allow the 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. The metacognition about the visual acquisition patterns should be beneficial for decision-making and situation awareness; the pilots' underestimation of the environmental clues and the Visual Landing Aids is to take in consideration when the performance is subjectively assessed. Furthermore, our findings suggest a more critic consideration of subjects ’self-reported data, i.g. by interviews.[16,21] Thus, the combination of both subjective and objective measurements is recommended to analyse the users' needs and the system's requirements.

The present study has some limitations. Although past research used fewer subjects to study pilot’s 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 thirty subjects[18]. Regarding technical aspects, it is important to mention that the eye-tracker was not performing well in low light (high task demands) condition as in daylight (low task demands) condition. Additionally, although the cut-off on 1500 flight hours was agreed by the instructor, a more outdistance 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 behaviour of novice and expert helicopter pilots landing on ships. Implications are relevant for organisations involved in developing systems and interfaces to reduce pilots 'mental workload, improving their visual scanning behaviour and ultimately increasing the pilot's safety and operations success rate.

ACKNOWLEDGMENTS

Authors and affiliations: Giuseppe Rainieri, M.S., Federico Fraboni, M.S., Martin Tušl, M.S., Alessia Tessari, Ph.D, and Luca Pietrantoni, Ph.D, Department of Psychology, University of Bologna, Bologna, Italy; Gabriele Russo, M.S., Department for Life Quality Studies, University of Bologna, Rimini, Italy; and Andrea Pingitore, M.S., Centro Sperimentale Aeromarittimo c/o Maristaeli Luni, Stato Maggiore Marina, Italian Navy, Italian Navy, Sarzana, Italy.

REFERENCES

1. Bellenkes AH, Wickens CD, Kramer AF. Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. Aviation, space, and environmental medicine. 1997; 68(7), 569-579.
2. Carico D, Ferrier B. Evaluating Landing Aids to Support Helicopter/Ship Testing and Operations. IEEE Aerospace Conference. 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. Journal of Cognitive Engineering and Decision Making. 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. Aerospace medicine and human performance. 2015; 86(7), 647-651.
5. European Aviation Safety Agency (EASA). Annual Safety Recommendations review. Retrieved from https://www.easa.europa.eu/sites/default/files/dfu/safety-and-research-docs-safety-recommendations-review-2010-2010---Annual-Safety-Recommendations-Review.pdf. 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. Aerospace medicine and human performance. 2015; 86(12), 1052-1057.
7. Gibb R, Ercoline B, Scharff L. Spatial disorientation: decades of pilot fatalities. Aviation, space, and environmental medicine. 2011; 82(7), 717-724.

8. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology. 1988; 52, 139-183. Doi: 10.1016/S0166-4115(08)62386-9.

9. Hoencamp A, Holten T, Prasad VR. Relevant aspects of helicopter-ship operations. 34th European Rotorcraft Forum. 2008.

10. Hodge SJ, Forrest JS, Padfield GD, Owen I. Simulating the environment at the helicopter-ship dynamic interface: research, development and application. Aeronautical Journal. 2012; 116(1185), 1155.

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). 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. 2014; (pp. 1151-1160). ACM.

13. Kirby CE, Kennedy Q, Yang, JH. Helicopter pilot scan techniques during low-altitude high-speed flight. Aviation, space, and environmental medicine. 2014; 85(7), 740-744.

14. Lee D, Horn JF. Simulation of pilot workload for a helicopter operating in a Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering. 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. 2020.

16. Minotra D, Feigh K. Studying Pilot Cognition in Ship-Based Helicopter Landing Maneuvers. Proceedings of the American Helicopter Society International Forum 74. 2018.

17. Peißl S, Wickens CD, Baruah R. Eye-tracking measures in aviation: a selective literature review. The International Journal of Aerospace Psychology. 2018; 28(3-4), 98-112.

18. Robinski M, Stein M. Tracking visual scanning techniques in training simulation for helicopter landing. Journal of Eye Movement Research 2013; 6(2), 1-17.

19. Schriver AT, Morrow DG, Wickens CD, Talleur DA. Expertise differences in attentional strategies related to pilot decision making. Human 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. Aviation, space, and environmental medicine. 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. Journal of Cognitive Engineering and Decision Making. In press.

22. van de Merwe K, van Dijk H, Zon R. Eye movements as an indicator of situation awareness in a flight simulator experiment. The International Journal of Aviation Psychology. 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 Journal. 2013; 4(4), 385-396

24. Wickens CD. Situation awareness and workload in aviation. Current directions in psychological science. 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. Aviation, space, and environmental medicine. 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. Aerospace medicine and human performance. 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. (eds) Engineering Psychology and Cognitive Ergonomics: Performance, Emotion and Situation Awareness. 2017; EPCE 2017.

28. Zimmer A, Stein M. Information Systems in Transportation. In Information Ergonomics. 2012; (pp. 1-22). Springer Berlin, Heidelberg.

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

30. Ziv, G. (2020). The Need for Eye Tracking Studies in Helicopter Pilots: A Position Stand. In 1st International Workshop on Eye-Tracking in Aviation.

TABLES

		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	Take-off	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)

Table I. – Means and standard deviations of Number of Fixations and Duration of Fixations according to Location and Phase.

		Phase			
		Take-off	Cruise	Approach	Landing
FL	Expertise				
OTW	-	9.90 (4.38)	19.25 (9.56)	39.05 (7.15)	11.55 (4.95)
ITC	-	.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)

Table II. – Means and standard deviations of Number of Fixations during the Phases of Flight according to the Fixation Locations and the pilots ’Expertise.

		Subjective Evaluation [%]	Discrepancy [%]
Condition	Low-demand Task condition	73.96 (17.23)	-
	High-demand Task condition	66.88 (21.85)	-
Phase	Take-off	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)

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

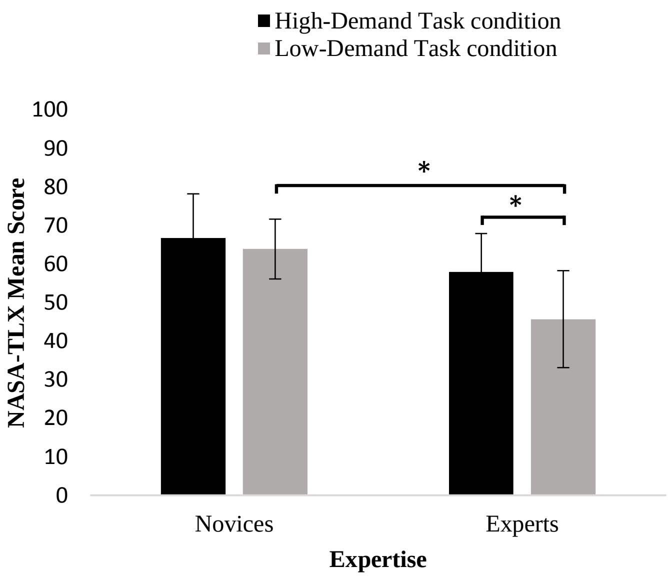
CAPTIONS FOR FIGURES

Figure 1. – Means and standard errors of NASA-TLX scores according to Flight condition and Expertise. *: $p < 0.05$.

Figure 2. – Means and standard deviations of Fixation Duration ITC and OTW across the Phases of Flight. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

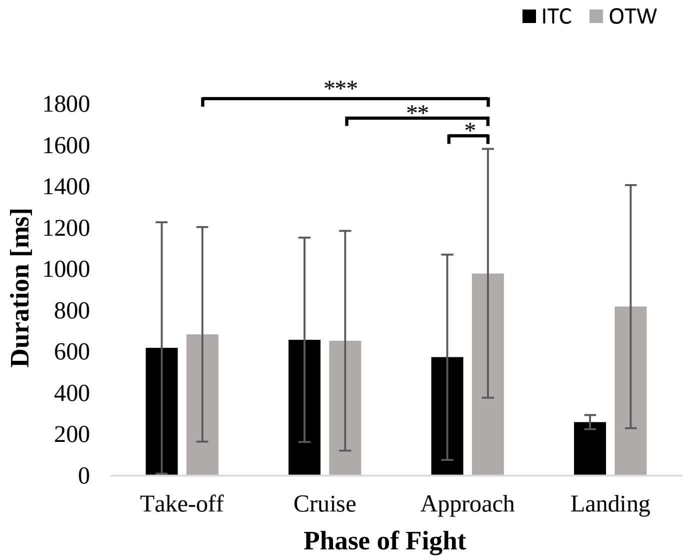
Figure 3. – Means and standard deviations of Fixation duration ITC and OTW in high and low-demand task conditions. *: $p < 0.05$.

Figure 1



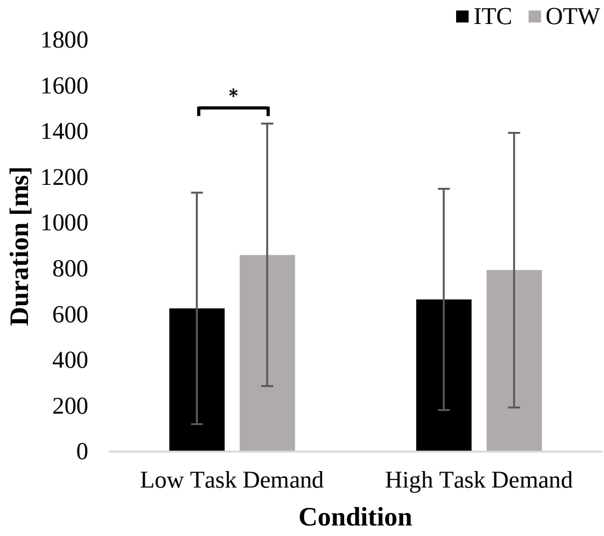
High-Demand Task condition
Low-Demand Task condition
100
90
80
70
60
50
40
30
20
10
0
NASA-TLX Mean Score
*
*
Novices
Experts
Expertise

Figure 2



ITC
OTW
***
**
*
1800
1600
1400
1200
1000
800
600
400
200
0
Duration [ms]
Take-off
Cruise
Approach
Landing
Phase of Fight

Figure 3



ITC
OTW
1800
1600
1400
1200
1000
800
600
400
200
0
Duration [ms]
*
Low Task Demand
High Task Demand
Condition