MEASURING THE LOOKOUT BEHAVIOR OF STUDENT PILOTS IN A VIRTUAL REALITY FLIGHT SIMULATOR

Ivo V. Stuldreher
Erik Van der Burg
Wietse D. Ledegang
Mark M.J. Houben
Eric L. Groen
TNO Human Performance, Soesterberg, The Netherlands
Danny van der Horst
Erik A.M. Starmans
Guido Almekinders
Royal Netherlands Air Force, Soesterberg, The Netherlands

Learning adequate gaze behavior is essential in flight training. In this exploratory study we investigated the development of gaze behavior in flight training in a virtual reality (VR) flight simulator. Following standardized study material, fifteen participants without flying experience repeatedly practiced three basic flight maneuvers in a VR simulator of a small aircraft. During some runs, participants performed an additional N-back task to measure cognitive spare capacity. From the recorded gaze data we computed the percentage of time during which the gaze was directed outside the cockpit, i.e., the "Lookout". This outside dwell ratio differed between flight maneuvers. A higher outside dwell ratio was associated with better flight performance. Remarkably, the outside dwell ratio increased with the additional N-back task. A heatmap indicated staring behavior during the N-back. In a follow-up study we will extend the analysis of gaze behavior with more dynamic measures than only the dwell ratio.

Aspiring military pilots within the Royal Netherlands Air Force (RNLAF) undergo Elementary Military Pilot Training (EMVO) in a turbo-prop trainer aircraft (the Pilatus PC-7). During their flight training, the student pilots also learn how to perform adequate visual scanning during flight, as this is leads to improved flight performance (Ziv, 2016). During the EMVO, emphasis is made on directing a large proportion of gaze to the outside environment. Student pilots are instructed to perform a structured lookout procedure during which they systematically scan the horizon, alternated by brief cross-checks of the relevant flight instruments inside the cockpit.

The RNLAF is interested in the possibility to incorporate Virtual Reality (VR) as training means within the EMVO. VR means have already been implemented by the Royal Air Force, United States Air Force and Royal Australian Air Force (Pope, 2019; Air Education and Training Command, 2020; Lewis & Livingston, 2018; Pennington et al., 2019). Nowadays VR systems have a built-in eye tracker, which allows for the monitoring of gaze behavior of the student pilots. There are indications that pilots whose gaze behavior better corresponds to that of expert pilots show better flight performance (Wickens et al., 2008). Besides learning how to control the aiplane it is thus also important to learn how to direct your gaze during flight.

In this study, we investigated the development of gaze behavior of student pilots in a VR flight simulator during a mini-training of three sessions in which they practiced three basic flight maneuvers. In a related paper, we discuss the flight performance measures (Ledegang et al., in press). Here, we examine 1) how gaze behavior develops over sessions in which student pilots learn to fly, 2) how their gaze bahavior relates to flight performance, and 3) how their gaze behavior is affected by additional cognitive load.

Method

Participants

Fifteen military cadets (12 males and 3 females) of the Royal Military Academy participated in this study. The participants had a mean age of 23.7 years (± standard deviation of 2.4 years), an average of 3.6±7.8 hours of flight experience on powered- and glider aircraft and 2.4±7.7 hours on flight simulators. Prior to the experiment, all participants signed an informed consent, stating that the details of the experiment had been sufficiently explained, and that they participated voluntarily. The experiment was conducted with approval of the institutional ethics committee and was in accordance with the revised Helsinki Declaration.

Materials

The simulator environment (see Figure 1), developed by the company multiSIM BV, consisted of a fixed-base cockpit (front-seat) of the Pilatus PC-7 turboprop trainer aircraft, includingcontrol devices with control loading. A VARJO-Aero VR device with built-in eye-tracker (200Hz) was used to present the cockpit and virtual environment near Woensdrecht Air Force Base, the Netherlands, rendering at 90Hz. The flight model characteristics were comparable to the PC-7 aircraft and were validated by EMVO flight instructors. During the experiment, audio instructions and an auditory secondary task were presented through a headphone.

Procedure

The participants repeatedly practiced three basic flight maneuvers: Straight-and-Level flight (SAL), Speed Change (SC) and Level Turn (LT). Each manoeuvre was performed three times in runs of 210 seconds each, followed by a test run in which the same manoeuvre was performed while simultaneously executing an additional N-back memory task as a measure of cognitive spare capacity. Each session of four consecutive runs was repeated three times, divided over two days, cumulating to twelve runs per manoeuvre.

The primary task consisted of manual control of the aircraft, including the instrument scan. This instrument scan was part of the lookout procedure, during which the particpant scanned the horizon and performed an instrument crosscheck each time when gaze passed the airplane nose. As secondary task, an auditory N-back memory task (Kirchner, 1958) was used, which required the participant to continuously update their working memory. The applied 2-back task required the participant to remember the last two letters of an auditory sequence of continuously changing letters at a fixed 3-seconds interval with a 25% repetition probability. The participant was instructed to make a self-paced response by pressing a dedicated button on the throttle if the letter heard was identical to the letter two trials back and to withhold a response if the letter was different.

Prior to each block, the participant was asked to study the instruction material (including a video from a flight instructor explaining each maneuvre), so that the instructions were identical for each participant. The participant received no feedback during the experiment.

Figure 1. (A) Setup of the VR simulator during the experiment, with the experimental test leader behind the instructor station and the participant inside the cockpit mock-up. (B) Pilot view with gaze direction overlaid in green. Note: the gaze direction was not shown during the experiment.

Measurements

From the recordings we analyzed the gaze direction in pitch, roll and yaw directions. We processed data to obtain the episodes where the gaze was directed inside the cockpit, and episodes where the gaze was directed to the outside environment. We calculated the outside dwell ratio, i.e., the percentage of time that the gaze was directed to the outside environment, during each run.

Objective measures of flight performance were also extracted, normalized and combined into one normalized performance measure on scale from zero to one per maneuver. This procedure is described in the accompanying paper by Ledegang et al. (in press).

Statistical analysis

We conducted a repeated-measures ANOVA on the outside dwell reatio with run and flight maneuver as within-subject variables to examine how the outside dwell ratio varied over runs across the three flight maneuvers. Note that runs 4, 8 and 12 were excluded from these analyses, as for these runs participants also perfomed the additional N-back task.

We then used Pearson correlations to investigate whether the outside dwell ratio related to flight performance, both averaged across flight maneuvers.

Finally, we examined whether the outside dwell ratio differed between runs with the N-back task (runs 4, 8, and 12) and the preceding runs without the N-back task (3, 7, and 11, respectively) using three separate two-tailed t-tests. For all analyses, alpha was set to .05.

Results

Figure 2(A) shows the group mean outside dwell ratio over the twelve runs for the SAL, SC and LT flight maneuvers, separately. A repeated measures ANOVA revealed that the outside dwell ratio varied between flight maneuvers, $F(2, 8) = 25.659$, $p < .001$ and varied over runs, $F(2, 8) = 2.069$, $p = .043$. Post-hoc paired t-tests revealed that the outside dwell ratio was significantly lower during the SC, $t(16) = -7.35$, $p < .001$, and LT $t(16) = -6.35$, $p < .001$, maneuvers compared to the SAL maneuver, but did not vary between SC and LT maneuvers, $t(16) = 0.99$, $p = .335$. A second set of post-hoc paired t-tests showed that for SAL, SC and LT maneuvers in none of the runs the outside dwell ratio significantly differed from the outside dwell ratio of run 1. This indicates that there was no systematic learning effect, even though the outside dwell ratio varied over runs.

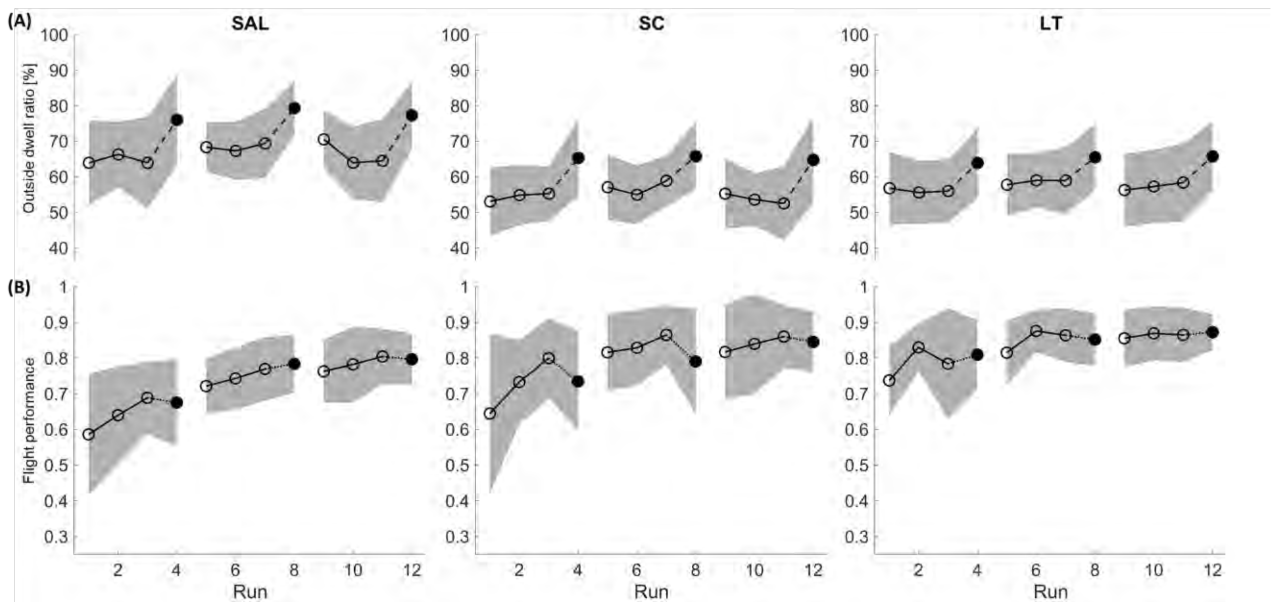

Figure 2. (A) Mean outside dwell ratio and (B) Flight performance as a function of the twelve runs for SAL, SC and LT flight maneuvers. Shading reflects standard deviation across participants. Filled markers correspond to runs with N-back task..

Figure 2(B) illustrates the group mean normalized flight performance over the twelve runs for SAL, SC and LT flight maneuvers, seperately. We found a significant positive correlation ($r = 0.18$, $p = .045$) between the normalized flight performance and the outside dwell ratio averaged across flight maneuvers, which means that participants with higher outside dwell ratios also showed better flight performance.

With respect to the effects of the additional N-back task, the black dots in Figure 2(A) illustrate that the outside dwell ratio was higher during runs with N-back task compared to runs without this task (open dots). Separate paired t-tests revealed that this was the case for each flight maneuver and each session, $p < .007$. To further investigate this effect, we examined differences in gaze direction between runs with and without N-back task in a post-hoc analysis. The results are shown in Figure 3(A), depicting a heatmap of the difference in gaze proportion during runs with versus runs without the N-back task. Figure 3(B) depicts the difference in gaze proportion in four areas-of-interest for the three flight maneuvers. It shows that with the additional N-back task a larger proportion of the gaze is being directed outside, and mostly straight ahead. This seems to indicate staring behavior.

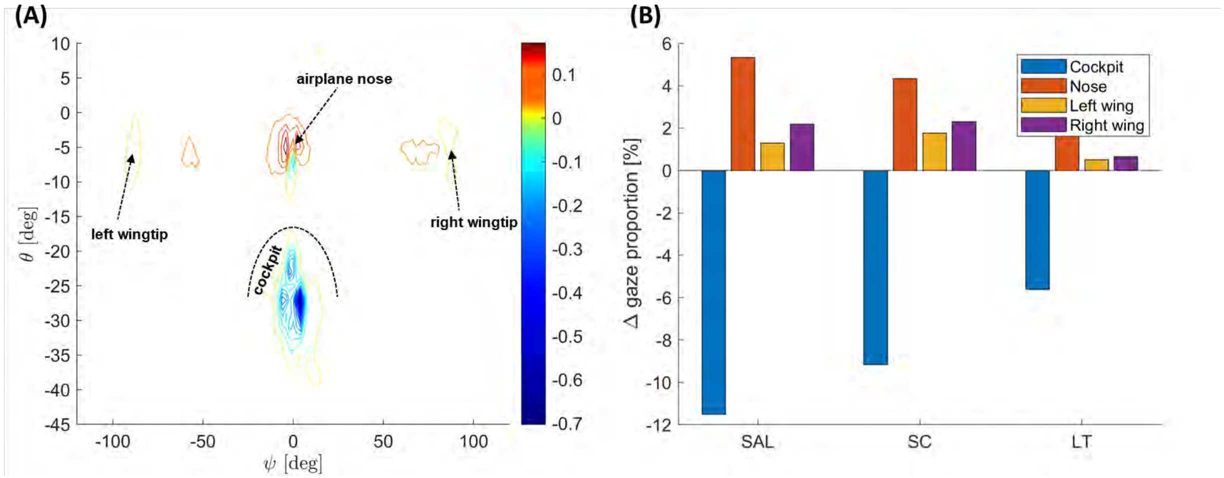

Figure 3. (A) Heatmap of difference in gaze proportion between runs with and without N-back task, averaged over all maneuvers. A positive number means a larger proportion of gaze is directed in that area during runs with N-back. (B) Difference in gaze proportion between runs with and without N-back for four main areas of interest: the cockpit, the outside view around the airplane nose, the outside view over the left wing and the outside view over the right wing.

Discussion

We examined how the Lookout behavior of student pilots changes during flight training in a VR flight simulator, how this Lookout behavior relates to flight performance and how extra cognitive load affects gaze behavior. Regarding the first research question, the outside dwell ratio varied significantly between flight maneuvers. In the SC and LT conditions, the outside dwell ratio was significantly lower than in the SAL condition. We explain this difference by the observation that compared to SAL, during the SC and LT maneuvers pilots have to monitor their instruments more closely to check the progression of their speed change or turn, respectively. Although the outside dwell ratio varied significantly over runs, we did not find evidence that the outside dwell ratio systematically increased over runs. Regarding the second research question, we found that the outside dwell ratio was positively correlated with flight performance. Thus, it seems that the participants with better flight performance were also directing a larger proportion of their gaze towards the outside environment.

Regarding the third research question we observed that with an additional N-back task the outside dwell ratio increased, indicating that the participants were staring. This suggests that with extra cognitive load the participants did not have the cognitive capacity to process the information of the flight instruments inside the cockpit.

Conclusions

Our findings indicate that the dwell ratio did not show any progression in Lookout performance in student pilots during a mini-flight training course without feedback from an instructor. This suggests that the dwell ratio is too rudimentory to measure progression. Therefore, future work should also consider measures capturing the scan pattern dynamics when assessing the lookout and instrument cross-check while learning to fly.

Acknowledgements

This research was supported by the Defence Research and Development Programmes V1917 and V1903, and RNLAF AIR. The authors acknowledge all RNLAF subject matter experts from EMVO, FCL and CMA for their guidance and advise in the preparation of the experiment and interpretation of the results. Furthermore, we thank multiSIM BV for their VR simulator developments and all participants for their enthusiastic contribution.

References

Air Education and Training Command. (2020). About PTN

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*(4), 352.

Ledegang, W. D., Van der Burg, E., Stuldreher, I. V., Houben, M.J., Groen, E. L. Van der Horst, D., Starmans, E. A. M, Almekinders, G. (in press). Acquiring manual flying skills in a virtual reality flight simulator. *International Symposium on Aviation Psychology*

Lewis, J., & Livingston, J. (2018). Pilot Training Next: Breaking institutional paradigms using studentcentered multimodal learning. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*, Orlando, FL.

Pennington, E., Hafer, R., Nistler, E., Seech, T., & Tossell, C. (2019, 26/04). Integration of advanced technology in initial flight training. *Systems and Information Engineering Design Symposium (SIEDS)*, University of Virginia, USA.

Pope, T. (2019). A cost-benefit analysis of Pilot Training Next Air University. Wright-Patterson Air Force Base, OH, USA.

Sheets, T. H., & Elmore, M. P. (2018). Abstract to action: Targeted learning system theory applied to adaptive flight training.

Wickens, C. D., McCarley, J. S., Alexander, A. L., Thomas, L. C., Ambinder, M., & Zheng, S. (2008). Attention-situation awareness (A-SA) model of pilot error. *Human Performance Modeling in Aviation*, 213-239.

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