Analysis and Control of Pilot Gaze Behavior to Support Basic Flight Training

D. Verkooij

Technische Universiteit Delft

TU Delft

Analysis and Control of Pilot Gaze Behavior to Support Basic Flight Training

by

D. Verkooij

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on March 29, 2022.

Student number:	4564251	
Project duration:	February 1, 2021 – March 29, 2022	
Thesis committee:	Prof. Dr. Ir. M. Mulder,	TU Delft
	Dr. Ir. M.M. van Paassen,	TU Delft
	Ir. O. Stroosma,	TU Delft
	Dr. Ir. M. Stroosma,	MultiSim B.V.
	Dr. O. A. Sharpans 'kykh,	TU Delft

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Acknowledgments

With this thesis, my long journey at the TU Delft comes to an end. I have thoroughly enjoyed this road towards becoming an engineer, but I am glad it has finally come to an end. It has taught me a lot about myself, my strengths, but also my weaknesses.

I would like to start by thanking my supervisors, Prof. Dr. Max Mulder, Ir. Olaf Stroosma and Dr. Rene van Paassen. They have provided me with valuable insights and guidance, and have managed to keep me motivated throughout this project. Special thanks to Dr. Mark Wentink from MultiSIM for providing me the opportunity for this project. Thanks to Dr. O. A. Sharpans 'kykh for freeing up time to be part of my committee.

In addition, I would like to thank everyone at MultiSIM for their assistance. Joris for the help with implementing the gaze tracking tools and VR. Martijn for teaching me how to work with MultiSIM's software. Stephen for the help with simulation hardware, and transporting the entire simulator setup to Delft. Maarten for the valuable tips on performing research. Ivan, for the help with the development and tuning of the PC-7 simulation.

Thank you to all my friends who were there when I needed them, and for listening to me talk about my thesis. My roommate Martijn for the support during our daily lunches. Sera for the occasional coffee break accompanied with some mutual support.

Lastly, I would like to thank my parents Danny and Jolijn for the crucial moral support. Thank you Laura for listening to me talk about this project and supporting me almost every day.

D. Verkooij
Delft, March 2022

Contents

Analysis and Control of Pilot Gaze Behavior to Support Basic Flight Training

Daan Verkooij*

It has been shown that expert gaze behavior can be used to enhance training for novices, among others in the field of laparoscopic surgery. The research in this paper focused on the effect of changing pilot gaze behavior on flight performance and detection task performance. Two groups of novices were shown a recording of expert behavior, including an indicator of the expert's gaze. The actively trained group was additionally shown a gaze director during training, to suppress the field of view outside of a set gaze point. Said gaze point moved around in a basic scanning sequence, encouraging this scanning pattern in the participants. Both groups were then evaluated on flight and detection performance in a segment of straight level flight, during which objects around the aircraft had to be detected. The active training group showed a larger change in gaze behavior between the before and after training tests, accompanied by a reduction in flight performance. Detection task performance was comparable for both groups. The gaze director thus seemed to distract participants from the tasks. One suggestion for future improvement is to focus on tuning the gaze director's timing using questionnaires. Alternatively, further research could focus more on the passive exposure of novices to expert gaze behavior, and compare this to a control group.

I. Introduction

Starting as early as the 1950s, eye tracking technology has been used to learn more about pilot behavior in the cockpit (Fitts, Jones, & Milton, 1950). Eye tracking relies on the analysis of the direction of foveal attention, the fovea being the part of the eye that provides the sharpest vision. The eye-mind hypothesis states that the location that is currently being fixated on provides an insight into what the mind is currently occupied with (Just & Carpenter, 1980, 1976).

Due to advances in technology, eye trackers are now compact and can be found integrated in consumer virtual reality (VR) headsets. VR is also increasingly prevalent in training, as shown by recent qualification of the first VR training simulator (Northcote, 2021). Additionally, VR is being considered for a more extensive role in some military flight schools, including the EMVO (Elementaire Militiaire Vliegers Opleiding, Elementary Military Flight Training), part of the Royal Netherlands Airforce. This provides an opportunity to investigate the possible applications of eye tracking within training programs.

The use of eye-tracking in augmented and virtual reality has already been shown to be valuable for training in both in aviation and in other fields (Vlasblom & Rooij, 2020; Kang & Landry, 2014; Jarodzka et al., 2012). Furthermore, it has been shown that inadequate gaze behavior in the cockpit has been the cause of several accidents as shown by NTSB (1994) (National Transportation Safety Board). The link between inadequate monitoring and scanning patterns is solidified by Zaal et al. (2021), further showing the potential benefit in investigating scanning behavior.

Beyond the tracking of gaze behaviors, research into controlling pilot gaze behavior has shown promising results (Dubois, Blättler, Camachon, & Hurter, 2015). Little investigation has been done into how this would affect flight performance and scanning performance. For this purpose this study presents a 'gaze director' system that aims to train gaze behavior. The gaze director attempted to encourage participants to follow a predefined scanning pattern during training using a spotlight. The pattern used was designed to increase the amount of time spent looking outside of the cockpit. To test the developed gaze director an experiment was set up. For this experiment, two metrics were taken into account; flight performance and detection task performance.

Section II shows the background, focusing on the differences between novices and experts and the application of eye tracking. Section III discusses the gaze director design in more detail. Section IV elaborates upon the experiment method. The results are presented in Section V and discussed in Section VI. Section VII shows the conclusion.

*MSc student, section Control and Simulation, Aerospace Engineering, TU Delft

II. Background

A. Novices versus Experts

Three theories are used here to help explain the difference between novice and expert behavior. First of which is the long-term working memory, which states that experts use retrieval cues, stored in their short-term memory to access processed information which is stored in their long-term memory (Ericsson & Kintsch, 1995).

Second is the information reduction hypothesis, proposing that experts neglect information not relevant to the current task, reducing the total amount of information that needs to be processed (Haider & Frensch, 1996). Lastly, the holistic model of image processing states that experts also use information from their parafoveal vision (Kundel, Nodine, Conant, & Weinstein, 2007). The information from the parafoveal vision is used for an initial global analysis of a situation after which relevant details are extracted more purposefully.

The difference between novices and experts are also further established using the zone of proximal development (Vygotsky, 1978). The zone of proximal development refers to what a trainee can do with guidance. Activities within this zone are optimal for learning. Anything beyond this zone, is not within reach of the learner without further training.

B. Areas of Interest

The movements of the eye can be described through saccades and fixations. A saccade is a rapid eye movement (Javal, 1878) during which no processing of visual information occurs (Fuchs, 1971). A fixation is the opposite of a saccade, where the eyes are focused on an object and detailed information is taken in.

Raw eye tracking data is often categorized into several Areas of Interest (AOI) to be able to extract useful information. Processing the eye tracking data into AOI data has the benefit of reducing data complexity through the removal of eye tremors and drifts (Ditchburn, 1980). An AOI can be any area in the field of view that is considered relevant to the experiment. One of the available algorithms for converting raw eye tracker data into a string of AOIs is I-AOI (Salvucci & Goldberg, 2000). I-AOI labels each point not situated within an AOI as a saccade, with other data points being considered fixations after a minimum duration is met. The minimum fixation duration for I-AOI is commonly set to 0.1 second (Vlasblom & Rooij, 2020), (Nyström & Holmqvist, 2010).

AOIs can then be used to generate several measures that provide an insight into the gaze behavior of participants. One of which is the percentage of fixation time spent in a specific AOI, which has been used in several studies (Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001), (Brams et al., 2018), (Dehais, Juaneda, & Peysakhovich, 2020). For the experiment in this paper, this measure will be used to investigate changes in gaze behavior. A low number of AOIs is considered, as this mitigates the effect of noisy eye tracker data.

C. Training Using Eye Tracking

Training using eye tracking and eye movements has already been applied in other fields, such as the medical field (Tien et al., 2014). One particular method is using Eye-Movement Modeling Examples (EMMEs). An EMME is used to show expert gaze behavior to novices to enhance training procedures. This method has been used successfully in virtual laparoscopic surgery (Wilson et al., 2011). Expert laparoscopic surgery gaze behavior is not task modulated, and consists mainly of focusing on the object being manipulated (Vine, Masters, McGrath, Bright, & Wilson, 2012). Describing expert gaze behavior in a cockpit is more complex, as attention has to be divided over the outside world and the instruments (Wickens, Goh, Helleberg, Horrey, & Talleur, 2003). These differences between the two fields warrant caution when using EMMEs in aviation applications, as expert pilot gaze behavior is also harder to replicate, in part due to the reasons mentioned in subsection II.A.

One method of applying EMMEs is the spotlight method. In the spotlight method, anything outside of the gaze area is obscured in some way (de Koning, Tabbers, Rikers, & Paas, 2010). In comparison with drawing a circle around the gaze area, the spotlight method was found to be more effective (Jarodzka et al., 2012).

It has been shown novice pilot gaze behavior can be changed. Applying aural or visual warnings after two seconds of head down time showed to reduce head down time (Dubois et al., 2015). A reduction in head down time for novices can be considered to be more expert-like behavior (Johnson, Wiegmann, & Wickens, n.d.).

As mentioned, little research has been done on the effect of attempts to change gaze behavior in aviation, while it is something that shows promise in other fields. For this reason, this study will investigate the effect of an EMME on flight performance. One group will be trained passively, meaning they are exposed to an expert's gaze behavior after some familiarization. The second group will be exposed in the same way, but will also be shown an EMME (referred to as the gaze director) during flight to train this group actively.

III. Gaze Director

The aim of the gaze director is to control participants' gaze behavior and to achieve more expert behavior in the process. More specifically, the aim is to increase the time spent looking outside, and increasing awareness of other objects in the airspace. As discussed, experts and novices process information differently, meaning that directly replicating the expert behavior is not desirable. For this reason the gaze director is focused on guiding towards a basic scan, rather than forcing an exact replication of expert behavior.

The procedure of this experiment is shown in Figure 1. After the briefing and familiarization, a detection task is performed to evaluate how well participants are scanning. This was followed by a training phase in which both groups were shown a recording of expert behavior. The active training group was additionally shown the gaze director described in this section. This was followed by another detection task. The gaze director was implemented by placing a simulated sphere over the pilot's head, with a clear area in the texture where the gaze director was pointing. The size of the clear area was set to be slightly larger than the primary flight display, to allow for completely unobstructed viewing of all frontal instruments when the director is pointing at it. The orb was assigned a transparency value of 0.97, this enabled the use of the peripheral vision of the pilots, while still strongly highlighting a point in the field of vision.

The gaze director is only shown to the active training group. The implementation of the gaze director is, however, also used for both groups to display the expert's gaze location in the simulator recording. During this recording a participant would be able to look around, but not control the simulation as they normally could.

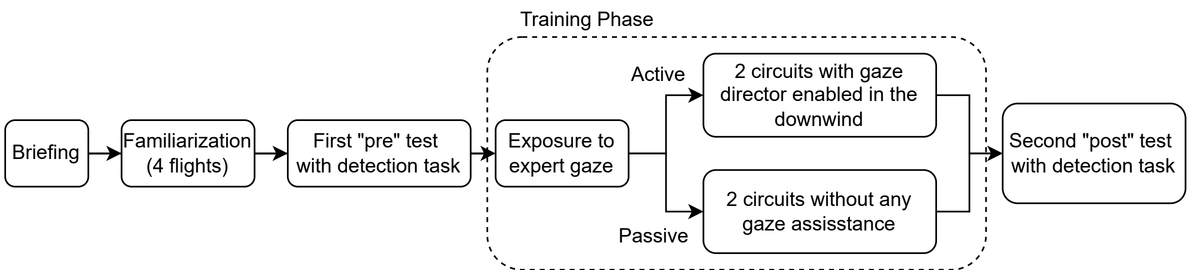


Fig. 1 Experiment procedure

The gaze director used for the active training group follows a predefined gaze pattern, which is indicated in Table 1. The pattern presented here consists of a full sweep of the horizon, with a check of the instruments when looking forward. This scanning behavior is also what is taught to students at the EMVO. EMVO students are, however, taught to check the ADI and no more than one other instrument before looking outside again. Due to the skill level of the participants, this part was left out of the gaze director's sequence.

In following this sequence, the gaze director moved at 50 deg/s, which is slower than the eye's maximum saccade speed of 300 deg/s (Fuchs, 1971). The reason for this slowdown is to encourage a scanning sweep over the horizon, discouraging jumping in gaze from one position from the next. This is meant to ensure objects between the two gaze points are also likely to be detected by participants. For example, if an object is positioned 45 degrees to the left from the current heading, a slow sweep would detect the object. In the scenario where a participant jumps in gaze from the forward position to the left tip, detection of this object would be unlikely, as little to no information is processed in a saccade (Fuchs, 1971).

Fig. 2 Gaze director and hot air balloon

Table 1 Gaze sequence during active training. Unlisted times are when the orb is moving at a constant speed.

Time[s]	Location
0	Forward
2-3	Over left wing
5-6	Forward
8-9	Over right wing
11-12	Forward
13-15	Instruments
16-17	Forward

IV. Method

A. Participants

Twenty students from Delft University of Technology without prior flight experience participated in this experiment. Two groups of ten participants were formed randomly and both groups experienced a different training method, as described in Section III.

B. Apparatus

The experiment was performed on a stationary research simulator provided by MultiSim. Figure 3 shows the full setup used during this experiment. Label **A** shows the VR headset used for this experiment, the HP G2 Omnicept edition. Its diagonal field of view is 114 degrees, and each eye has a screen with 2160 x 2160 pixels.

(a) Overview **(b) Close-up**

Fig. 3 Simulator setup during the experiment. A: VR headset, B: Cockpit instruments, C: Control stick, D: Participant's view, E: Instructor control panel, F: Throttle

B indicates the cockpit instruments. **C** shows the control stick, which was linked to a Brunner CLS-P cyclic control loading system. The force-displacement curve was linear, and scaled with speed to simulate increasing control forces. This scaling was tuned with PC-7 instructors for the flight conditions in this experiment. The hat switch on the control stick controls the simulated aircraft pitch and roll trim. **F** presents the throttle.

The screen with label **D** serves to show the experiment operator the perspective of the participant, including a small torus indicating the current gaze location of the participant. Label **E** was the interface for the observer of the experiment, and interfaces with a scenario system to control the experiment conditions.

This simulator simulated a Pilatus PC-7, with propwash effect removed from the flight model to reduce the difficulty of keeping the aircraft in a stable flight. The flight model was developed as an extension of MultiSIM's generic flight model. The instruments available to the participants can be seen in the setup with label B, the primary flight display can also be found in Figure 5.

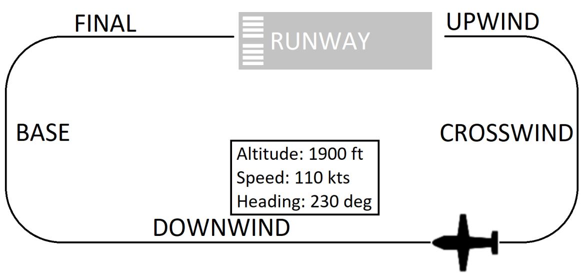


Fig. 4 The basic flight circuit flown in this experiment

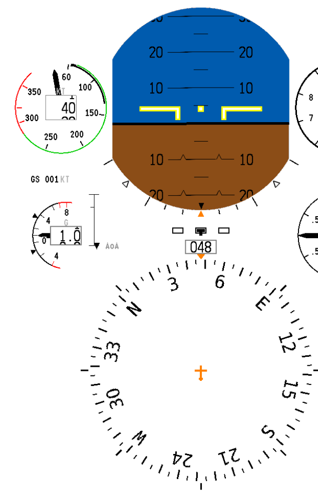


Fig. 5 Primary flight display for the PC-7 simulation

C. Experiment Procedure

After the briefing, participants were given the time to perform four flights in a familiarization phase. Four flights amounted to about 20 minutes of flying and was considered to provide adequate time to learn how to fly the aircraft with consistent performance.

After familiarization, the first test flight is done. Participants were instructed to look for intermittently appearing hot air balloons around the aircraft, somewhere in the forward hemisphere. Upon spotting such a hot air balloon, the participants had to press the trigger on the control stick while looking in the general direction of the balloon. Hot air balloons, 10 meters tall, were placed at a 2,500m distance in the simulation. In terms of FoV, the balloons covered about 4 degrees. Figure 2 shows an example of the hot air balloon. The task was only active for 120 seconds during the downwind leg of the circuit, and started as soon as the wings were level after the turn to the downwind leg. Two fixed sequences of 6 balloon presentations were used in the test; the second sequence was the mirror image, in time, of the first. During the first detection test, half of the participants in each experiment condition were presented with the original sequence, the other half with the reversed. For the second detection test, after the training, participants were tested with the other condition. This was done to ensure no bias, due to one of the sequences possibly being easier, could occur.

The first test was followed by an exposure phase. In this phase, a recording of an expert flying the basic circuit is shown. The expert used was an instructor at the EMVO. During the recording, participants could look around in this simulation as they normally would, while also being able to see the gaze location of the expert. As eye movements can be up to 300 deg/s, this indication would be difficult to follow unfiltered. Therefore, a second order low pass filter with $\zeta = 1, \omega_n = 15rad/s$ was implemented on the attitude of the gaze director.

After the exposure to expert gaze behavior, the active training group performed two more flights with the gaze director showing the gaze sequence on the downwind leg of the circuit. This started as soon as the wings were level after the turn to the downwind leg, just as the . The passive training group instead performed two more flights without gaze director.

Lastly, both groups performed one more test flight, with detection task and without gaze director.

D. Simulated Scenario

Each phase of the experiment consisted of one or more flights, such a flight is shown in Figure 4. Additionally, participants were given an altitude (1,900ft), speed (110kts) and a heading (230 deg, parallel to runway) to fly at for the downwind leg of the circuit. The deviation from these parameters served to gain insight into the participants' flight performance. The basic circuit was selected to allow for a realistic flight. However, as the participants are inexperienced, performing a stable landing can be considered to be outside of their capabilities, meaning the downwind leg was the focus for this experiment.

E. Dependent Measures

The full simulator time trace was logged for each participant, but measures are only taken from the downwind leg. The measures can be divided in the following three categories:

Flight performance was investigated through three parameters, the altitude, the heading and the speed. Together these three parameters set a 'correct' track to be flown over time. The way this track is defined means the distance between the aircraft and the runway is not a factor, allowing participants some room for error. The absolute error between the real value and the expected value was relevant for each of these.

Detection task performance: The detection task provided two objective measures for investigating performance. First, the time needed to spot objects in the vicinity of the flown aircraft. Second, the number of objects that were successfully detected.

Gaze behavior: Several measures were used to quantify the participants' gaze behavior. First of which is the dwell time percentages per AOI. The AOI data are calculated using I-AOI, using a minimum fixation duration of 0.1 second. AOIs investigated are the left, right and forward directions outside, and the instruments.

Additionally, the common N-gram patterns and the number of transitions are looked at to classify changes in behavior.

F. Hypothesis

The use of the gaze director for the active training group was expected to lead to larger changes in gaze behavior, with a shift towards looking outside. More time spent looking outside, would then mean better performance on the detection task for the active training group. A shift towards looking outside would also be accompanied by less time spent looking at instruments, leading to less accurate instrument flying.

H1: The active training group will show larger change in scanning behavior.
Metrics: Percentage of fixation time, number of fixations, common patterns

H2: The active training group will show a larger increase in detection task performance.
Metrics: Detection time, Missed detections

H3: The active training group will show a decrease in flight performance compared to the passive training group.
Metrics: Heading RMSE, Altitude RMSE, Speed RMSE

H4: The active training group will show more expert like gaze behavior than the passive training group.
Metrics: Percentage of fixation time, common patterns

G. Data Analysis

A one-sample Kolmogorov-Smirnov test or K-S test has been used to test the normality of the samples. After confirmation that the samples are not normally distributed, a MannWhitney U test for significance between the two groups after training was performed. For the paired comparison within both groups (before and after training), a Wilcoxon signed-rank test is performed, with $\alpha = 0.05$.

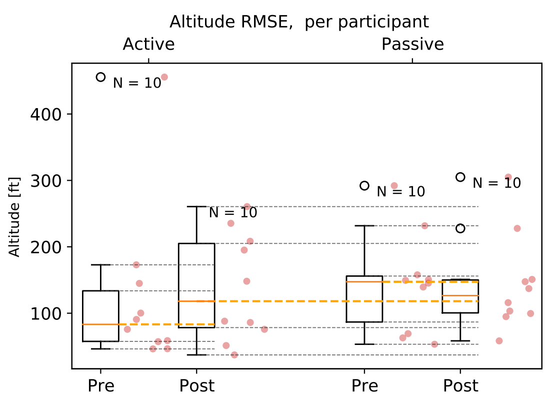
For the percentage of fixation time spent looking at instruments, multiple tests are performed on a set of variables that are not fully independent. Since three related percentages are tested, a Bonferroni correction is applied. The Bonferroni correction leads to a new α of 0.01667.

V. Results

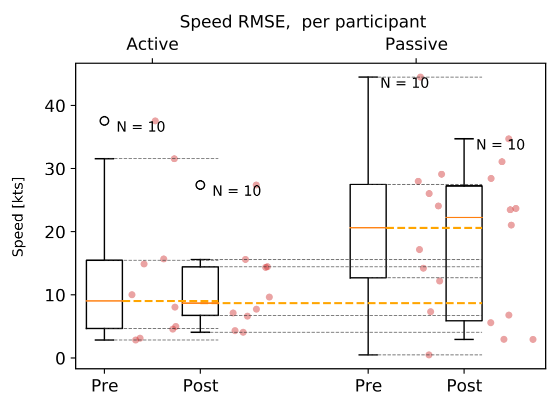
A. Flight Performance

Figure 6 presents the flight performance results for the active and passive training group, both before and after the training phase. Figure 6a shows the altitude root mean square error (RMSE). In the upper half the active training group is presented. Both the median (83.2ft → 118ft) and spread of the altitude RMSE increased for this group (Z = 19, p = 0.431). For the passive training group a decrease in both median (147ft → 126ft) and spread is visible (Z = 25, p = 0.8457). The difference between the two groups after training is not significant (U = 38, p = 0.182). For the speed RMSE, presented in Figure 6b, the active training group median remains almost the same (9.04kts → 8.69kts), while the spread has been reduced (Z = 25, p = 0.846). The median for the passive training group experienced a similarly small change (20.6kts → 22.3kts), accompanied with a reduction in the spread (Z = 27, p = 0.352). The difference between the two groups after training is not significant (U = 44, p = 0.325). Figure 6c shows the heading RMSE. Here an increase in spread is visible for both groups. The active training group (3.61deg → 4.17deg) also shows a slight increase in the median error (Z = 24, p = 0.770), the passive training group again shows a reduction (4.63deg → 3.90deg) (Z = 17, p = 0.322). The difference between the two groups after training is not significant (U = 47, p = 0.410).

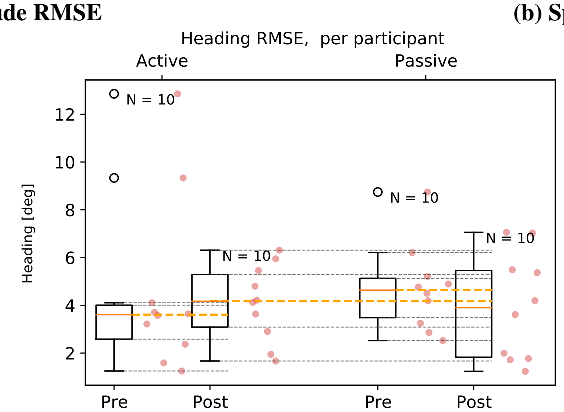
These results seem to indicate an overall reduction in flight performance for the active training group. The passive training group showed a small improvement.



(a) Altitude RMSE



(b) Speed RMSE

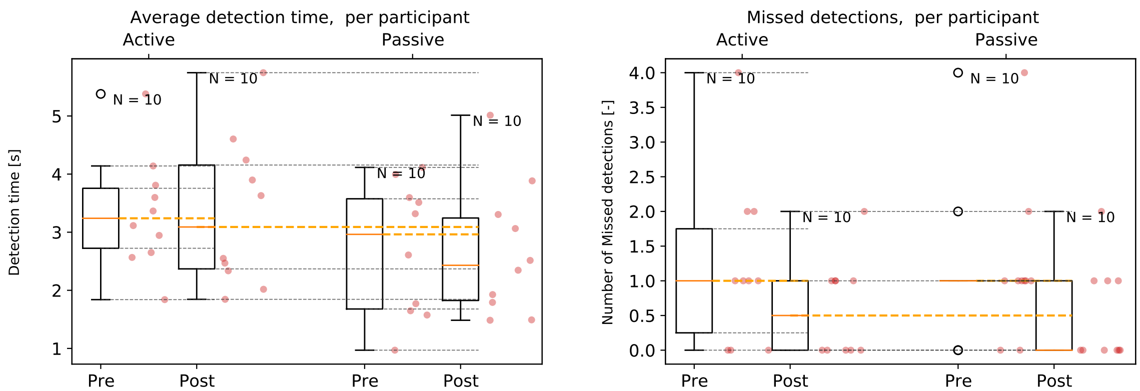


(c) Heading RMSE

Fig. 6 Flight Performance Results

B. Detection Task

Figure 7 presents the results of the detection task. Figure 7a presents the average time between the appearance of the balloon and the detection of said balloon per participant. The median (3.24s → 3.09s) sees a minor decrease for the active training group before and after the training session (Z = 25, p = 0.846). The passive training group sees a larger decrease in median (2.96s → 2.42s) (Z = 26, p = 0.922). The statistical difference between the groups for the average detection time is not significant (U = 33, p = 0.0993). Figure 7b presents the number out of six hot air balloons that have not been detected by a participant. Both the active training group (1.0 → 0.5) and passive training group (1.0 → 0.0) show a decrease in the median, with a tighter spread around this new median. The improvement in the passive training group seems to be larger, with a slightly lower median. The median seems to be a function of the number of participants that had no missed detections. The active training group had 5, and the passive training group had 6 participants without missed detections in the second test. This could indicate that a larger number of balloons would be needed to effectively use this measure for comparison. The reduction in the missed detections is not significant for the active training group (Z = 12, p = 0.190), the passive training group (Z = 3.5, p = 0.0578), nor for the comparison between the groups (U = 45.5, p = 0.351).



(a) Time between hot air balloon appearing and detection **(b) Number of hot air balloons not detected**

Fig. 7 Detection task results

C. Gaze Behavior

Figure 8 presents the results for the percentage of time spent fixating in that AOI. Figure 8a shows the percentage of time spent looking left, which shows a increase in median (9.73% → 15.4%) for the active training group. A smaller increase (11.6% → 12.8%) can also be seen for the passive training group.

The active training group shows a significant difference for the left AOI (Z = 4, p = 0.0137), considering the p value of 0.0167 due to the Bonferroni correction. The difference is not significant for the passive training group (Z = 10, p = 0.0840), nor for the comparison between groups (U = 32, p = 0.0868).

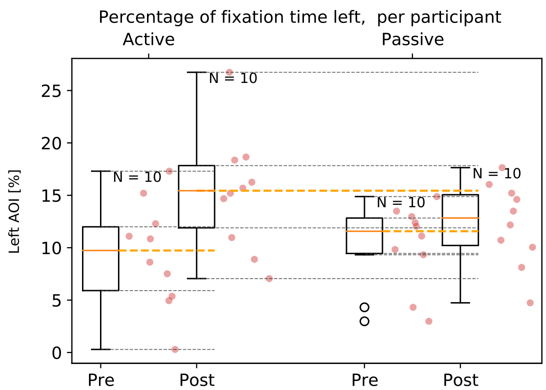
Figure 8b presents the same information for looking forward. The active training group shows a decrease in median (43.8% → 34.1%), with the maximum value also decreasing. A similar decrease (47.6% → 39.7%) is found for the passive training group , with the median staying slightly higher than for the active training group. The maximum shows a larger decrease.

The change for the passive training group is significant (Z = 2, p = 0.0058), while the change for the active training group is not significant (Z = 9, p = 0.0644). The difference between groups is not significant (U = 36, p = 0.145).

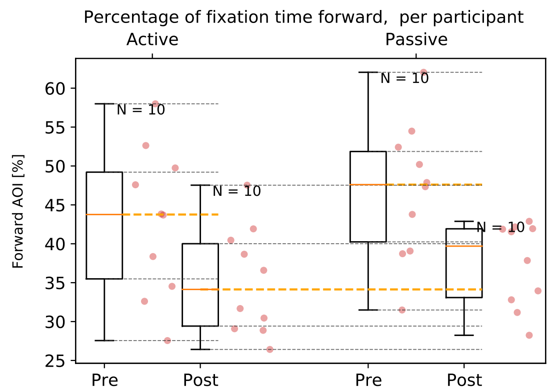
Figure 8c shows the time spent looking at the instruments. The active training group shows a reduction in maximum, with only a minor change in median (31.8% → 32.8%). The passive training group shows a small reduction in the spread around the median, and a similarly small change in the median (32.5% → 33.1%).

For both the active training group (Z = 21, p = 0.557) and the passive training group (Z = 17, p = 0.322) these small changes are not significant. The same can be said for the difference between the groups (U = 42, p = 0.273)

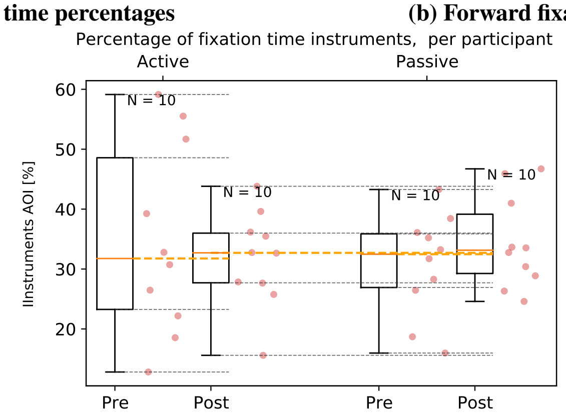
Overall, the active training group can be seen to be looking around more, with a decrease in time spent looking forward, while the median for the time spent looking inside barely changes. This means that more time is spent looking to both the left and right when looking outside.



(a) Left fixation time percentages



(b) Forward fixation time percentages

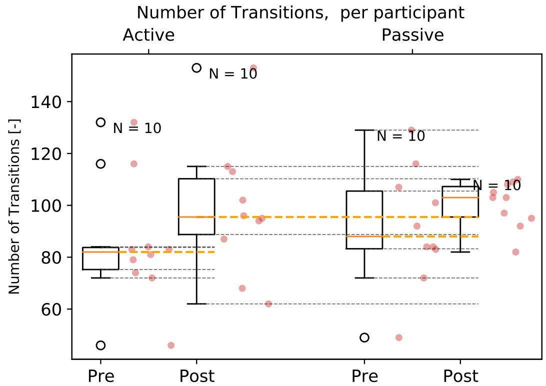


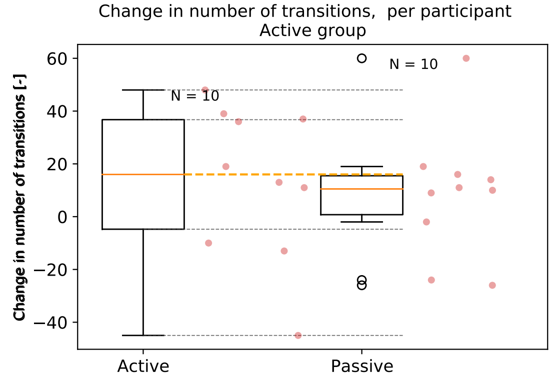
(c) Instrument fixation time percentages

Fig. 8 Gaze behavior fixation time percentages

Figure 9 presents boxplots with information on the number of transitions between AOIs per participant. Figure 9a shows the number. The active training group shows an increase in median (82.0 → 95.5) for this parameter, and a wider spread around this median. The passive training group has a slight increase in median (88.0 → 103) with a tighter spread around this median. Additionally, Figure 9b shows the difference between the pre-training and the post-training scenarios, here it can be seen that the active training group has a much larger median than the passive training group.

For both the active training group (Z = 13.5, p = 0.160) and the passive training group (Z = 18, p = 0.357) the difference in the number of transitions is not significant. The difference between the groups is not significant either (U = 42.5, p = 0.285).



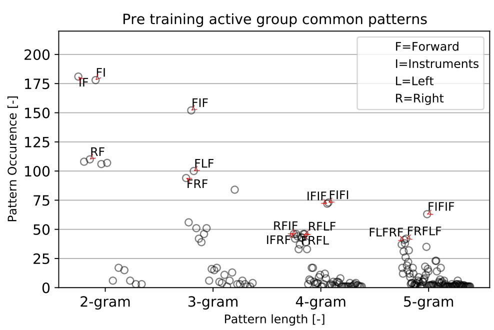


(a) Number of transitions

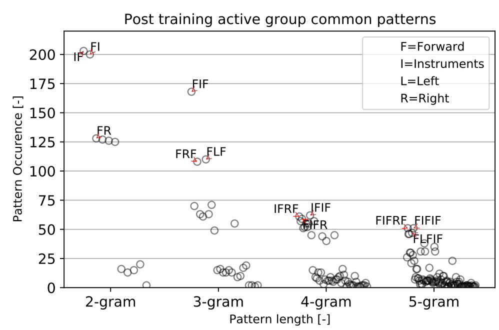
(b) Change in number of transitions

Fig. 9 Number of transitions between AOIs per participant

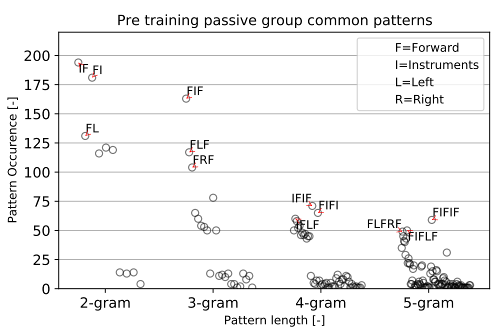
Figure 10 shows the distribution of the N-gram patterns found in the AOI data. Indicated in each plot are the three most common patterns, and any tiebreakers. Most notable for the active training group are the changes for the 4-grams and 5-grams. The patterns consisting of repeated switching between the instruments and looking forward occur less, and patterns relating to the sides show a minor increase. These changes are not visible for the passive training group. For the passive training group it should also be noted that the repeated switching patterns are already less pronounced, so any changes would be too.



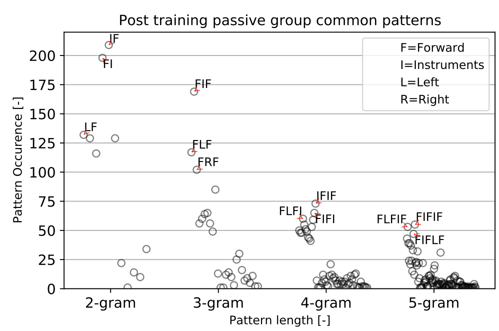
(a) For the active training group before training



(b) For the active training group after training



(c) For the passive training group before training



(d) For the passive training group after training

Fig. 10 Commonly occurring gaze patterns (N-Gram patterns)

D. Further Observations

Additional observations were also made during the experiment. Some participants were more likely to successfully perform a stable turn from the crosswind to downwind after having viewed the recording of expert behavior. Their gaze behavior also changed to include a fixation on the runway. A more stable turn lead to a smaller heading error at the start of the downwind, as presented in Figure 11. It is noteworthy that this effect was only observed in the passive group.

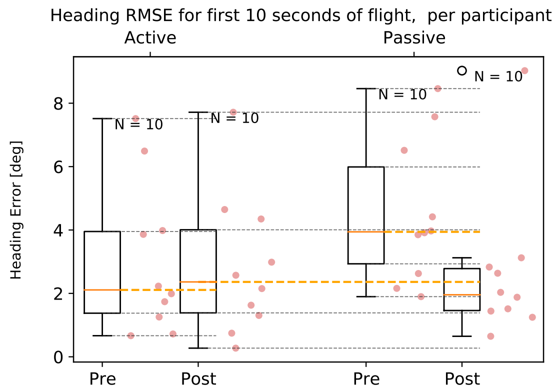


Fig. 11 Heading RMSE in first ten seconds of downwind

During the active training phase it also became clear that the gaze director went too slow for some participants. An example of when this became clear is when the orb was in the 'forward' position, and it started moving left, some participants would perform a sweep that went much faster than the gaze director. All the director seemed to do in these cases is serve as an interruption.

The mismatch between participant scanning speed and gaze director could have two possible causes. First is that the gaze director is moving too slow. Second, the participants might not be performing a sweep, but are moving their gaze from one point to the next without looking across the horizon.

The former would indicate that more tuning would be needed on the sequence shown by the gaze director. The cause would be the decision to set gaze director's speed to 50 deg/s, which was done in order to encourage a scanning sweep.

The latter would indicate that the gaze director is not forcing a slow scan enough. Participants are simply following the sequence they have remembered. No scan across the horizon would mean a reduction in detection performance, which is consistent with the results showing a decreasing trend on detection task performance. A possible solution is increasing the opacity of the gaze director, forcing participants to follow it more precisely.

VI. Discussion

To investigate the effectiveness of training tools for gaze behavior and the effect of changing gaze behavior on performance, an experiment in flight training was performed with two experimental groups that differed in the training of gaze behavior. The active training group was trained with a newly developed tool, the gaze director, while the passive training group was trained without this tool. The gaze director attempts to encourage participants to follow a predefined scanning pattern during training using a spotlight.

Evident from the lack of statistical significance is that an experiment such as this one would require more participants. The need for larger groups also becomes evident when looking at the difference between the two groups before the training phase. The groups are not homogeneous. Looking, for example, at the Altitude RMSE and Speed RMSE, differences are already visible before training.

In addition to requiring more participants, this experiment would have benefited from a trial experiment on just the active group. This would likely have shown the same downwards trend on detection task performance that was found in this experiment. Using this information, the experiment could have been altered to find more promising results.

After gaze training, the time spent by the active training group looking at the instruments shows barely any changes, while the time spent looking forward shows a decrease. This means the time spent looking to the left and right of the aircraft has increased. While this is in line with the desired changes towards more expert behavior, the passive training group shows similar results, albeit with less extensive changes between the pre and post training tests.

The change in the active training group's scanning behavior is also visible when looking at the N-gram patterns. The patterns with the left and right AOIs included become more pronounced. However, while both metrics seem to indicate changes towards more expert behavior, the differences between the two groups are not significant and hypothesis **H4** is rejected.

When looking at the number of transitions between AOIs, a larger change is visible in the active training group compared to the passive training group. Combining the information of the percentages of fixation time and the number of transitions, the active training group can indeed be said to have a larger change in scanning behavior than the passive training group. However, as none of these trends or changes are significant, it cannot be said that hypothesis **H1** is accepted.

The active training group shows a negative trend for flight performance. The altitude and heading RMSE both show an overall increase between the two tests, and no improvement is visible in the speed RMSE. The passive training group shows a general positive trend for flight performance instead. This would suggest that the active training using the gaze director is detrimental to the flight performance, which is in line with the expectations set by **H3**. Again, these differences are not significant, meaning hypothesis **H3** is not accepted.

The active training group's reduction in flight performance can be explained by the changes in gaze behavior. More time is spent looking outside, making instrument flying more difficult. Flying with the more 'expert'-like gaze behavior is therefore outside the zone of proximal development. Further research might focus on finding the location of the expert gaze behavior relative to the zone of proximal development. Such an investigation might benefit from a questionnaire investigating how the participants experienced the gaze director. One could for example ask whether participants feel the gaze director is interrupting fixations. The answer could provide information on the tuning of the gaze director. For example, if the participants feel the gaze director is interrupting fixations specifically on the instruments, the gaze director could remain on this area longer.

In the detection task, both groups seem to have shown a similar level of improvement. It is therefore likely that the gaze director, while seemingly changing the gaze behavior, did not lead to better detection task performance. Hypothesis **H2** is therefore rejected. One possible cause for the lack of improvement is that participants end up looking, but not seeing, since they are spending too much effort on exactly following the prescribed gaze pattern. In doing so, they do not actually end up looking for balloons more actively. In addition, this focus on the gaze pattern leads to neglecting the actual flying part of the circuit.

An option for further research would be to investigate the increase in performance seen in the passive training group. From this experiment, it is not clear whether this increase in performance is due to the learning curve, or due to the exposure to expert gaze behavior. Such research could try to record the learning curve over several sessions for a group exposed to gaze behavior compared to a control group.

Further research with an increased number of participants would be necessary to investigate if the differences found in this study are actual differences between the two sampled populations. However, since the results do not show promise of enhancing training, any further research would likely have to take a different approach. As it was noted that participants sometimes ignored and skipped over the gaze director during the training phase, a solution might be a more opaque gaze director, forcing a sweep over the horizon.

Moreover, additional work might focus on the apparent benefit noted in the further observations subsection of the results. Participants were found to copy certain behavior from the expert's flight recording. Specifically, the fixation on the runway during the turn onto the downwind. This seemed to lead to a more stable turn, which was accompanied by a more accurate endpoint for this turn. As a result of this, the starting heading on the downwind is closer to the target value, leading to a lower RMSE. Therefore use of an expert recording in training might be further investigated. Dutch military first time flight training already uses a 360 degree full flight recording in their procedure training. It might be interesting to investigate the benefits of displaying the gaze location of the expert in this recording for further educational value, as was done for both groups in this experiment. However, due to the lack of a control group, the effect of this recording was not properly quantified. The shown effects might not differ from a normal learning curve.

When looking at the trends of the data, one can say that the change in the active training group's scanning behavior was indeed larger. However, due to the lack of significance, and the seemingly detrimental effect of the gaze director on detection task performance, changes to the system would be required before any further experiments.

VII. Conclusion

The purpose of this research was to investigate the possibility and effect of changing gaze behavior on flight performance and detection task performance using a gaze director. The gaze director attempted to encourage participants to follow a predefined scanning pattern during training using a spotlight. While the gaze director seemed to successfully change gaze behavior, this change was also accompanied by a reduction in flight performance. Detection task performance on the other hand seemed largely unaffected when compared to the group that did not get the gaze director. Based on these results, the gaze director seems to reduce the performance of participants. A likely cause for this is that the participants' focus shifts towards following the scan pattern of the gaze director, causing participants to look, but not see. Additionally, participant were found to immediately look over the wing instead of performing a sweep across the horizon. One possible solution is to encourage scanning is to increase the opacity of the gaze director, to ensure participants follow the movement across the horizon more precisely. Another suggestion for further research is to focus on tuning the timing of the gaze director using questionnaires. Lastly, additional work could also focus on the passive exposure to gaze behavior, as the passively trained group showed some improvement.

References

Brams, S., Hooge, I. T. C., Ziv, G., Dauwe, S., Evens, K., De Wolf, T., . . . Helsen, W. F. (2018). Does effective gaze behavior lead to enhanced performance in a complex error-detection cockpit task? [Journal Article]. *PLOS ONE*, *13*(11), e0207439. Retrieved from `https://doi.org/10.1371/journal.pone.0207439` doi: 10.1371/journal.pone.0207439

Dehais, F., Juaneda, S., & Peysakhovich, V. (2020). Monitoring eye movements in real flight conditions for flight training purpose [Conference Paper]. In V. Peysakhovich, D. Rudi, M. Causse, P. Kiefer, F. Dehais, & M. Raubal (Eds.), *1st international workshop on eye-tracking in aviation (etavi 2020)* (p. 60). ISAE-SUPAERO, Université de Toulouse; Institute of Cartography and Geoinformation (IKG), ETH Zurich. Retrieved from `http://hdl.handle.net/20.500.11850/407652` doi: 10.3929/ethz-b-000407652

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2010). Attention guidance in learning from a complex animation: Seeing is understanding? [Journal Article]. *Learning and Instruction*, *20*(2), 111-122. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0959475209000103` doi: https://doi.org/10.1016/j.learninstruc.2009.02.010

Ditchburn, R. W. (1980). The function of small saccades [Journal Article]. *Vision Research*, *20*(3), 271-272. doi: 10.1016/0042-6989(80)90112-1

Dubois, E., Blättler, C., Camachon, C., & Hurter, C. (2015). Eye movements data processing for ab initio military pilot training [Conference Proceedings]. In *Kes international conference on intelligent decision technologies* (Vol. 39). doi: 10.1007/978-3-319-19857-6_12

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory [Journal Article]. *Psychological Review*, *102*(2), 211-245. doi: 10.1037/0033-295X.102.2.211

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches [Journal Article]. *Aeronautical Engineering Review*, *9*, 24-29.

Fuchs, A. F. (1971). The saccade system [Journal Article]. *The control of eye movements*, 343-362. Retrieved from `https://ci.nii.ac.jp/naid/10008713140/en/`

Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition [Journal Article]. *Cognitive Psychology*, *30*(3), 304-337. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0010028596900092` doi: https://doi.org/10.1006/cogp.1996.0009

Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye-movement modelling examples [Journal Article]. *Instructional Science*, *40*(5), 813-827. Retrieved from `https://doi.org/10.1007/s11251-012-9218-5` doi: 10.1007/s11251-012-9218-5

Javal, E. (1878). Essai sur la physiologie de la lecture [Journal Article]. *Annales d'Ocilistique*, *80*, 97-117.

Johnson, N., Wiegmann, D., & Wickens, C. (n.d.). Effects of advanced cockpit displays on general aviation pilots' decisions to continue visual flight rules flight into instrument meteorological conditions [Conference Proceedings]. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, p. 30-34). Sage Publications Sage CA: Los Angeles, CA.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes [Journal Article]. *Cognitive Psychology*, *8*(4), 441-480. Retrieved from `https://www.sciencedirect.com/science/article/pii/0010028576900153` doi: https://doi.org/10.1016/0010-0285(76)90015-3

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension [Journal Article]. *Psychological Review*, *87*(4), 329-354. doi: 10.1037/0033-295X.87.4.329

Kang, Z., & Landry, S. J. (2014). Using scanpaths as a learning method for a conflict detection task of multiple target tracking [Journal Article]. *Human Factors*, *56*(6), 1150-1162. Retrieved from `https://doi.org/10.1177/0018720814523066` doi: 10.1177/0018720814523066

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during vfr flight [Conference Proceedings]. In *Proceedings of the 11th international symposium on aviation psychology* (Vol. 6, p. 325-335). Citeseer.

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: Gaze-tracking study [Journal Article]. *Radiology*, *242*(2), 396-402. Retrieved from `https://pubs.rsna.org/doi/abs/10.1148/radiol.2422051997` doi: 10.1148/radiol.2422051997

Northcote, J. (2021, 4). Easa approves the first virtual reality (vr) based flight simulation training device [Press Release]. *EASA*. Retrieved from `https://www.easa.europa.eu/newsroom-and-events/press-releases/easa-approves-first-virtual-reality-vr-based-flight-simulation`

NTSB. (1994). A review of flightcrew-involved major accidents of u.s. carriers, 1978 through 1990. document ntsb/sss-94-01. [Journal Article]. *National Transportation Saftey Board*.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data [Journal Article]. *Behavior Research Methods*, *42*(1), 188-204. Retrieved from `https://doi.org/10.3758/BRM.42.1.188` doi: 10.3758/BRM.42.1.188

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols [Conference Paper]. In *Proceedings of the 2000 symposium on eye tracking research and applications* (p. 7178). Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/355017.355028` doi: 10.1145/355017.355028

Tien, T., Pucher, P. H., Sodergren, M. H., Sriskandarajah, K., Yang, G.-Z., & Darzi, A. (2014). Eye tracking for skills assessment and training: a systematic review [Journal Article]. *Journal of Surgical Research*, *191*(1), 169-178. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0022480414004326` doi: https://doi.org/10.1016/j.jss.2014.04.032

Vine, S. J., Masters, R. S. W., McGrath, J. S., Bright, E., & Wilson, M. R. (2012). Cheating experience: Guiding novices to adopt the gaze strategies of experts expedites the learning of technical laparoscopic skills [Journal Article]. *Surgery*, *152*(1), 32-40. Retrieved from `https://www.sciencedirect.com/science/article/pii/S003960601200044X` doi: https://doi.org/10.1016/j.surg.2012.02.002

Vlasblom, J., & Rooij, J. v. (2020). Augmented eye: From theory to practice [Conference Paper]. In *AIAA Scitech 2020 forum.* Retrieved from `https://arc.AIAA.org/doi/abs/10.2514/6.2020-0169` doi: 10.2514/6.2020-0169

Vygotsky, L. (1978). Mind in society [Journal Article]. *Mind in society the development of higher psychological processes. Cambridge, MA: Harvard University Press*.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology [Journal Article]. *Human Factors*, *45*(3), 360-380. Retrieved from `<GotoISI>://WOS:000187038200002` doi: 10.1518/hfes.45.3.360.27250

Wilson, M., Vine, S., Brewer, J., Bright, E., Masters, R., & McGrath, J. (2011). Gaze training improves technical performances and resistance to distractions in virtual laparoscopic surgery [Conference Proceedings]. In *Sages annual meeting 2011.* Society of American Gastrointestinal and Endoscopic Surgeons (SAGES). Theă.

Zaal, P., Lombaerts, T., Mumaw, R., Billman, D., Torron, I., Jamal, S., . . . Feary, M. (2021). Eye-tracking analysis from a flight-director-use and pilot-monitoring study [Conference Paper]. In *AIAA Aviation 2021 forum.* American Institute of Aeronautics and Astronautics. Retrieved from `https://doi.org/10.2514/6.2021-2995` doi: doi:10.2514/6.2021-299510.2514/6.2021-2995

A. Tables of generated measures

Table 2 Detection times for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	0.65	0.85	3.61	1.52
	0.62	3.16	2.27	1.64
	8.66	9.08	9.07	1.96
	6.63	7.35	1.84	1.33
	0.69	2.62	6.89	1.06
	7.06	2.39	1.01	1.4
	2.88	3.96	4.72	3.42
	3.73	0.77	3.95	0.48
	1.89	5.2	0.93	2.69
	1.42	1.12	1.41	0.57
	1.7	8.32	6.97	1.09
	1.85	2.41	1.13	0.7
	0.76	2.15	7.27	0.81
	6.07	0.81	2.54	1.53
	2.45	4.4	2.33	0.97
	8.78	0.65	3.32	0.87
	1.34	1.39	1.44	0.55
	9.55	1.67	7.34	6.83
	6.91	3.1	4.46	5.39
	2.22	4.74	0.63	1.25
	3.48	0.73	6.1	3.28
	9.94	2.25	0.84	3.3
	1.06	0.86	1.03	0.6
	1.31	0.87	5.8	1.25
	4.03	3.89	8.9	3.58
	6.03	1.23	0.99	1.46
	0.48	2.51	1.86	4.84
	3.04	1.59	4.25	1.7
	3.27	1.46	0.87	1.0
	2.01	1.41	0.65	1.54
	5.14	1.57	1.22	5.57
	2.23	7.2	1.49	3.29
	0.66	1.11	0.59	1.99
	2.71	3.13	2.47	0.83
	7.71	1.05	2.32	5.21
	2.26	1.39	2.17	8.98
	0.84	8.85	0.85	7.47
	9.1	7.62	0.78	0.74

Continued on next page

Table 2 Detection times for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	2.16	1.35	1.27	6.85
	4.74	4.67	4.74	6.01
	1.15	0.93	0.93	2.4
	2.17	0.52	5.32	4.5
	1.51	4.93	0.92	2.9
	1.82	1.15	1.02	1.39
	0.97	2.61	2.47	1.18
	1.66	9.79	0.88	4.59
	5.57	2.58	0.7	6.43
	3.23	0.83	2.25	2.65
		1.25		1.22
		8.57		4.53
		6.82		3.2
		8.74		2.31
		1.6		1.86
		5.82		1.07
				0.52
Sum	166	177	137	146.0
Mean	3.46	3.28	2.85	2.66

Table 3 Heading rmse for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	3.22	4.22	4.89	1.77
	12.86	5.45	4.77	4.19
	2.37	1.95	3.25	1.72
	4.11	3.63	4.5	1.24
	3.64	4.8	4.19	7.06
	1.25	1.67	5.21	7.03
	3.57	4.12	8.74	5.37
	1.59	2.91	6.21	5.49
	3.71	5.95	2.86	3.61
	9.33	6.31	2.52	2.0
Sum	46.0	41.0	47.0	39.0
Mean	4.56	4.1	4.71	3.95

Table 4 Speed rmse for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	15.69	14.46	17.18	5.59
	14.9	14.36	14.22	23.69
	2.85	4.34	12.19	21.05
	10.02	6.63	0.49	2.96
	31.56	27.41	29.12	23.49
	3.14	7.13	26.05	34.72
	8.06	9.65	24.08	28.44
	4.58	4.08	44.51	6.8
	5.0	15.61	7.33	2.94
	37.56	7.74	27.99	31.1
Sum	133.0	111.0	203.0	181.0
Mean	13.34	11.14	20.32	18.08

Table 5 Altitude rmse for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	100.19	51.33	139.42	94.86
	172.88	195.11	62.86	58.36
	144.9	148.17	53.25	305.01
	58.72	88.0	69.22	147.5
	90.61	208.37	145.28	150.92
	57.18	75.75	157.81	227.73
	46.29	86.04	292.08	103.29
	46.48	37.21	149.44	115.93
	75.75	235.37	150.48	99.56
	455.65	260.55	231.73	137.14
Sum	1249.0	1386.0	1452.0	1440.0
Mean	124.86	138.59	145.16	144.03

Table 6 Dwell percentages outside for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	60.75	72.16	64.79	73.68
	87.19	84.41	66.77	69.6
	77.81	72.35	73.54	75.41
	69.27	64.52	68.28	59.02
	81.46	63.84	81.31	71.12
	48.32	60.38	61.57	67.25
	40.87	67.34	71.71	54.08
	67.21	74.24	84.03	66.46
	44.48	67.25	56.73	53.28
	73.53	56.19	63.91	66.35
Sum	651.0	683.0	693.0	656.0
Mean	65.09	68.27	69.26	65.62

Table 7 Average detection time for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	4.14	4.24	4.12	1.48
	2.94	3.63	3.6	1.49
	2.57	1.84	3.32	1.93
	5.38	2.34	3.99	3.3
	3.81	2.02	3.51	2.35
	3.36	2.55	1.77	2.52
	3.11	3.9	1.65	5.01
	3.6	2.47	2.61	3.06
	1.84	4.6	0.97	3.88
	2.65	5.74	1.58	1.79
Sum	33.0	33.0	27.0	27.0
Mean	3.34	3.33	2.71	2.68

Table 8 Missed detections for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	2.0	0.0	0.0	0.0
	0.0	0.0	1.0	0.0
	1.0	0.0	1.0	0.0
	0.0	1.0	1.0	2.0
	0.0	1.0	1.0	1.0
	2.0	1.0	1.0	0.0
	1.0	0.0	0.0	0.0
	1.0	0.0	1.0	0.0
	4.0	1.0	4.0	1.0
	1.0	2.0	2.0	1.0
Sum	12.0	6.0	12.0	5.0
Mean	1.2	0.6	1.2	0.5

Table 9 LZC normalized x1000 for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	108.43	86.38	106.64	93.21
	105.83	91.75	92.2	91.16
	108.81	88.23	104.29	90.02
	111.74	96.13	89.39	89.72
	101.26	111.13	102.05	93.72
	102.05	94.04	84.03	101.64
	141.69	100.63	85.48	99.97
	75.28	107.04	88.72	90.78
	90.04	113.77	116.31	89.46
	85.48	68.45	108.05	92.07
Sum	1031.0	958.0	977.0	932.0
Mean	103.06	95.75	97.72	93.18

Table 10 **Dwell percentages left for each of the experiment conditions**

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	7.51	8.89	12.4	17.64
	17.3	18.36	14.88	12.17
	11.1	15.2	12.98	15.2
	8.62	15.68	4.31	4.74
	15.2	18.65	9.83	10.04
	5.37	14.68	11.11	16.04
	4.95	16.25	12.03	10.72
	12.29	26.74	13.49	14.61
	0.29	10.97	2.98	8.12
	10.84	7.05	9.33	13.49
Sum	93.0	152.0	103.0	123.0
Mean	9.35	15.25	10.33	12.28

Table 11 **Dwell percentages forward for each of the experiment conditions**

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	32.61	41.92	31.49	33.95
	52.63	47.53	39.07	37.85
	57.99	38.66	47.34	41.87
	43.83	29.08	52.42	42.17
	43.69	26.42	62.04	41.54
	34.53	31.68	38.72	28.24
	27.56	30.45	43.77	31.17
	47.58	28.87	54.48	41.94
	38.36	36.59	50.2	32.8
	49.75	40.46	47.87	42.89
Sum	429.0	352.0	467.0	374.0
Mean	42.85	35.17	46.74	37.44

Table 12 Dwell percentages right for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	20.63	21.35	20.9	22.09
	17.26	18.51	12.82	19.58
	8.73	18.49	13.23	18.33
	16.82	19.76	11.55	12.11
	22.56	18.76	9.43	19.54
	8.42	14.03	11.74	22.97
	8.36	20.64	15.91	12.19
	7.33	18.63	16.05	9.91
	5.82	19.69	3.55	12.36
	12.94	8.67	6.71	9.97
Sum	129.0	179.0	122.0	159.0
Mean	12.89	17.85	12.19	15.9

Table 13 Dwell percentages instruments for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	39.25	27.84	35.21	26.32
	12.81	15.59	33.23	30.4
	22.19	27.65	26.46	24.59
	30.73	35.48	31.72	40.98
	18.54	36.16	18.69	28.88
	51.68	39.62	38.43	32.75
	59.13	32.66	28.29	45.92
	32.79	25.76	15.97	33.54
	55.52	32.75	43.27	46.72
	26.47	43.81	36.09	33.65
Sum	349.0	317.0	307.0	344.0
Mean	34.91	31.73	30.74	34.38

Table 14 Mean altitude error for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	14.15	38.66	115.78	85.44
	14.71	177.29	25.11	10.27
	141.71	139.43	12.73	129.17
	43.5	74.13	68.49	84.28
	59.83	161.54	123.93	69.23
	29.11	33.43	140.4	55.27
	19.86	14.55	209.3	46.37
	32.07	23.91	49.53	81.26
	45.4	191.75	78.07	66.5
	316.21	179.06	81.3	114.55
Sum	717.0	1034.0	905.0	742.0
Mean	71.66	103.37	90.47	74.23

Table 15 Mean heading error for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	1.99	3.19	4.3	0.94
	10.41	4.68	2.14	3.67
	1.77	0.24	2.54	0.56
	3.72	2.29	3.14	0.03
	2.78	4.21	2.0	3.14
	0.84	1.21	3.85	6.38
	2.32	1.93	8.29	4.83
	1.32	1.11	2.0	4.82
	1.9	5.28	0.0	2.41
	8.43	5.61	1.48	1.16
Sum	35.0	30.0	30.0	28.0
Mean	3.55	2.97	2.97	2.79

Table 16 Mean speed error for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	12.75	14.04	13.77	3.81
	14.08	12.32	8.83	20.07
	1.86	3.52	11.01	18.49
	8.68	5.1	0.12	0.22
	31.19	23.55	28.52	19.57
	1.1	3.31	25.86	32.83
	6.62	7.59	16.72	23.75
	0.94	1.36	42.73	3.49
	2.67	5.56	3.92	1.06
	34.82	3.04	27.4	31.03
Sum	115.0	79.0	179.0	154.0
Mean	11.47	7.94	17.89	15.43

Table 17 Number of Transitions for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	79.0	115.0	92.0	108.0
	83.0	96.0	101.0	110.0
	74.0	113.0	84.0	103.0
	84.0	95.0	84.0	95.0
	81.0	68.0	83.0	97.0
	83.0	102.0	129.0	103.0
	46.0	94.0	116.0	92.0
	132.0	87.0	107.0	105.0
	72.0	62.0	49.0	109.0
	116.0	153.0	72.0	82.0
Sum	850.0	985.0	917.0	1004.0
Mean	85.0	98.5	91.7	100.4

Table 18 Heading RMSE for first 10 seconds of flight for each of the experiment conditions

Group:	Active Pre	Active Post	Passive Pre	Passive Post
	2.23	0.27	7.57	0.65
	3.86	0.75	3.91	1.44
	3.98	2.57	2.16	1.88
	0.72	4.35	6.51	3.12
	0.67	1.63	3.85	9.03
	1.74	2.99	1.9	1.25
	7.51	4.65	4.42	2.64
	1.25	1.3	8.46	2.03
	1.99	2.15	2.63	1.51
	6.49	7.72	3.97	2.83
Sum	30.0	28.0	45.0	26.0
Mean	3.04	2.84	4.54	2.64

B. Boxplots of generated measures

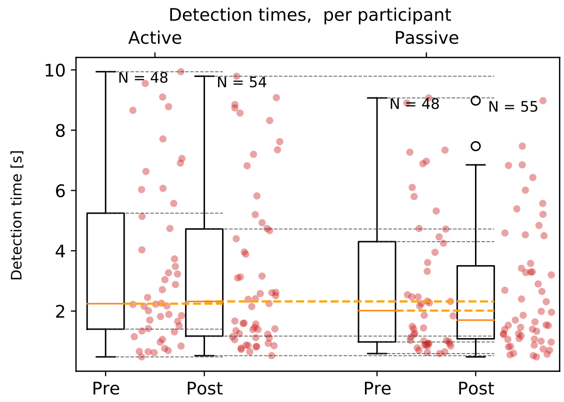


Fig. 12 Boxplot for Detection times

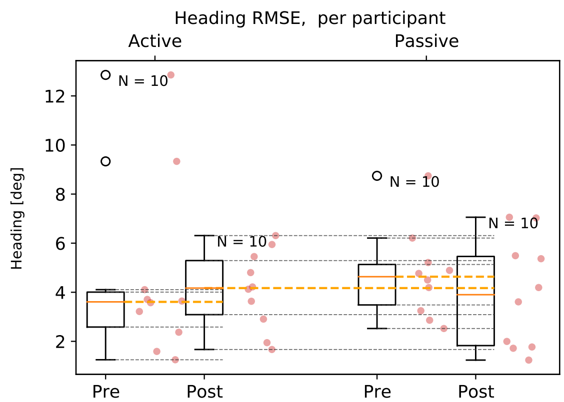


Fig. 13 Boxplot for Heading rmse

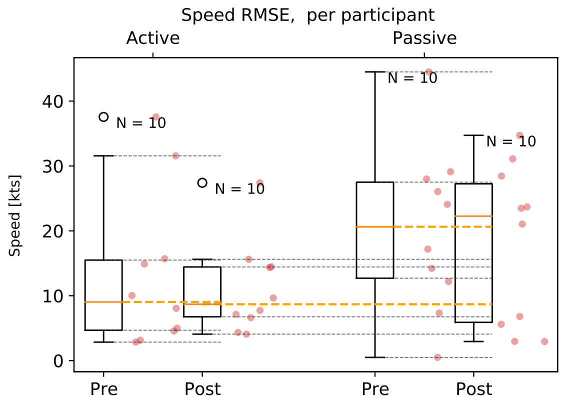


Fig. 14 Boxplot for Speed rmse

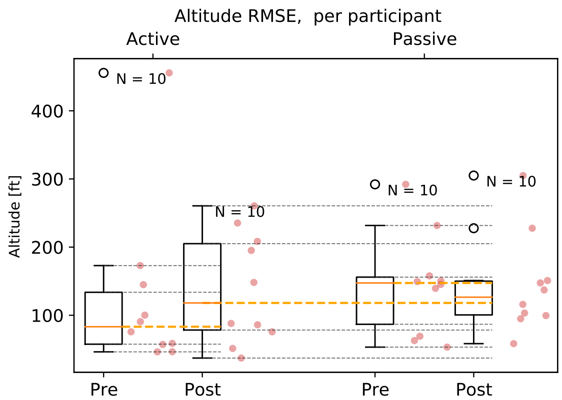


Fig. 15 Boxplot for Altitude rmse

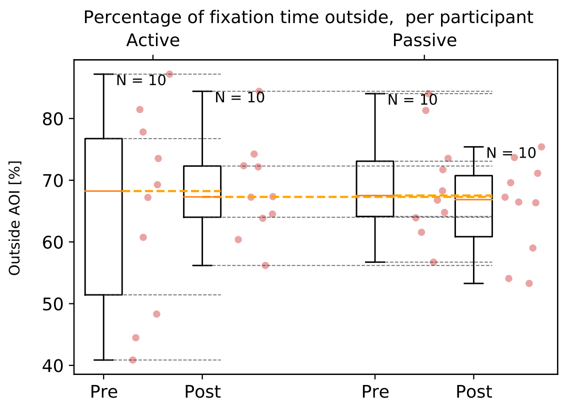


Fig. 16 Boxplot for Dwell percentages outside

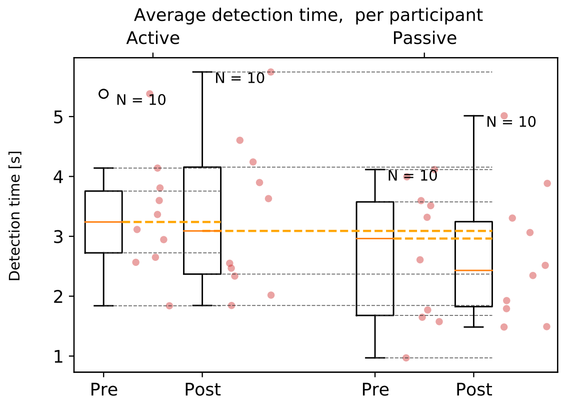


Fig. 17 Boxplot for Average detection time

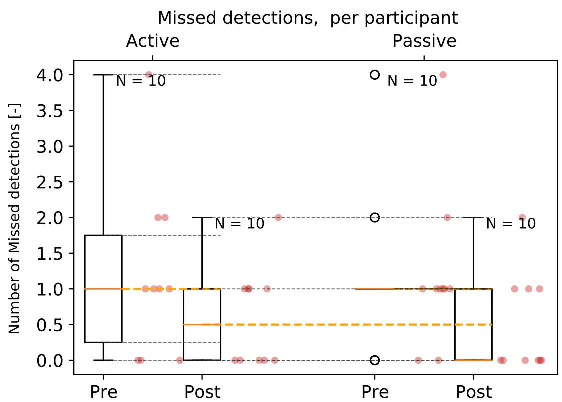


Fig. 18 Boxplot for Missed detections

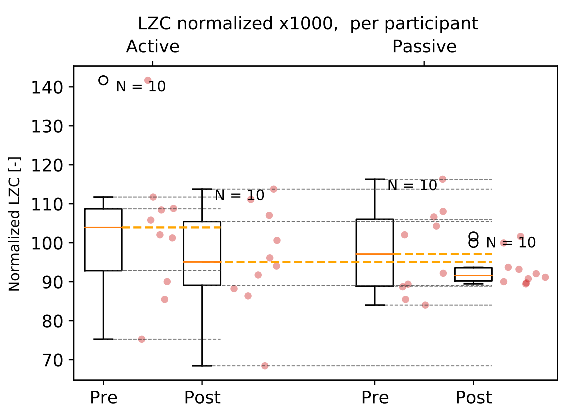


Fig. 19 Boxplot for LZC normalized x1000

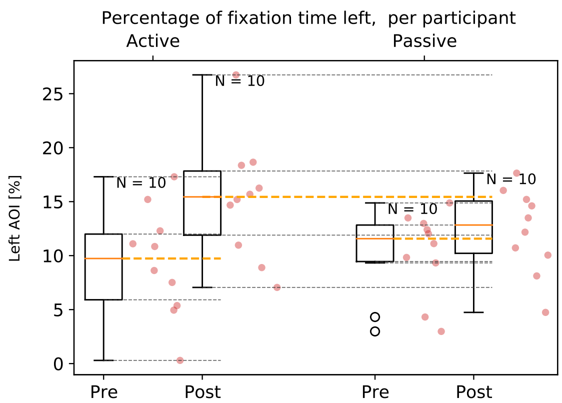


Fig. 20 Boxplot for Dwell percentages left

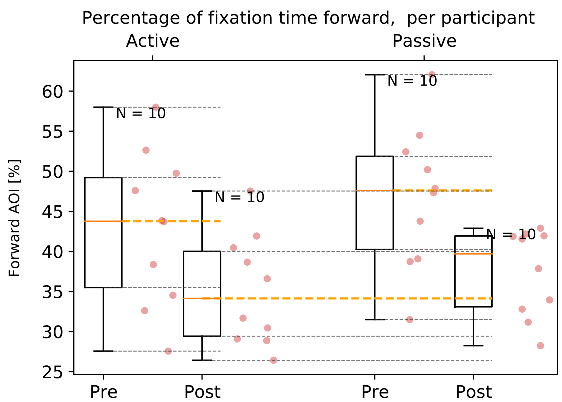


Fig. 21 Boxplot for Dwell percentages forward

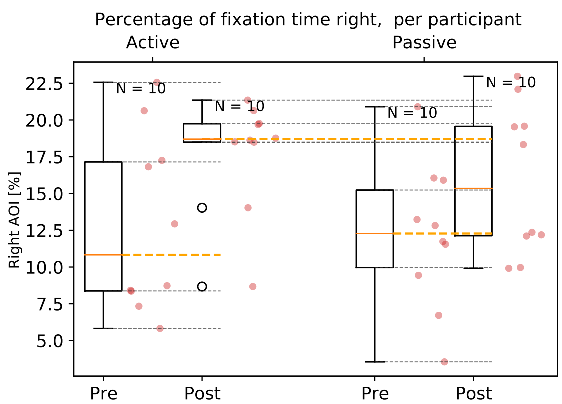


Fig. 22 Boxplot for Dwell percentages right

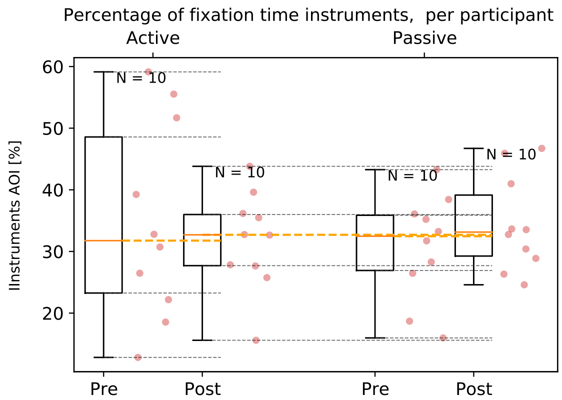


Fig. 23 Boxplot for Dwell percentages instruments

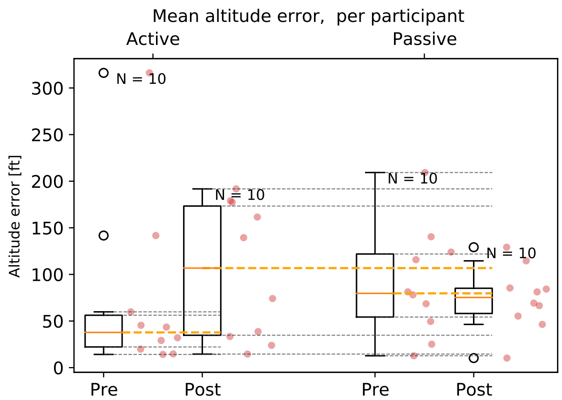


Fig. 24 Boxplot for Mean altitude error

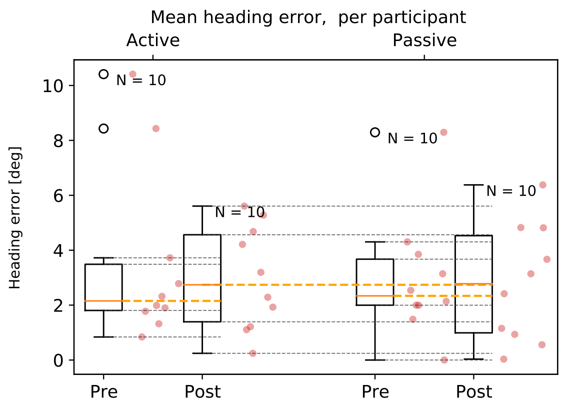


Fig. 25 Boxplot for Mean heading error

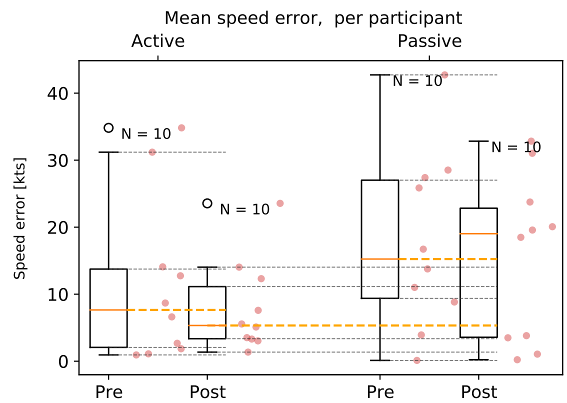


Fig. 26 Boxplot for Mean speed error

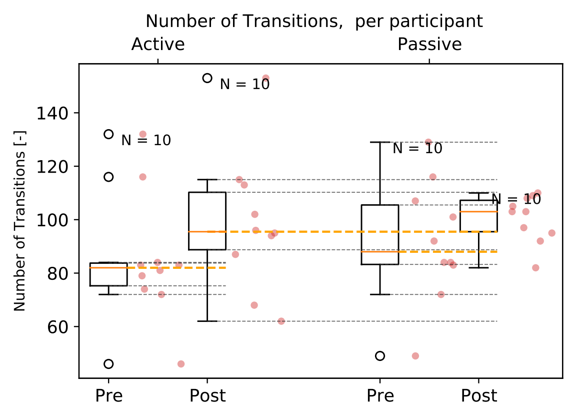


Fig. 27 Boxplot for Number of Transitions

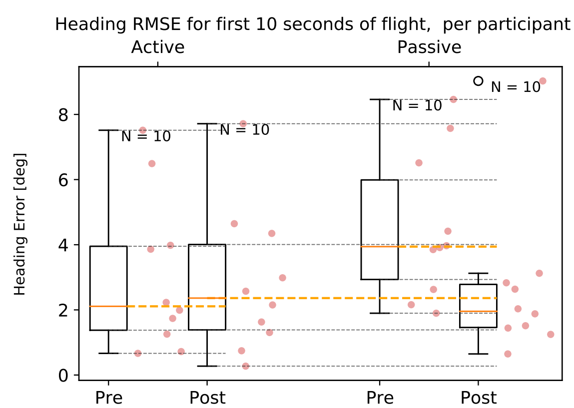


Fig. 28 Boxplot for Heading RMSE for first 10 seconds of flight

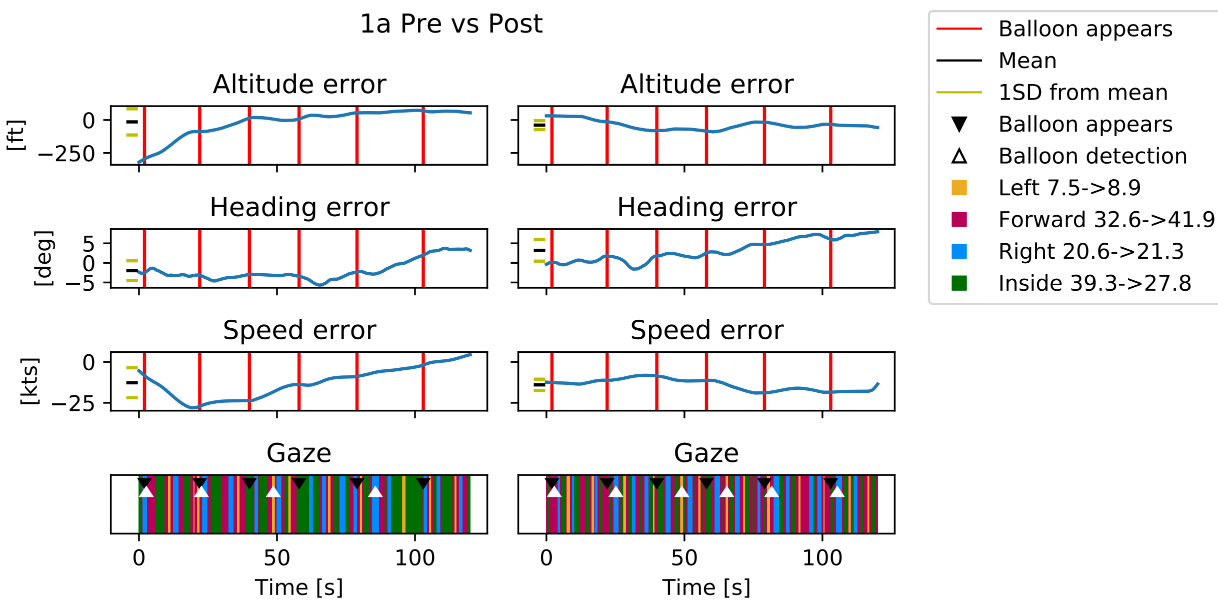


Fig. 29 Time traces for participant 1

C. Time traces for each participant

Fig. 30 Time traces for participant 2

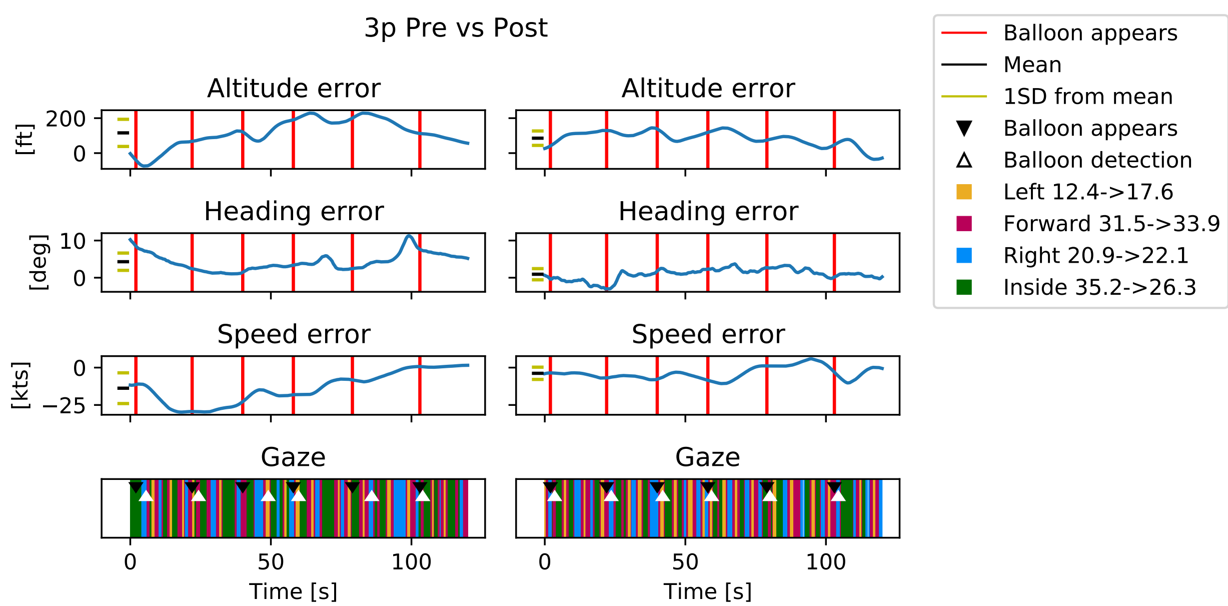


Fig. 31 Time traces for participant 3

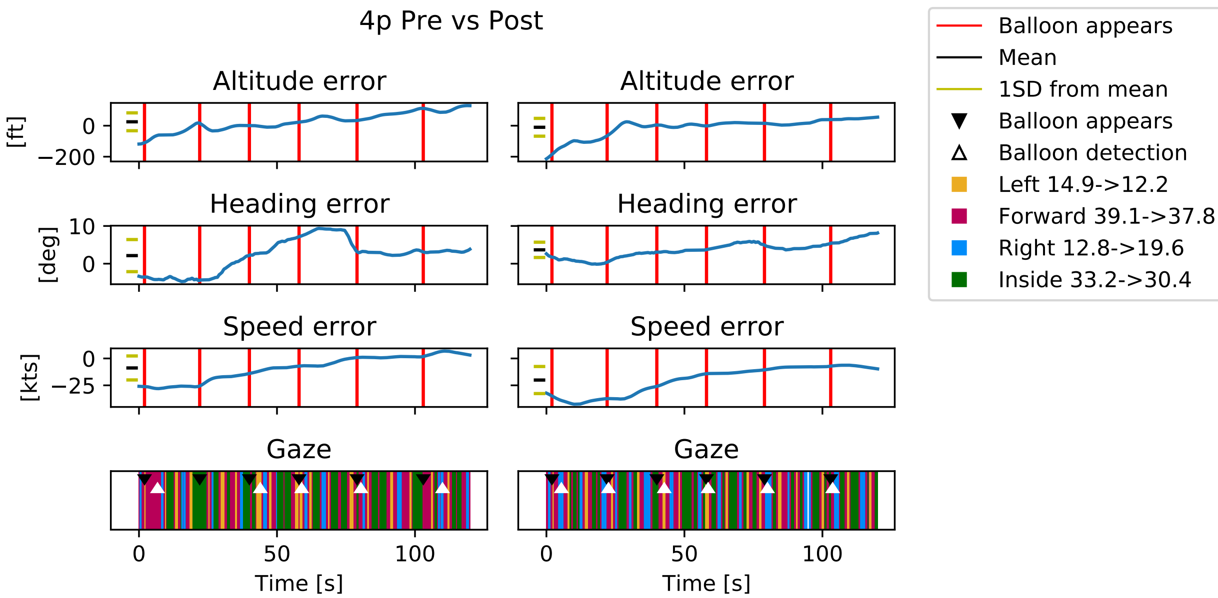


Fig. 32 Time traces for participant 4

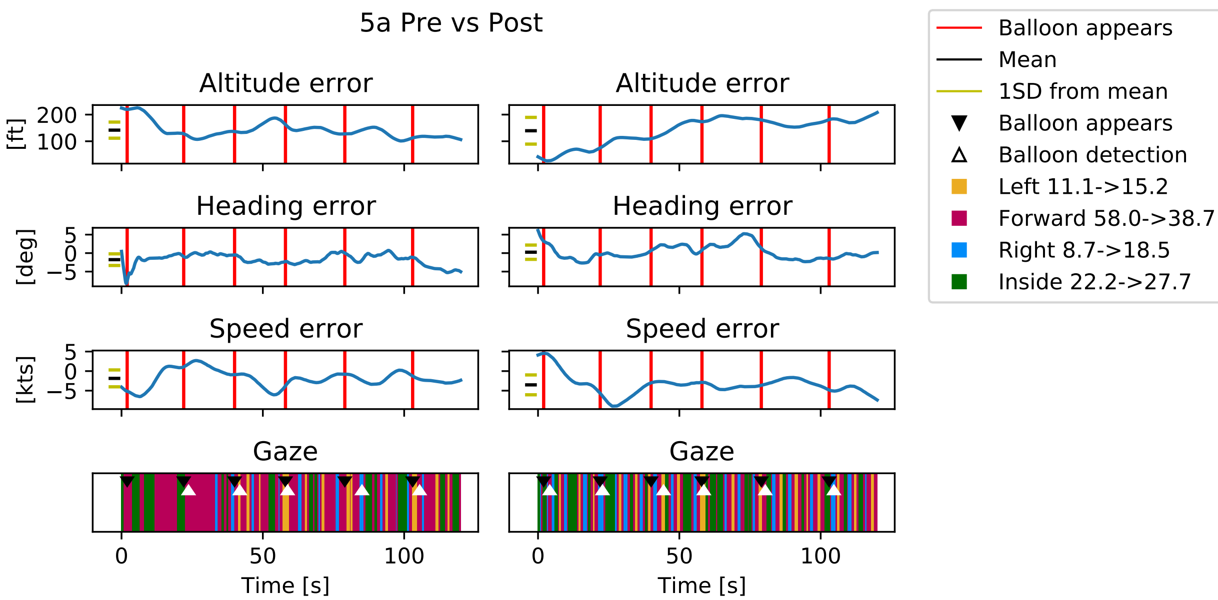


Fig. 33 Time traces for participant 5

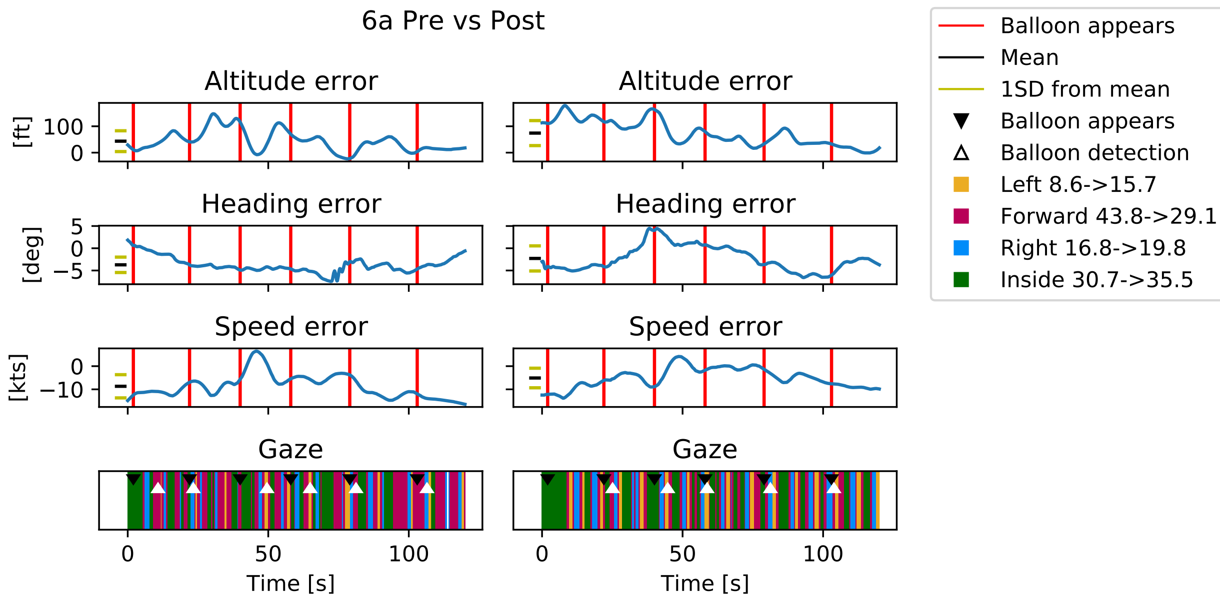


Fig. 34 Time traces for participant 6

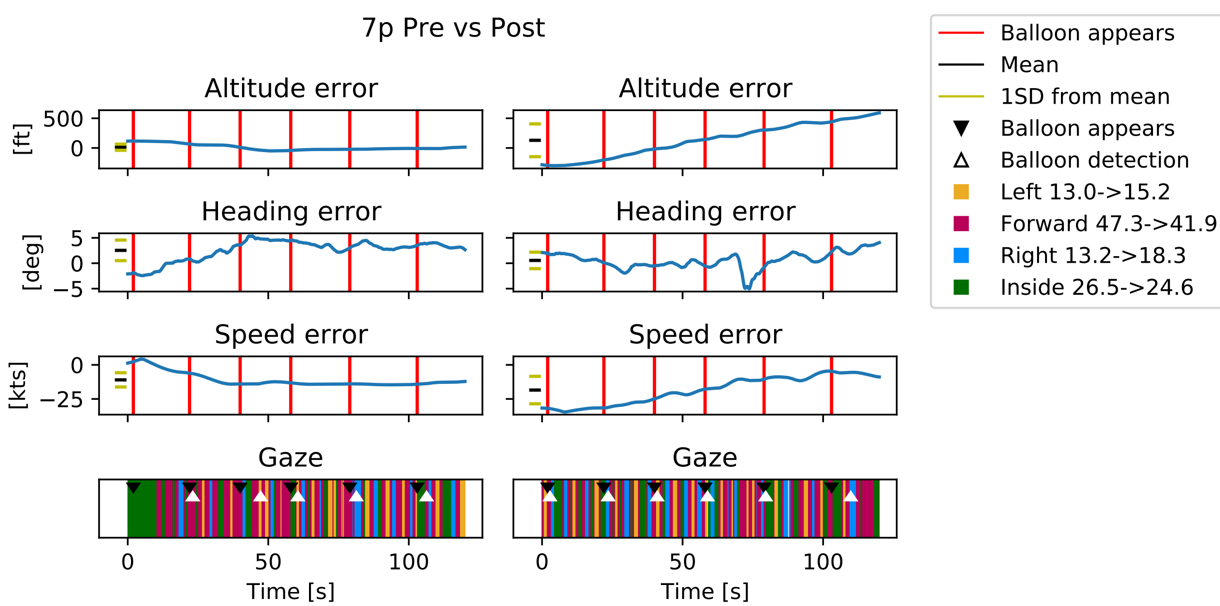


Fig. 35 Time traces for participant 7

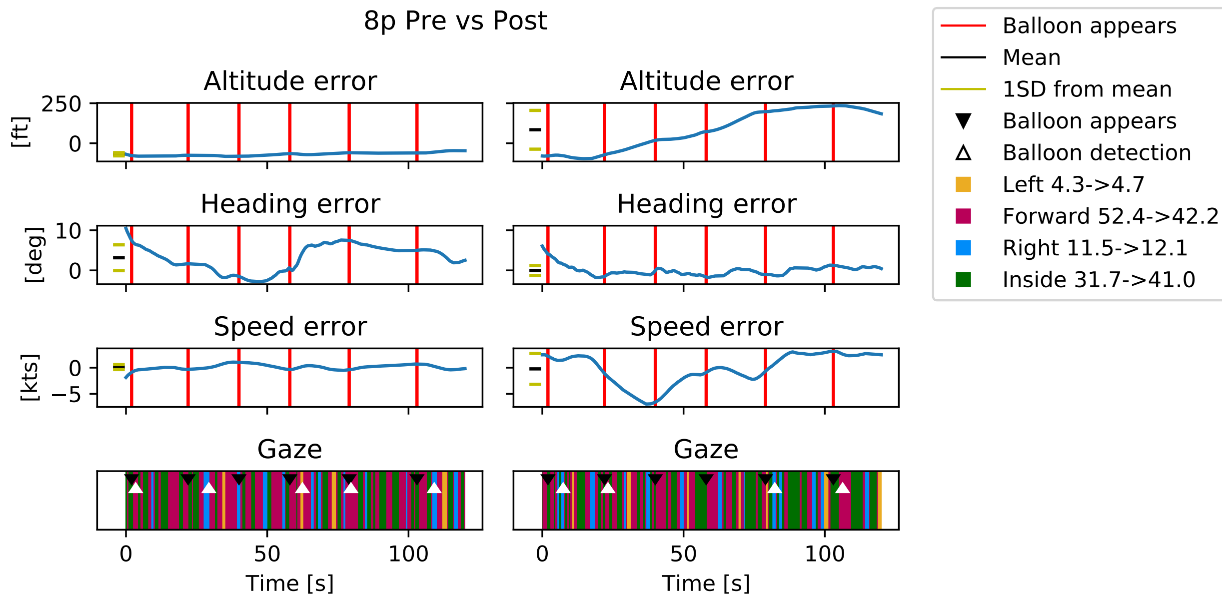


Fig. 36 Time traces for participant 8

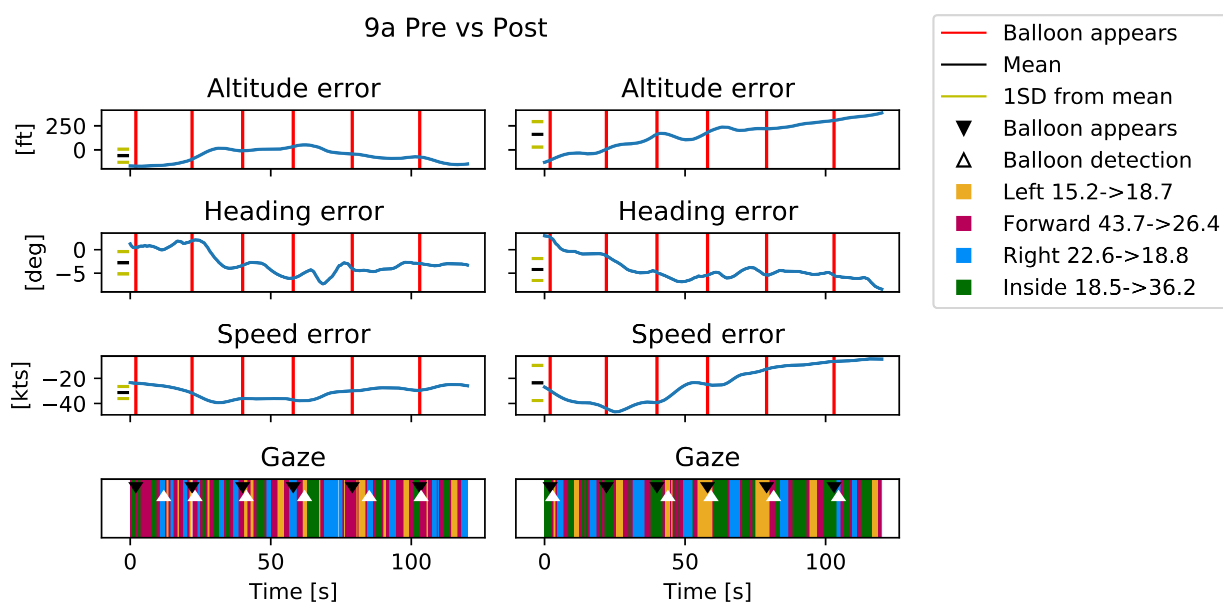


Fig. 37 Time traces for participant 9

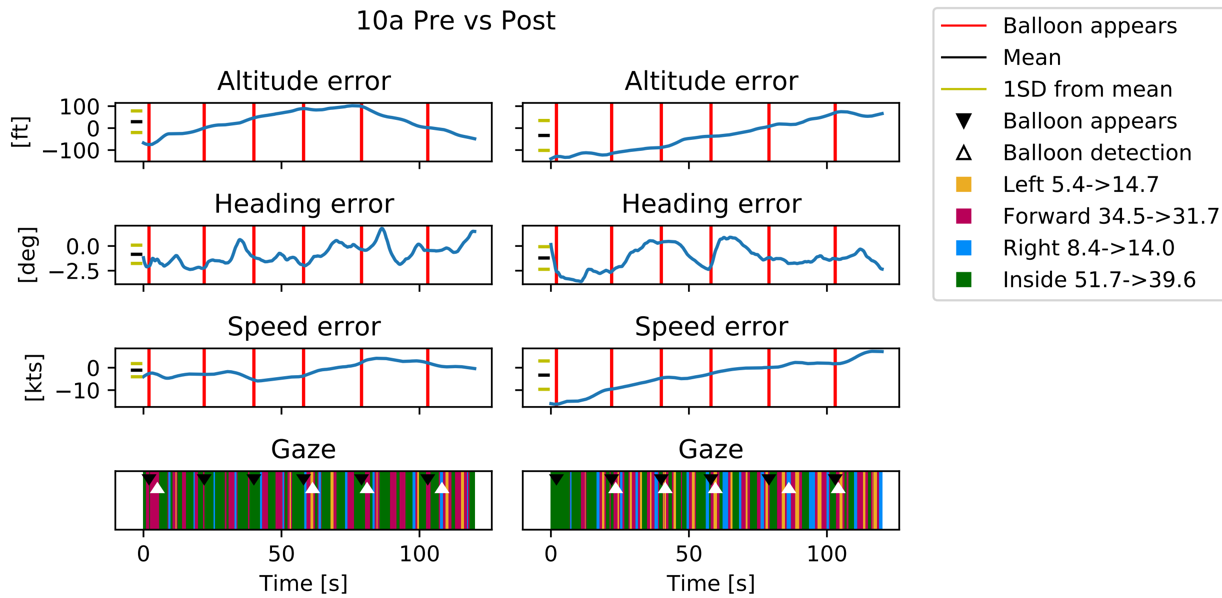


Fig. 38 Time traces for participant 10

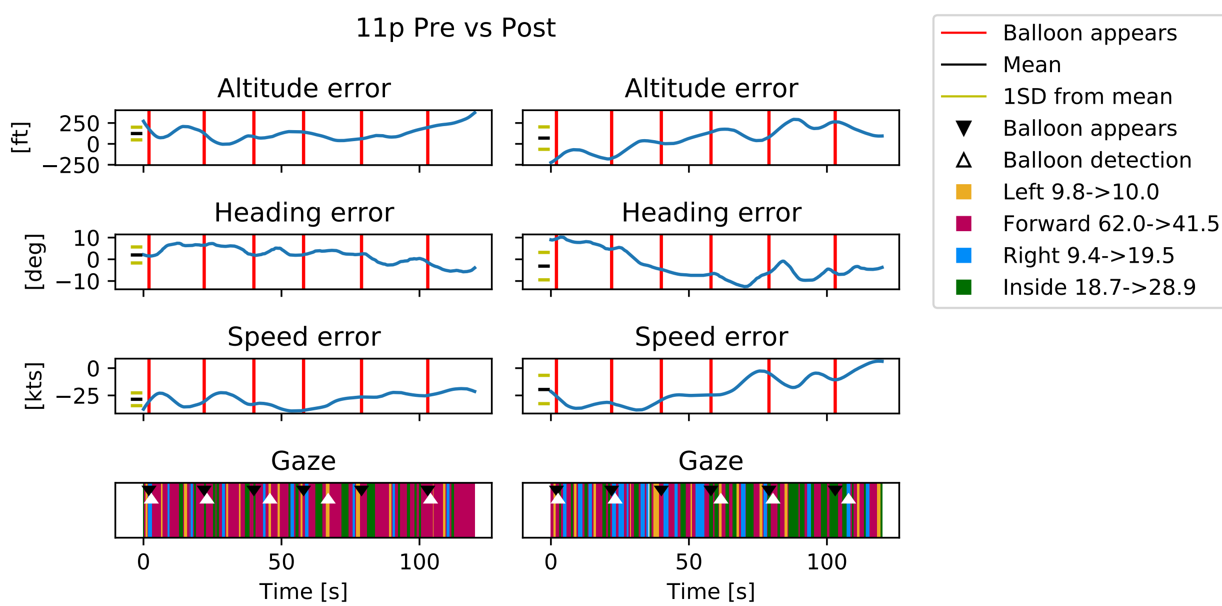


Fig. 39 Time traces for participant 11

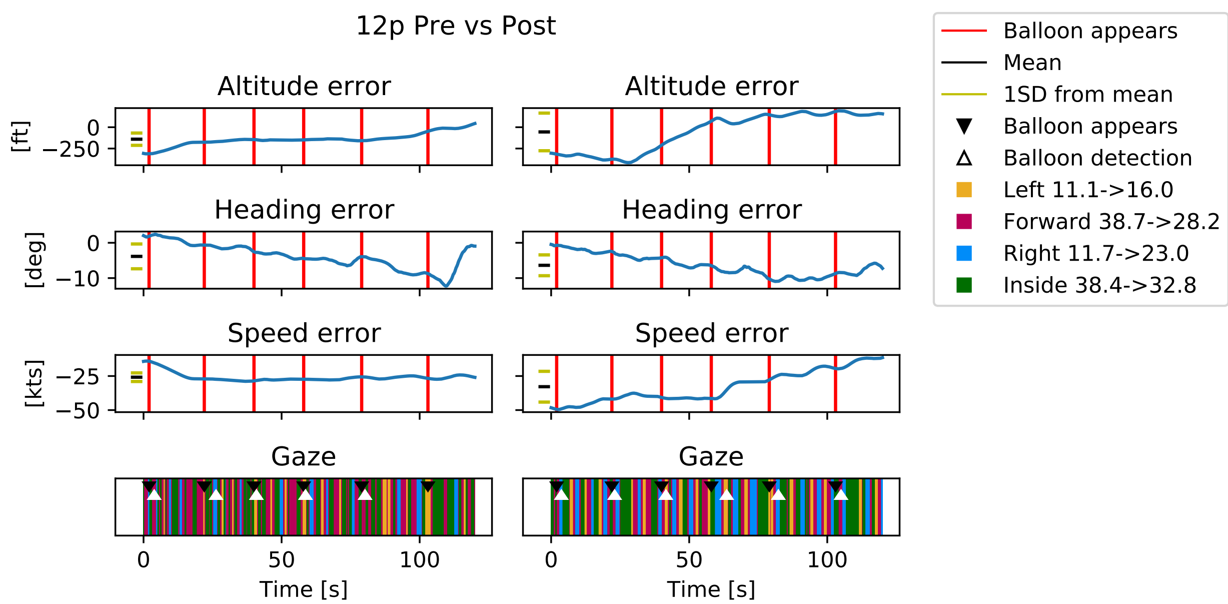


Fig. 40 Time traces for participant 12

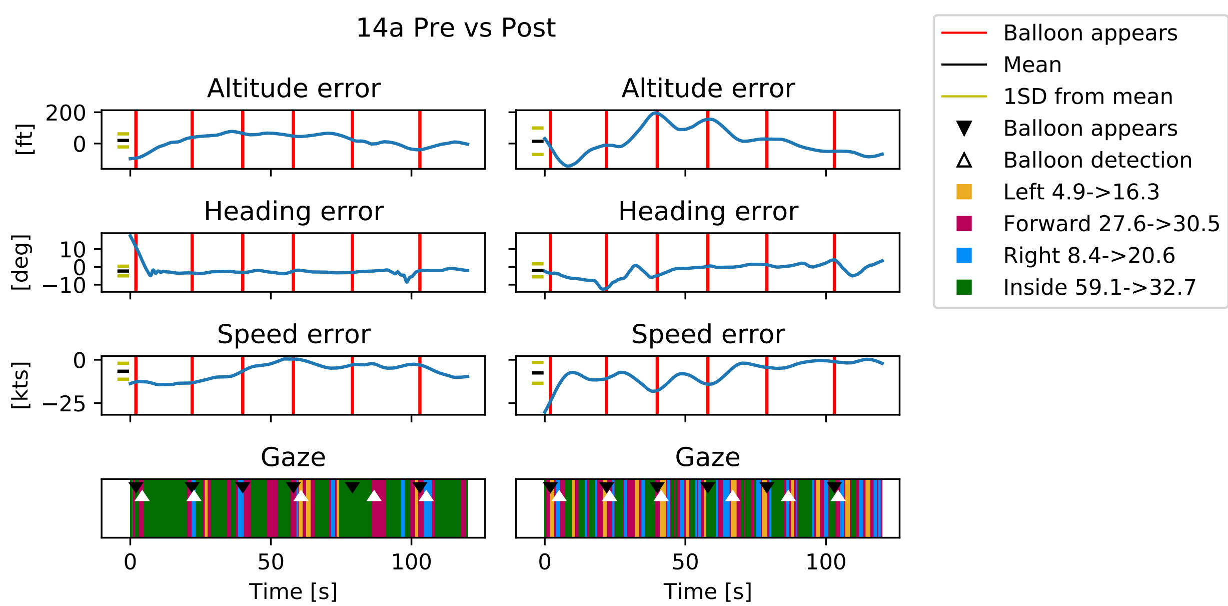


Fig. 41 Time traces for participant 14

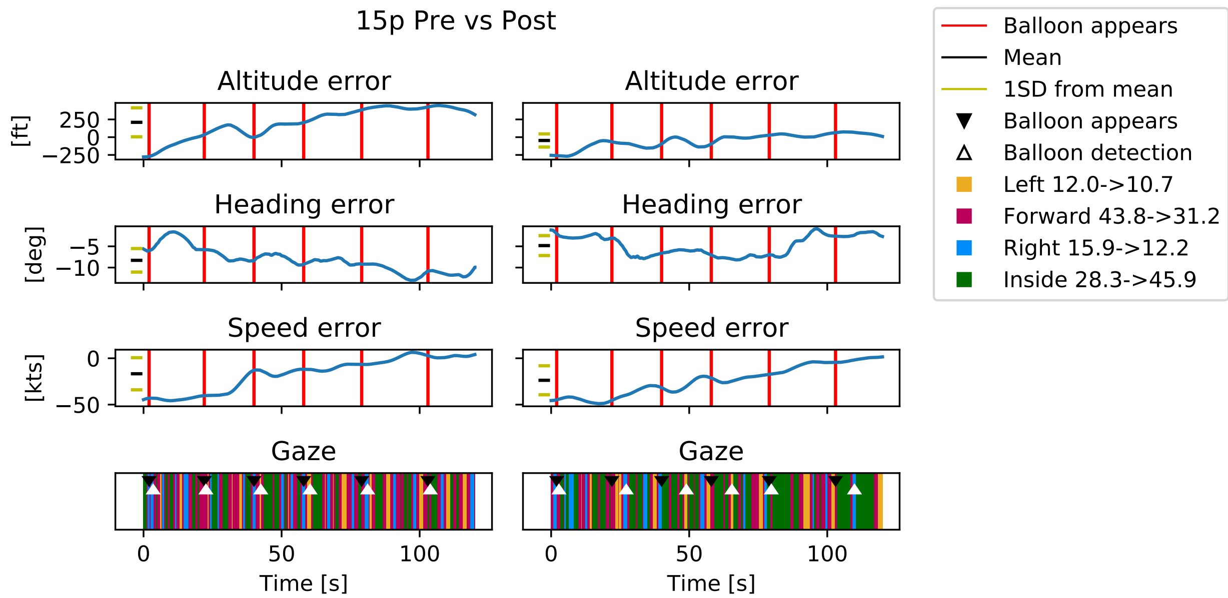


Fig. 42 Time traces for participant 15

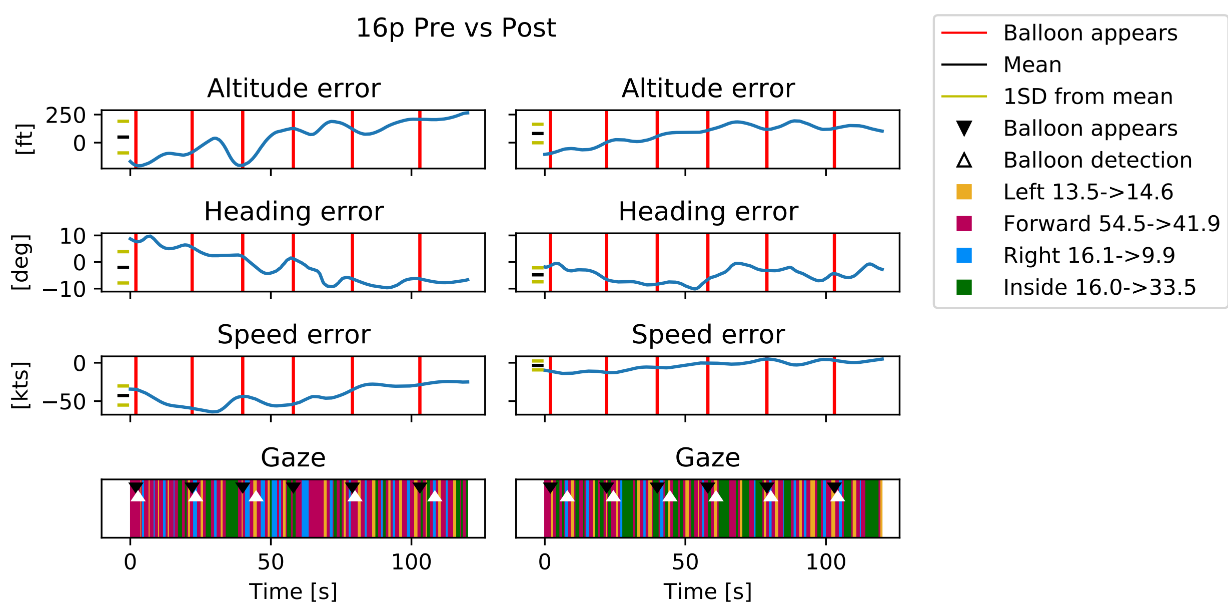


Fig. 43 Time traces for participant 16

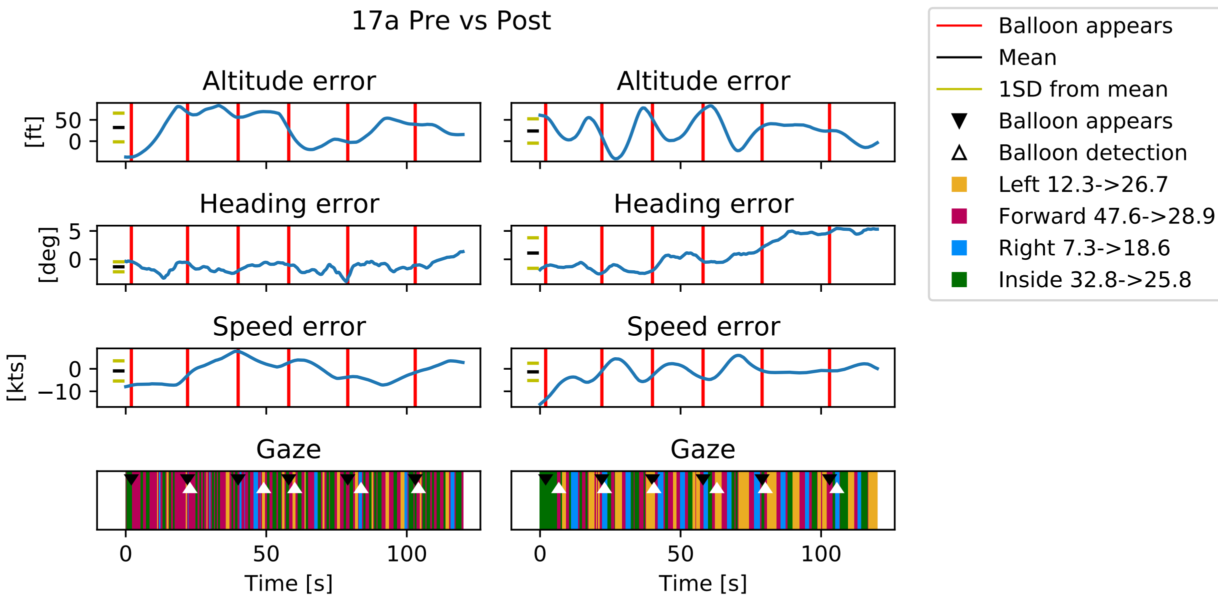


Fig. 44 Time traces for participant 17

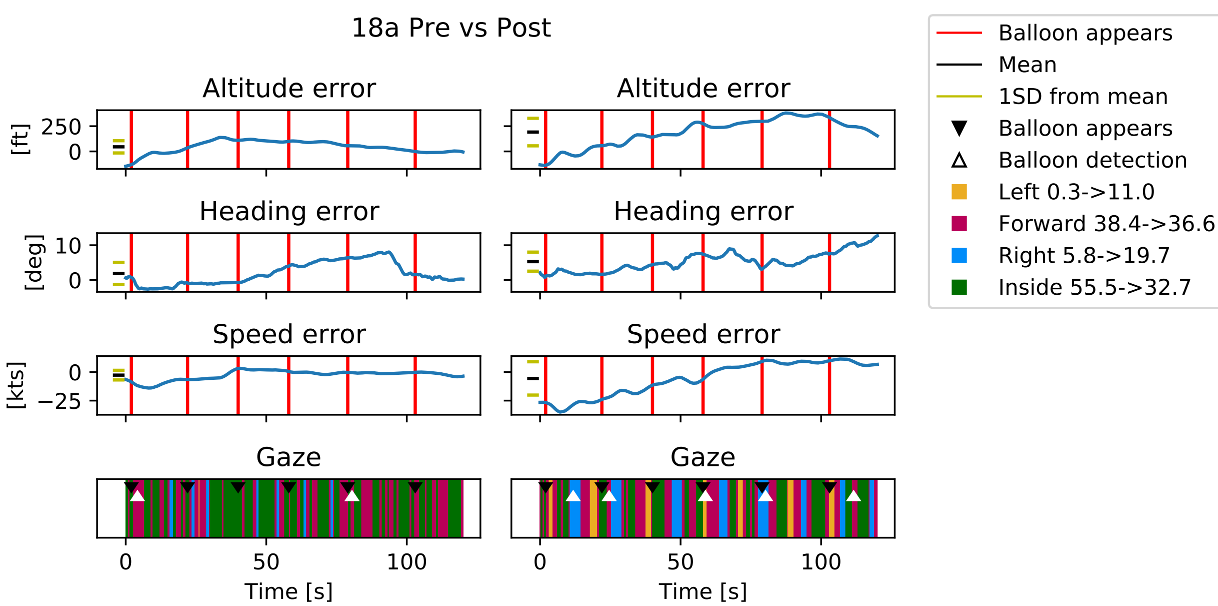


Fig. 45 Time traces for participant 18

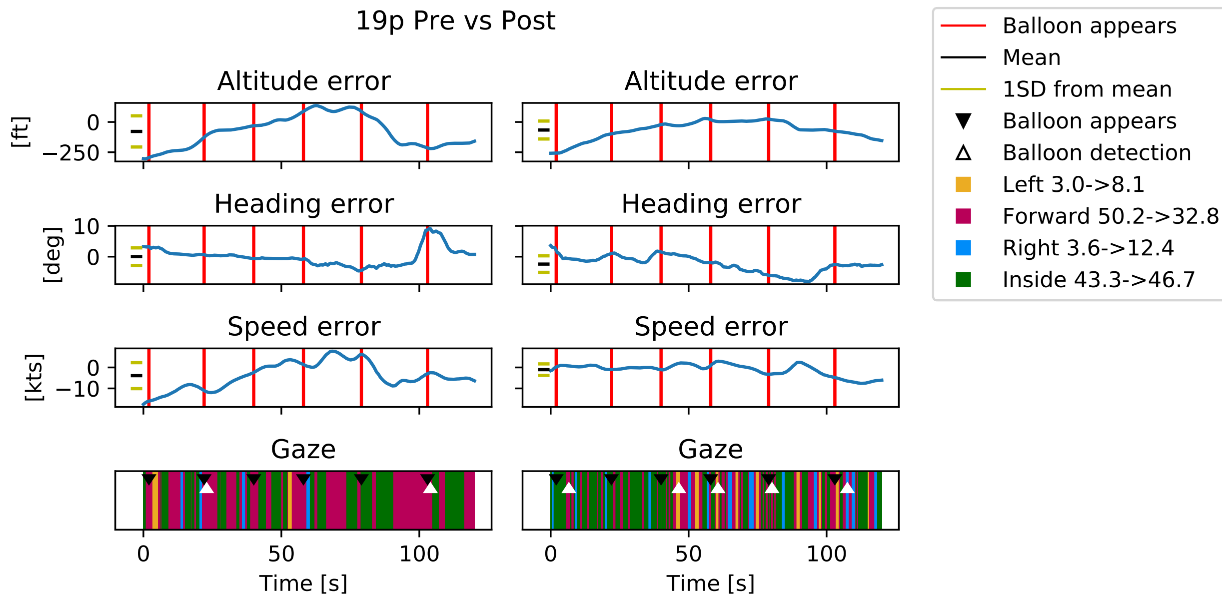


Fig. 46 Time traces for participant 19

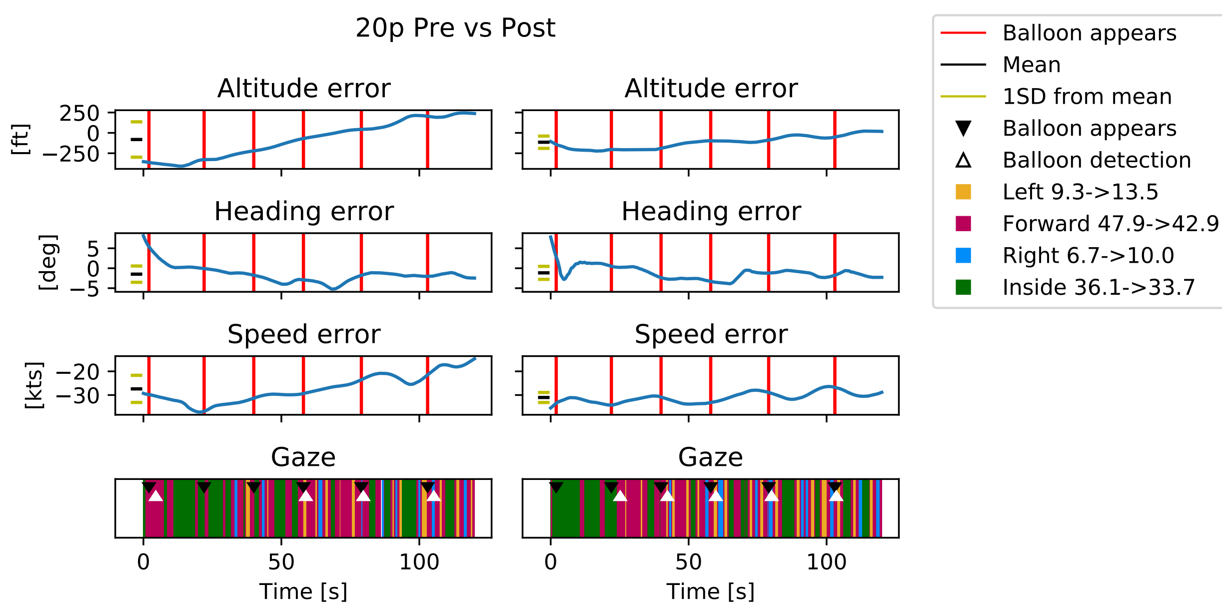


Fig. 47 Time traces for participant 20

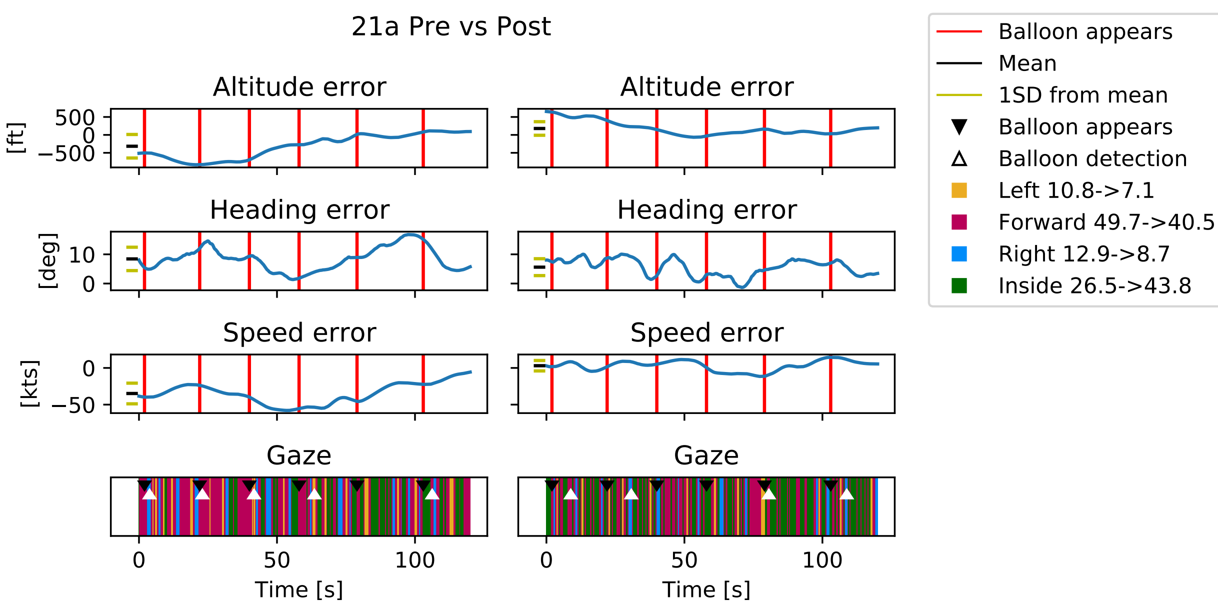


Fig. 48 Time traces for participant 21

D. Briefing

1

Covid

- Distance will be maintained between participant and researcher
- Cleanbox to disinfect VR headset
- Simulator disinfected between participants

multiSIM

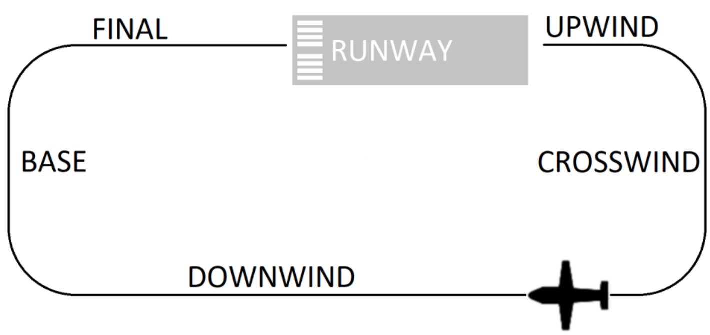
TU Delft

Experiment Phases

- Familiarisation
- Detection Test
- Gaze Training
- Detection Test

multiSIM
TUDelft

Basic circuit

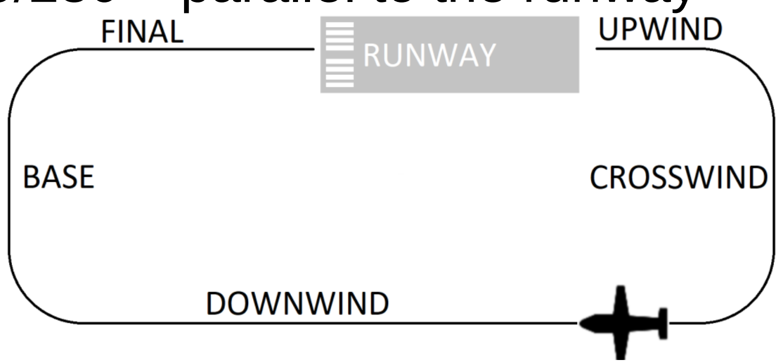


- Right handed
- Take-off at 80kts
- Cruise speed 110kts
- Cruise altitude 1500ft -> 1900ft baro
- Heading of the downwind leg: just under 50/230 deg
  - Parallel to the runway

multiSIM
TUDelft

Familiarisation

- Fly the basic circuit 4 times, landing in between
  - Right hand sided
  - 1500ft -> 1900ft on the barometer
  - Heading 50/230 – parallel to the runway
  - 110kts



multiSIM
TUDelft

Testing phase

- Stop on the runway
- Recalibrate the eye tracking procedure
- When entering the downwind leg, detection task will begin

multiSIM
TUDelft

Detection task

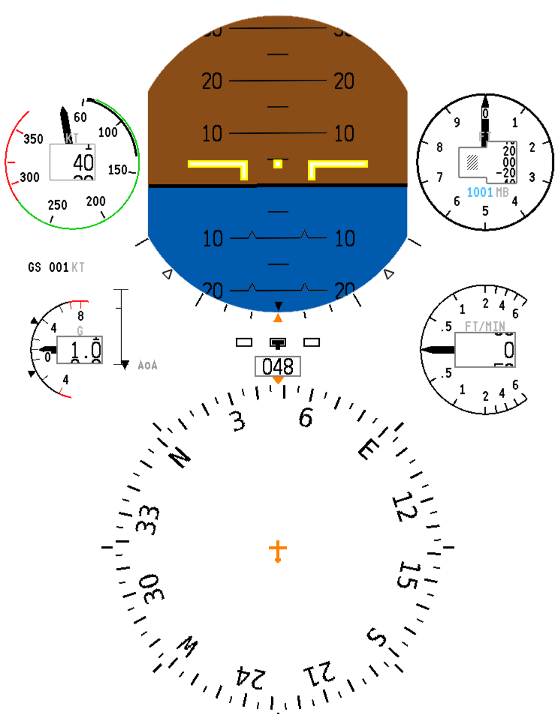
- A hot air balloon will be shown
- Upon spotting the balloon, pull the stick's trigger
  - Do this while still looking at the balloon
- The balloon will disappear

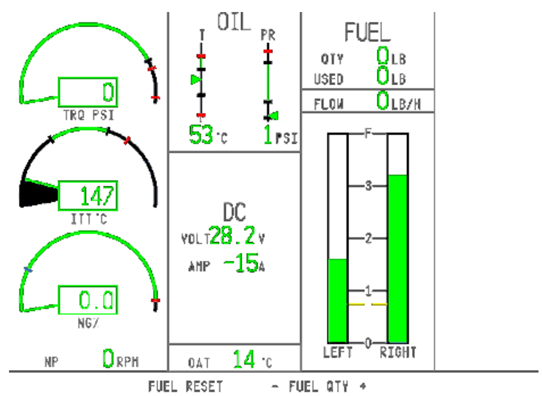
Training Phase

- Stop on the runway
- Will be shown a recording of expert behaviour
- Try to emulate the expert's gaze behaviour in your next flights
  - One more detection test

multiSIM
TU Delft

Instruments





multiSIM
TU Delft

Informed consent

- Anonymized data containing flight performance and eye tracker output
  - Shared with MultiSIM B.V., developer of the simulator
- Data will be used by the TU Delft for a MSc thesis, possibly papers
- Consent forms with personal information will be securely stored on TU Delft project storage
  - Only accessible by researcher

multiSIM
TUDelft

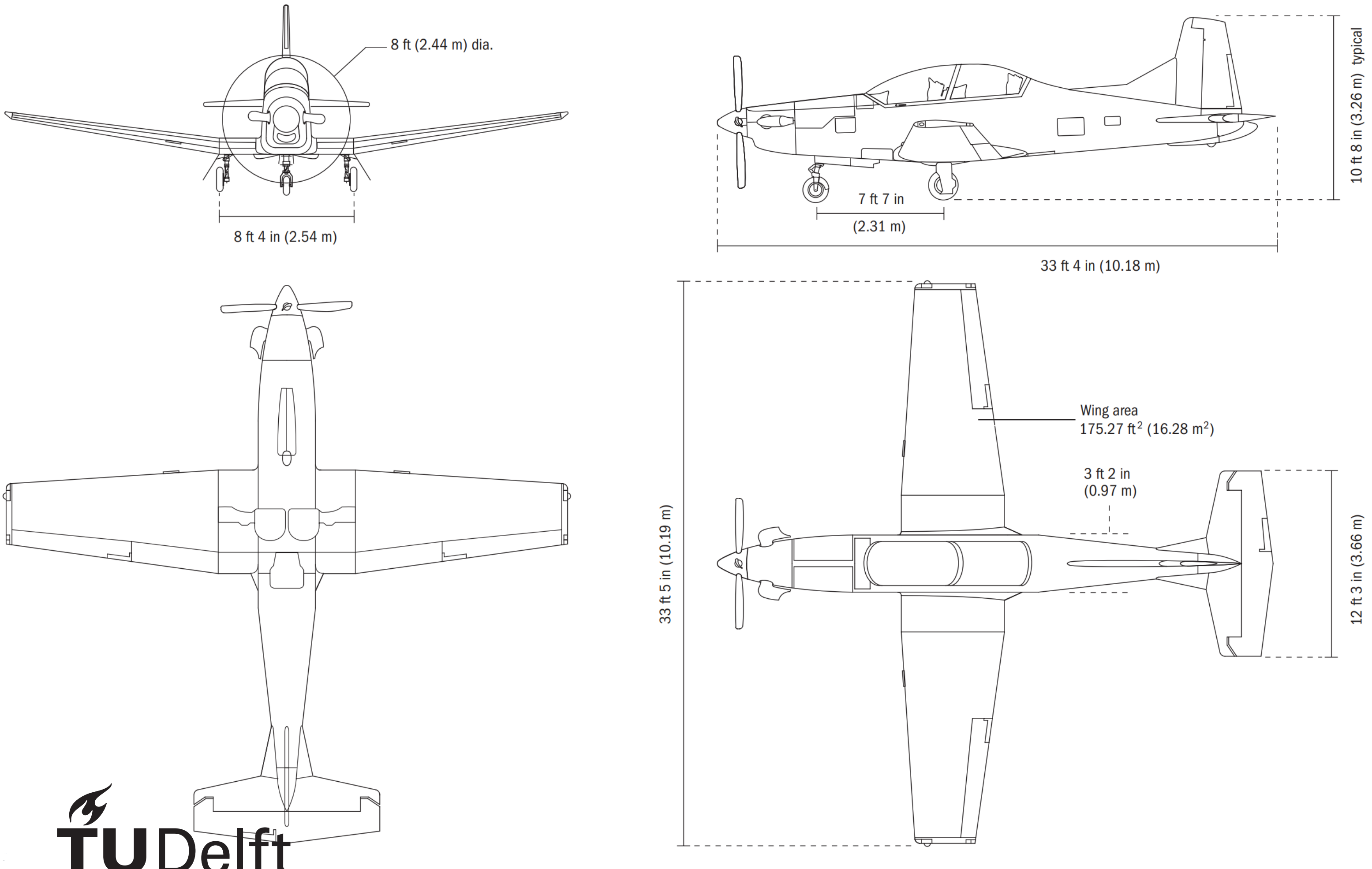
E. Preliminary Report

Eye Tracking

In ab initio pilot training

D. Verkooij

This report has already been graded as a part of AE4020 (Literature Study)



TU Delft

Eye Tracking

In ab initio pilot training

by

D. Verkooij

to obtain the degree of Master of Science
at the Delft University of Technology,
THIS REPORT HAS ALREADY BEEN GRADED AS PART OF AE4020

Student number:	4564251	
Project start:	February 12, 2021	
Thesis Supervisors:	Prof. dr. ir. M. Mulder,	TU Delft, supervisor
	Ir. O. Stroosma,	TU Delft
	Dr. ir. M.M. van Paassen,	TU Delft
	Dr. Ir. M. Wentink,	MultiSIM B.V.

TU Delft

Abstract

Scan patterns form an important part of pilot performance. A lack of proper scan patterns has been shown to be associated with several crashes in the past. Current training methods do not make active use of eye tracking in evaluating scan patterns, although previous research has shown that visualizing scanning patterns in augmented reality (AR) leads to promising results within the training process. Apart from AR, virtual reality (VR) is also increasingly likely to become a large part of training pilots in the future. The first VR flight training simulator has already been certified by the European Aviation Safety Agency. Modern VR headsets also contain eye tracking hardware, allowing for the integration of eye tracking without interfering in the training.

The training of first time pilots in the Dutch Military is currently done using the Pilatus PC-7 turbo-trainer aircraft. The PC-7 will however be replaced in the coming years, and the role of simulation within the training program is being reevaluated as well. Also providing an opportunity for introducing eye tracking. This thesis is a part of this investigation and details the development of a PC-7 simulation.

This thesis aims to review the current state of the art in eye tracking and use this to characterize the scan patterns of experts. The goal is to use this characterization to evaluate trainees, giving important insights into the behavior and areas of improvements for said trainees and enhancing flight safety in the process.

Additionally, this report details the initial plans for an experiment investigating the gaze behavior of experts. A group of expert flight instructors, is also available, allowing for the collection of valuable expert data. Ideally, this data will lead to a model of ‘proper’ gaze behavior. This report shows the (eye-tracking) tools developed for this experiment, together with the verification of these tools.

Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
ADI	Attitude Direction Indicator
AoA	Angle of Attack
AOI	Area of Interest
AR	Augmented Reality
ASI	Air Speed Indicator
ATC	Air Traffic Control
EASA	European Union Aviation Safety Agency
EFB	Electronic Flight Bag
EICAS	Engine-Indicating and Crew-Alerting System
EMME	Eye-movement modeling example
EMVO	Elementaire Militaire Vliegers Opleiding
GTE	Gaze Transition Entropy
HSI	Horizontal Situation Indicator
IFR	Instrument Flight Rules
LZC	Lempel Ziv Complexity
PFI	Primary Flight Instruments
SEEV	Saliency, Effort, Expectancy and Value
VFR	Visual Flight Rules
VR	Virtual Reality
VSI	Vertical Speed Indicator
VVI	Vertical Velocity Indicator

I

Research Context and Literature Study

1 Introduction

Simulation has already adopted a critical role in the training of pilots due to its safety, cost effectiveness and efficiency as illustrated by Allerton (2010). Virtual reality (VR) has not yet seen a large amount of use within pilot training. Virtual reality refers to the use of VR headsets in simulation to provide an immersive simulation environment at a lower cost than, for example, a full dome would provide. The slow rate of adoption for VR is in part due to the strict regulations set upon pilot training, but also due to the challenges in providing high fidelity interaction in VR. In April 2021, the first VR-based flight simulation training device was approved by the European Union Aviation Safety Agency (EASA). In recent years, VR has rapidly developed and seen increasing use in research.

Eye tracking and scanning patterns/gaze behavior, on the other hand, have already been the subject of research since the 1950's (Fitts, Jones, and Milton (1950)). This involved studying eye positions in pictures of the pilot after the experiment. Over the last decades eye-tracking technology has evolved to allow real-time tracking and has become more widespread. The use of eye-tracking in AR/VR has already been demonstrated to be an asset for training by Vlasblom and Rooij (2020). Furthermore, it has been shown that "inadequate monitoring and/or cross checking" has been the cause of several accidents as shown by NTSB (1994) (National Transportation Safety Board). The link between inadequate monitoring and/or cross checking and scanning patterns is solidified by Zaal et al. (2021), further showing the potential benefit in investigating scanning behavior.

An example of the current use of simulators within training is training procedures in a replica cockpit instead of inside the real aircraft. Other applications include maneuvers that carry a certain amount of risk, such as Upset Prevention and Recovery Training. Currently, instructors already affect the scan behavior of their students. Students are taught to make a sweep from right to center, look at the Attitude Direction Indicator (ADI), next the Air Speed Indicator (ASI) or Barometer, back to the ADI, and finally finish the sweep from center to left. In most turbo-trainer aircraft, the instructor is positioned behind the trainee. From this position, they cannot see where the trainee is looking currently, beyond the direction of their head. As mentioned, VR is also currently being adopted by the industry. Combining this with the fact that new models of VR headsets have an integrated eye-tracker, now is the perfect time to further investigate pilot gaze behavior for use in training.

1.1. Research Objective and Scope

The specific aim of this project is to use eye-tracking data to assist in the training of novice pilots, in particular those who have not yet fully completed training. Novice pilot gaze patterns will be analyzed, and attempts will be made to affect these gaze patterns to more resemble experts. Experts such as the instructors in the EMVO (Elementaire Militiaire Vliegers Opleiding) program or experienced jet fighter pilots.

The Royal Netherlands Airforce (RNLAF) currently uses Pilatus PC-7 aircraft for first time fliers at the EMVO. For this reason, a PC-7 simulation is developed as part of this project, and any experiments will be done using this simulation. Development of this simulation will be done in collaboration with MultiSim B.V., making use of their simulation software and hardware. MultiSIM is already involved with the training process of the EMVO, meaning knowledge on the current training procedures is available. Furthermore, the research done and tools developed can be used to assist other researchers in adopting eye tracking as a tool for research.

1.2. Research Question

The research question for this project is as follows; *To what degree can expert instrument scanning patterns be used to aid novice pilot training in VR simulation?* And as part of answering this question, the following two sub questions are investigated; *How can instrument scanning patterns (of expert pilots) be characterized?* and, *how can instrument scanning patterns be affected?*.

In order to move towards an answer for these questions, a 'preliminary' experiment was also performed. This experiment served to verify and validate the functionality of the tools developed in preparation for the final experiment.

1.3. Report Structure

First the literature is presented in Part I. First is Chapter 2, detailing the anatomy and functions of the eye. Chapter 3 discusses in more detail how eye tracking has been used, and how it can be quantified. Chapter 4 then discusses pilot training, and the existing uses of simulation and eye tracking within training. Lastly, Chapter 5 briefly discusses the possibility of predicting pilot eye movements. Part II discusses the preliminary research, which is done to prepare for the final experiment. Chapter 6 will detail the apparatus and general approach. Chapter 7 then details a preliminary experiment performed to test the established methodology. Lastly, Chapter 8 the current plans for the future of this project.

Appendix A elaborates upon the PC-7 simulation that was developed for use in this research. Appendix B provides some documents used in the mentioned preliminary experiment.

2
The Human Eye

This chapter serves as a brief introduction on the human eye. The anatomy of the eye will be covered, knowledge of which can aid in the understanding of the movements of the eye. The movements of the eye are then discussed. In eye tracking experiments the movements of the eye are the dependent variable being recorded, and knowing what eye movements to expect aids in processing this data.

2.1. Anatomy of the Eye

A cross-section of the human eye is shown in Figure 2.1. As explained by Kolb (1995), the human eye consists of 3 different layers. First is the external layer consisting of the sclera and cornea. Next is an intermediate layer, with an anterior (iris and ciliary body) and posterior (choroid) part. Last there is an internal layer containing the retina, which is the sensory part of the eye.

The retina consists of rods and cones, the latter of which is found in a higher concentration in the back of the eye at the fovea as explained by Hecht (1937). In contrast, rods become increasingly more common towards the periphery of the retina. Lamb (2016) compares the function of cones and rods. Cones can achieve more visual acuity, while rods have the benefit of being more efficient and functioning with less illumination. Light rays entering the eye are focused on the fovea. The fovea is the part of the eye which can perceive the highest detail due to a higher concentration of cones, as stated by Kolb, Nelson, Ahnelt, Ortuño-Lizarán, and Cuenca (2020). The rim of the retina also contains cones, as established by Williams (1991). The exact function of this ring of cones is still disputed, but it might aid in motion detection in the peripheral vision.

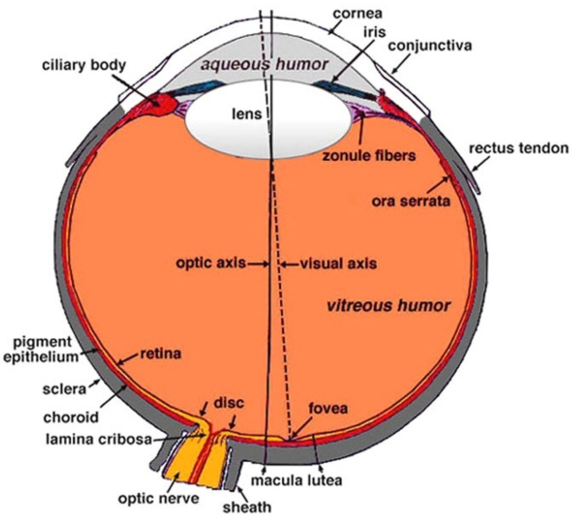


Figure 2.1: Cross-sectional view of the human eye [source: Kolb (1995)]

2.1.1. The Field of View

Understanding the field of view of the eye helps in understanding what information is available when the eye is focused at a certain point. Additionally, VR headsets do not provide full field of view, and understanding how the eye's field of view is built up can aid VR headset selection as in Subsection 6.2.2

According to Engbert, Longtin, and Kliegl (2002) the fovea is the inner $2°$ of visual angle from the center. The area around the fovea, described as the parafoveal area, is also able to take in information within $5°$ of visual angle from the center. This 'parafoveal preview' can already provide information. An example of parafoveal preview is the ability to read the words to the left or right of the word one is currently looking at, without moving one's eyes.

The remaining part of the field of view is referred to as the peripheral vision. The peripheral vision is perceived as more 'blurry' than the central part of the field of view. However, as Strasburger, Rentschler, and Jüttner (2011) describe, this is not a complete description of the peripheral vision. In fact, the peripheral vision is most capable when detecting moving stimuli as noted by To, Regan, Wood, and Mollon (2011). The detection of moving stimuli is likely related to the aforementioned rim of cones.

2.2. Eye Movements

Knowing what movements the eye can make helps making an informed decision in what algorithm to use in processing this data, as will be discussed in Section 3.2. The eye's movements can generally be classified in three different categories.

First, saccades are rapid eye movements, discovered by Javal (1878). A saccade can have a velocity of over 300 deg/sec, and during a saccade little to no visual processing takes place as established by Fuchs (1971).

Second, fixations on the other hand are when the eyes are not moving (or slowly moving to track an object in motion), during fixations detailed information can be taken in from the fovea.

Lastly, glissades are a third form of eye movement. Weber and Daroff (1972) discovered glissades (or glissadic corrective movement) to be a corrective movement at the end of a saccade, when the intended amplitude of the saccade is not achieved. According to Nyström and Holmqvist (2010), glissades occur in half of the saccades and therefore need to be taken into account in the processing of eye tracking data. One needs to decide whether or not these glissades count for the following fixation's duration or not.

3 Eye Tracking

The literature reviews on eye tracking by Ziv (2016) and Peißl, Wickens, and Baruah (2018) were used to find further literature.

Gegenfurtner, Lehtinen, and Säljö (2011) highlight that eye tracking allows for making observations on how experts, defined as those with “consistently superior performance on a specified set of representative tasks for a domain”, achieve superior performance. These observations on expert performance can help in further improving training protocols, establishing the importance of understand the use of eye tracking for this research.

3.1. Link Between Eye Movements and Pilot Information Intake

The use of eye-tracking in research is in part based upon the mind-eye hypothesis by Just and Carpenter (1976, 1980). The mind-eye hypothesis states that the object the eye is currently fixated on is also what the mind is processing. The mind-eye hypothesis is the basis of the assumption that the area currently fixated on is also the only source a pilot is taking information from. This hypothesis has been shown to be invalid by Fox, Merwin, Marsh, McConkie, and Kramer (1996), in which the removal of any instruments not currently being looked at had their information hidden. A change in gaze behavior and a reduction in performance was found. Ramón Alamán, Causse, and Peysakhovich (2020) also found that some information can be extracted from the peripheral. Pilots noticed changes in instruments that were located outside of the center of the pilot’s gaze location.

3.2. Identifying Fixations

The first step in the processing of eye tracking data is the identification of fixations, as shown by Salvucci and Goldberg (2000). During a fixation, the eye is kept stationary and information is processed. The use of fixations reduces data complexity (removal of tremors and drifts as elaborated upon by Ditchburn (1980)) and separates the saccade (eye movement) data from the fixation data. For this purpose, many eye tracking studies work with fixation within areas of interest (AOIs, sometimes referred to as Regions of Interests or ROIs), as is explained by Holmqvist and Nyström (2011). An AOI is an area in which some task-related data can be found. AOIs vary greatly between experiments, Dubois, Blättler, Camachon, and Hurter (2015) used two AOIs for example, those being the inside of the cockpit and the outside of the cockpit. However, research working with eye tracking involved a multitude of AOIs can also be found. This includes Van Dijk, Van de Merwe, and Zon (2011) tracking 5 AOIs, Xiong, Wang, Zhou, Liu, and Zhang (2016) tracked 4 AOIs and Ho, Su, Li, Yu, and Braithwaite (2016) tracked 5. Lefrancois, Matton, Gourinat, Peysakhovich, and Causse (2016) tracked total fixation duration over 10 AOIs, providing detailed insight into what each pilot is looking at. Altogether, it can be said that a low number of AOIs is more prevalent in research, while larger numbers do occur.

As explained by Nyström and Holmqvist (2010), the choice of algorithm for fixation identification also has an impact on the value of the average fixation duration. Algorithms that are not robust enough might for example not be able to deal with glissades, which are minor eye movements that can occur at the end of a saccade.

Two algorithms laid out by Salvucci and Goldberg (2000) are further highlighted here, due to their simplicity and robustness. These are velocity threshold identification (I-VT) and area of interest identification (I-AOI). Both can run in real-time while also providing fairly accurate results. I-VT dismisses all data points with a velocity below a certain value as fixations, the remaining points are classified as saccades. I-VT is shown in Algorithm 1. A value commonly used for the threshold in I-VT is 20 deg/s, a value that was found by Sen and Megaw (1984).

Algorithm 1 I-VTSalvucci and Goldberg (2000)

```
procedure Velocity threshold identification
    for each point do
        get velocity ((point-previous point) / dt)
        if velocity < threshold (20 deg/s) then
            label as fixation
        else
            label as saccade
    for each group of fixations do
        store centroid of fixation group
    return fixation group centroids
```

I-AOI, shown in Algorithm 2, labels all groups of points exceeding a minimum duration within an AOI as a fixation. Any other points are labeled as a saccade. I-AOI therefore also has the effect of getting rid of any fixations that are not within an AOI. The minimum duration for I-AOI is commonly set to 0.1 sec (Vlasblom and Rooij (2020), Nyström and Holmqvist (2010)). The output of I-AOI is commonly represented as a sequence of AOI, which can be used as input for generating measures as will be discussed in the next section.

Algorithm 2 I-AOISalvucci and Goldberg (2000)

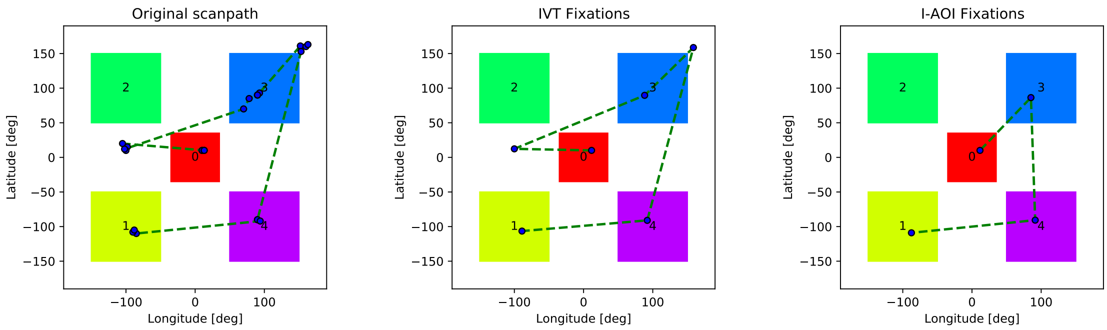
```
procedure Velocity threshold identification
    for each point do
        if point in AOI then
            store AOI in fixation group
    for each fixation group do
        if duration exceeds minimum duration (0.1sec) then
            store centroid of fixation group
    return fixation group centroids, corresponding AOIs
```

Figure 3.1 shows a comparison between an original scanpath (Figure 3.1a), a scanpath with I-VT applied (Figure 3.1b), and a path with I-AOI applied Figure 3.1c. This serves as an illustration of the effect each of these algorithms have on the scan path. To generate these images data points were generated with a sampling time of 0.05 seconds. An I-VT threshold of 20 deg/s was used, while I-AOI had a minimum duration of 0.1 second. The resulting sequence of AOI for the I-AOI case is 0-3-4-1, where each AOI that was visited for more than 1 time step counts as a fixation.

3.3. Quantifying Gaze Behaviour

This section discussed ways to characterize or quantify gaze behavior. Several measures were found in literature, each of which will be discussed here. Table 3.1 then collects these measures for convenience and shows an example of their usage in literature.

The simplest measure that can be used is the dwell time within an AOI. The dwell time is often expressed as the percentage of time spent fixating on an AOI. The methodology used research by Lefrancois et al. (2016) takes the mean dwell time over the duration of the task of pilots who performed well at an approach task as a baseline. This baseline is then compared with pilots who performed an unstable approach. With such a comparison, it becomes evident what can be improved in a specific pilot's behavior (over/under-focusing of certain elements). An example shown in the mentioned research shows that a pilot that did not successfully stabilize their approach massively under-focused on the area outside of the aircraft. What is still worth investigating here is if the gaze behavior is the



(a) Original scan path (b) I-VT fixations (c) I-AOI fixations

Figure 3.1: Comparison of original scan path and identified fixations

cause of the unstable approach, or if it is a consequence instead. If the deviating gaze behavior is just a consequence of the unstable approach, improving gaze behavior is also not likely to lead to a more stable approach.

Lounis, Peysakhovich, and Causse (2021) show several ways to evaluate eye tracking. These methods can be valuable in describing and evaluating gaze behavior. At the base of several eye tracking measures is the transition frequency matrix, $\mathbf{P}$, as is found using Markov Chains defined by Kang and Landry (2014). The transition frequency matrix can highlight which AOIs are closely intertwined. With i being the root AOI, j the destination AOI, S the range of AOIs, n_{ij} showing the number of transitions from i to j, and matrix element p_{ij} the probability of transitioning from i to j. The latter of which can be found using the following equation:

$$p_{ij} = \frac{n_{ij}}{\sum_{i \epsilon S} n_{ij}}, \quad i, j \epsilon S \tag{3.1}$$

If there are no transitions in a row, this leads to a division by zero. Such rows are set to $1/S$ to prevent dividing by zero. The following transition frequency matrix is the result of the 0-3-4-1 sequence found in the example in Figure 3.1c:

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Additionally, Goldberg and Kotval (1999) introduce the transition density matrix, which highlights which AOIs are not taken into account. The density matrix is found by replacing any nonzero n_{ij} with a one. This method, further elaborated upon by Ognjanovic, Thüring, Murphy, and Hölscher (2019) has seen most use within computer interface development. In aviation it can be used to highlight differences in instruments and transitions used between participants. This can highlight experience (e.g. an experienced pilot will have a relatively full row/column at the ADI, which is often used as a 'base' instrument). A particularly sparse matrix can also indicate tunneling behavior for a certain instrument. This density matrix can also be expressed as a percentage, representing what percentage of AOIs was visited by the participant.

The Gaze Transition Entropy (GTE) is another measure which can be derived from the transition frequency matrix. It was defined by Endsley (1995) based upon the entropy definition laid out by Shannon (1948), with the concept of entropy in the context of Markov chains being defined by Ekroot and Cover (1993). GTE is also sometimes referred to simply as entropy. The GTE evaluates the randomness of the process. Should the gaze behavior be fully deterministic (after looking at region 2, region 4 is always looked at), then the GTE is zero. If on the other hand, the gaze behavior is fully random (meaning each AOI has an equal chance of occurring from every root AOI), the GTE is equal to 1. As Holmqvist and Nyström (2011) state, the ideal measure for comparing scan patterns is a single value between 0 and 1, giving further credence to the GTE's value. The GTE is calculated with the following equation, in which $\hat{H}_t$ is the GTE and π_i are the stationary probabilities:

$$\hat{H}_t = -\sum_{i\epsilon S} \pi_i \sum_{j\epsilon S} p_{ij} \log_2 p_{ij}, \quad i,j\epsilon S \tag{3.2}$$

The GTE can be normalized as follows:

$$H_t = \frac{\hat{H}_t}{\log_2(S)} \tag{3.3}$$

The stationary probabilities, π_i, are elements of the stationary distribution π. Jarodzka, van Gog, Dorr, Scheiter, and Gerjets (2013) explain that the stationary distribution can be estimated through eigenvalue analysis. If $\mathbf{P}$ is regular, then over a large amount of transitions n, $\mathbf{xP}^n$ converges to π. With $\mathbf{x}$ being an arbitrary probability distribution. Thus, it holds that $\pi\mathbf{P} = \pi$, with π being the eigenvector corresponding to the eigenvalue equal to one.

Next is the Lempel-Ziv Complexity as defined by Lounis, Peysakhovich, and Causse (2020). LZC is a compression algorithm and can be used to measure the repetitiveness of the gaze data. This provides an insight into which patterns are present in the data. The downside with this method is that some very similar patterns are counted as different patterns.

Further, N-gram methods can highlight the basic patterns within a pilot's gaze behavior. An N-gram is a possible sequence or subset within a scanpath. A 3-gram is therefore a sequence of 3 consecutively scanned AOIs. In the example sequence of 0-3-4-1, the 3-grams of 0-3-4 and 3-4-1 both occur once. Reani, Peek, and Jay (2018) evaluated the effect of N-gram length within eye tracking analysis and concluded that shorter N-grams (3 being optimal) are better for analysis due to the increasing number of possible sequences with a larger N preventing the frequency from becoming large. Reani et al. (2018) then compare the frequency distribution of these N-grams for different groups. This statistical analysis was performed by evaluating the Hellinger distance of each group's N-gram distribution.

A consideration that needs to be made here is that a single incorrectly identified fixation will change an N-gram pattern. The algorithms presented by Salvucci and Goldberg (2000) should help prevent this, but it is something to keep in mind.

For the direct comparison of scanpaths between two participants in an experiment various algorithms exist. An example of such an algorithm is the string edit algorithm as presented by Levenshtein (1966). Holmqvist and Nyström (2011) explain that string edit methods are most suited for clearly defined, delimited AOIs with some separation between themselves. For this reason, the string edit method might not be appropriate for use within a cockpit with electronic flight instruments clustered closely together.

More refined sequence alignment measures also exist, based on the algorithms presented in Needleman and Wunsch (1970) and Chenna et al. (2003). These algorithms find the optimal alignment of two scan paths through a comparison matrix and gap penalties. An example of such an algorithm is the ScanMatch algorithm by Chenna et al. (2003). They found ScanMatch to be a robust alternative to the string edit methods in their evaluation through three experiments.

Finally, Krejtz, Duchowski, Krejtz, Szarkowska, and Kopacz (2016) show the method of Focal coefficients (or K-coefficients). Initially used for art, this coefficient aims to highlight ambient vs focal scanning. It indicates, in standard deviation units, the difference between a fixation's duration and the following saccade's amplitude. The ambient and focal modes referred to with the K-coefficient are characterized by Velichkovsky, Helmert, Joos, and Pannasch (2005). One problem with the focal coefficients is that due to the way they are calculated they may not be suitable for use in aviation, with looking outside leading to much larger saccades than on the instrument panel, likely leading to focal behavior inside and ambient behavior outside.

Table 3.1 presents the discussed measures and an example study that has used the measure to an example of its use. It also shows whether the measure is based on a string of AOIs or based on eye data (referring to fixations and saccades).

3.4. Experts vs Non-experts

Three theories in particular are referred to often in studies relating to eye tracking, together they explain how an expert achieves superior performance. Understanding the differences in how information is processed by experts and non-experts can help form a hypothesis. First of which is the long term working

Table 3.1: Relevant measures and a non-exhaustive list of studies in which they were used

Measure	AOI or Eye Data	Used in
Percentage of dwell time	AOI	Kasarskis, Stehwien, Hickox, Aretz, and Wickens (2001), Brams et al. (2018), Dehais, Juaneda, and Peysakhovich (2020)
Average fixation duration	AOI & Eye	Kasarskis et al. (2001)
Transition Frequency Matrix	AOI	Used as basis for other measures
Fixation Frequency	AOI & Eye	Faulhaber and Friedrich (2019)
Transition Density Matrix	AOI	Goldberg and Kotval (1999), Ognjanovic et al. (2019)
Gaze Transition Entropy	AOI	Krejtz et al. (2015)
Lempel Ziv Complexity	AOI	Lounis et al. (2020)
N-gram patterns	AOI	Reani et al. (2018)
Focal (K-) coefficients	Eye	Krejtz et al. (2016), Lounis et al. (2021)

memory found by Ericsson and Kintsch (1995). This states that experts are able to store processed information in their long-term memory and access this information through cues in their short-term memory. Second is the information reduction hypothesis as described by Haider and Frensch (1996), proposing that experts neglect task-irrelevant information, hereby reducing amount of information to be processed. Lastly, Kundel, Nodine, Conant, and Weinstein (2007) explain the holistic theory/model of image processing which states that experts also use the parafoveal area of their vision to gather information. The parafoveal information is used to make an initial global analysis of a situation after which finer details can be extracted more purposefully.

In addition to the three aforementioned theories, papers such as Reingold, Charness, Pomplun, and Stampe (2001) also indicate that experts process data differently when compared with novices. Vlasblom, Zon, and van der Pal (2020) also conclude that it is not desirable to have novices directly replicate expert behavior due to them not processing information as efficiently as experts would. In contrast, Chapter 4 established that showing expert gaze behavior to novices can lead to increased performance.

Brams et al. (2018) also confirms that the three theories capture aspects of the differences novices and experts. Additionally, it is found that less experienced pilots tend to have a longer, less efficient scanpath, moving between distant AOIs.

3.5. Relation Between Eye Tracking and Pilot Condition

Van Dijk et al. (2011) established that dwell times and GTE obtained with eye tracking can be used as an indicator of pilot situation awareness (using expert commercial pilots with mean flying hours = 7450h). In this study, it was found that increased gaze entropy indicated decreased situation awareness.

Faulhaber and Friedrich (2019) also evaluate pilot (mean flight hours = 4500h) workload using eye tracking. It was observed that an increase in workload led to a significant increase in fixation frequency. This link between workload and gaze patterns (and performance) is also established by Friedrich, Lee, Bates, Martin, and Faulhaber (2021), which evaluated novices, experts and pilots. Additionally, Allsop and Gray (2014) show that (novice) pilot anxiety causes an increase in gaze transition entropy. The two papers suggest that stressful and low situation awareness situations increase gaze entropy. However, Diaz-Piedra et al. (2019) showed that for fighter pilots, the gaze behavior became more deterministic when faced with complex control tasks (emergencies). This experiment was done by evaluating the GTE during different levels of complexity. The AOIs were defined by labeling each 1x1 degree area of the field of view as an AOI. The likely explanation for the reduction in entropy is that fighter pilots follow clearly defined checklists in case of emergencies.

Johnson, Wiegmann, and Wickens (2006) demonstrated that in cockpits with advanced display systems, pilot situation awareness is negatively affected by the advanced display systems drawing the pilot's gaze. As such, it becomes evident that too much time looking inside the cockpit is detrimental to flight performance. This is analogous with the aforementioned tunneling analysis.

Heiligers, Van Holten, and Mulder (2009) shows eye tracking can be used as an indicator of workload by looking at the distribution of dwell time spent on instruments in the primary scanning cycle. Although the hypothesis posing a relationship between workload and dwell times was not confirmed due to a small sample size.

4 Training

Training simulators for the PC-7 currently include a procedure simulator, which is a replica cockpit without any flight elements attached, a 360 degree video recording of a flight that can be watched in VR, and a full flight simulator. So while there is some use of simulation within the program already, flight training is done in the aircraft. The instructor and trainee share the controls, and basic maneuvers are taught to the trainees. In this gap between procedure simulation and full flight training, eye-tracking during a fully simulated flight might find its place. This could be with the instructor simply being shown the current gaze location of the student.

For training, it is important to consider how proper scan behavior could be used. It is very likely that simply ‘forcing’ a scan pattern will not lead to improved performance, as the optimal scan pattern is likely to be task specific. A question that logically follows is whether the difference in scan pattern is not simply a result of the difference in controller effectiveness. The latter would mean novices need to be better at the task before developing a ‘proper’ scan pattern.

Part of the challenge of using gaze behavior for training is making the eye tracking data usable for trainees and/or instructors without requiring a background in research. To this end, Niehorster, Hildebrandt, Smoker, Jarodzka, and Dahlström (2020) suggest aggregating eye tracking data at a higher abstraction level. Section 4.2 shows research within pilot training. After which, Section 4.3 shows some research in which the location of fixations are shared between different participants. Next, Section 4.4 shows research investigating the use of example gaze behavior models. Lastly, Subsection 4.4.1 briefly discusses the possibilities of providing pilots insight into their own scanning behavior.

4.1. Current EMVO Training

Currently, gaze behavior is part of the EMVO curriculum, with a standard scanning pattern being taught. However, the exact scanning behavior cannot be observed by the instructor as they are seated behind the trainee. Therefore, VR training can lead to an increased amount of information being available on the way a student scans. Additionally, the use of gaze behavior in the EMVO curriculum provides additional information which can be used in affecting the scanning behavior of students. For example, the default scanning pattern taught to students in the EMVO curriculum can be used as a basis for a scanning pattern to be taught to first time fliers.

The aforementioned standard scanning pattern involves making a sweep over the horizon from behind the right wing to behind the left wing. When looking straight ahead during this sweep, the ADI and a single other instrument (ASI, Baro) are checked, after which the sweep is continued.

4.2. Use of Eye Tracking within Pilot Training

Vlasblom and Rooij (2020) show the development of an AR eye tracking support tool called ‘augmented eye’. The tool aims to help make use of eye tracking data intuitive and easy. It focuses on two types of analysis, referred to as gap analysis and funneling analysis. The former looks at the under focusing of a certain instrument, that according to flight stage relevant checklists should be looked at. While the latter highlights the over focusing of an instrument instead. The prototype of this tool was received positively by instructors who participated in a trial phase, highlighting the value eye tracking can add to training procedures.

Dubois et al. (2015) have shown that novice pilot gaze behavior can be changed to be closer to expert behavior. Novices were shown to have a tendency to look at their instruments for too long when compared to experts. Over-focusing on instruments is established to be detrimental to performance by Johnson et al. (2006). Dubois et al. (2015) attempted to resolve this by implementing an aural alert when the time spent looking inside the cockpit exceeded 2 seconds, redirecting the participants' gaze outside. This did in fact help reproduce the desired 'expert' gaze behavior. It should be noted that there was no investigation in whether this led to improved gaze behavior after the removal of the alerts. Dubois et al. (2015) did not investigate whether a change in performance accompanied this change in scanning behavior.

4.3. Fixation Location Sharing

Chetwood et al. (2012) used an interface in which the instructor's eyes were tracked during laparoscopic surgery simulation. Trainees using the simulator could see a cursor where the instructor was looking, allowing for enhanced communication. The groups with the eye tracking interface performed significantly more efficiently than the group with verbal instructions only.

Similarly, Velichkovsky (1995) found that in a cooperative problem solving environment, in which one participant knew the solution while the second had to act, the sharing of fixation locations would increase the efficiency of distributed problem solving. This increase in efficiency was found regardless of which participant could see the other's gaze data.

4.4. Example-based Training

One method of using eye tracking to support training is using eye-movement modeling examples (EMMEs). EMMEs involve showing the gaze behavior of an expert to novices to enhance training procedures.

Kang and Landry (2014) have shown the potential of using gaze behavior within aircraft traffic control (ATC) training. Expert gaze behavior was shown to novices to improve their performance. This was successful, evidenced by the gaze-trained group's increased performance relative to the two different control groups (instruction only and no treatment or instruction).

Wilson et al. (2011) shows how a group of medical trainees was trained by showing the gaze of an expert performing the simulated task. In comparison with the second group, which trained on the control of the tools instead. Retention and transfer tests showed that the gaze trained group was able to perform the task more efficiently than the movement trained group.

Clinical reasoning based on visual observation of patients requires proficiency in visual search. Jarodzka et al. (2012) set up an experiment in which the control group was taught verbally, while two other groups were shown a visualization of an expert model's eye movements. One group had eye movements highlighted with a circle, while the other group was shown an image in which features not currently focused on were blurred. Jarodzka et al. (2012) found that participants from the latter condition, referred to as a spotlight EMME, showed improved visual search and interpretation performance relative to the other two groups.

Jarodzka et al. (2013) also used EMMEs to enhance the training in classifying fish locomotion of two out of three experiment conditions. One of the conditions indicated the gaze with a spotlight and one with a dot. Both EMME groups showed improved visual search and enhanced interpretation of relevant information.

It is likely that the complexity of the expert gaze behavior needed for the task affects the degree to which EMMEs can be used. As an example, expert gaze behavior during laparoscopic surgery can be described as focusing mainly on the object being manipulated, as explained by Vine, Masters, McGrath, Bright, and Wilson (2012) and Law, Atkins, Kirkpatrick, and Lomax (2004). Novices on the other hand tend to switch between the tools and the object being manipulated. The point being that pilots have more instruments to focus on, leading to a more complex 'expert' scan pattern.

4.4.1. Meta-cognitive Reasoning

Eye tracking can also be used to give pilots insight into their own scanning pattern. Van Gog, Paas, Van Merriënboer, and Witte (2005) state that this can enhance pilots' mega-cognitive reasoning, or self-reflection. This is also discussed by van Gog and Jarodzka (2013), stating that reviewing one's own gaze behavior can lead to more meta-cognitive statements about their own task execution.

5

Predicting Pilot Gaze behavior

In the process of forming an hypothesis it would be useful to have some predicted data available. Predicted data can give some insight in what is likely to happen in certain conditions. For this reason, research was done into the creation of an eye-tracking simulator that could predict how a pilot would behave in the cockpit.

5.1. Models

Carbonell (1966), building upon Senders (1964), presented one of the first methods used to explain the behavior of a pilot when sampling information from instruments. The method is based in queuing theory in which the next instrument to be looked at is the one with the highest value cost function. This cost function consists of the probability of exceeding a threshold and the cost of exceeding said threshold.

C. D. Wickens, Goh, Helleberg, Horrey, and Talleur (2003); Horrey, Wickens, and Consalus (2006) developed an extension of Carbonell's model. This extended model, called the SEEV model (Saliency, Effort, Expectancy and Value), uses more information such as the theories on salience presented by Itti and Koch (2000). The Expectancy and Value are the probabilities and values of the cost function in in Carbonell's model. These parameters are combined to retrieve the probability of looking at a specific instrument. Wickens' paper showing the usage of the SEEV Model in two experiments achieves a R^2 of 0.79-0.96 when predicting the % dwell time of the pilot over the three AOI. The required values for a simulation are the bandwith for each AOI, and the relevance of each AOI to each task.

A simulation with three AOIs required 11 parameters (Bandwidth for each AOI, relevance to two tasks for each AOI and the priority of two tasks). Furthermore, the different AOIs investigated (Instrument panel, outside world, cockpit display of traffic information) each had a distinct separate function, making it less suitable for situations in which specific instruments on the instrument panel are relevant. The priority of the tasks is ranked ordinally in the aforementioned simulation (C. D. Wickens et al., 2003). Aviate is assigned a priority value of 3, navigate a value of 2 and lastly communication is assigned a value of 1. The relevance values are then assigned using cognitive simulation. The bandwith values are then found using linear model fitting.

Further extensions of the SEEV model exist, such as NSEEV (Noticing SEEV) by C. Wickens (2015). NSEEV spefically aims to predict the latency and accuracy in the noticing of discrete events during scanning as predicted by SEEV.

5.2. Relevance

The intention of this chapter was to try to predict time-traces to aid in the forming of an hypothesis. With the research mentioned in this section this was deemed to not be suitable for the problem at hand as the models that exist are not expansive enough. As an example the SEEV models only output a percentage dwell time per distinct AOI.

As mentioned, an experiment will be presented in Chapter 7, this experiment has 11 participants and only one set of data per participant. The SEEV model for the 11 tracked AOIs would have 35 variables. When applying the model, the division of the considered would have to be reduced to be similar to those presented by C. D. Wickens et al. (2003). With each AOI serving a more distinct purpose. Doing so might make it viable to fit the model to the experiment group to gain insight into the priority of each AOI per task. The mentioned tasks, can be considered to be scanning outside and controlling the aircraft. One way of reducing the considered AOI would be reducing to the primary flight instruments, engine instruments and outside visuals. The variables that are set through cognitive analysis might still prove problematic in that case.

II

Preliminary Research

6 Approach

This chapter details the purpose of the experiment, the tools developed and used in the initial stage of this project as well as the hardware used. Section 6.1 identifies the research gap and research phases. Section 6.2 will briefly detail the simulation hardware and software, while Section 6.3 details the eye tracking software tools that were developed.

6.1. Research

The literature presented in Part I leaves a research gap. While attempts at changing gaze behavior are made, such as Dubois et al. (2015), they are not accompanied with an investigation on how this change affects flight performance. Such an investigation would determine whether or not the change in gaze behavior is accompanied by any negative side effects. This need for investigating this change is further highlighted by the difference in how experts and novices process and collect information, as described in Section 3.4.

This project, aimed at filling this research gap consists of two phases. The first is a preliminary phase in which research and much of the preparation is done. After the literature research, any tools required for the experiment are developed. This includes the development of a PC-7 VR Simulation and eye tracking tools within this simulation. Validation of these tools is then performed using a small experiment with commercial airline pilots. Besides validation this experiment also aids in providing some insight on an eye tracking experiment. It is essentially aiming to answer the following research question: ”What means are necessary to conduct an eye tracking experiment?”.

The second phase of this project can be defined as the experiment itself, along with the design of this experiment and analysis of results. The aim of this experiment is to fill the aforementioned research gap. The research question can then be said to be: ”What is the effect of changing gaze behavior on flight performance?”.

6.2. Apparatus

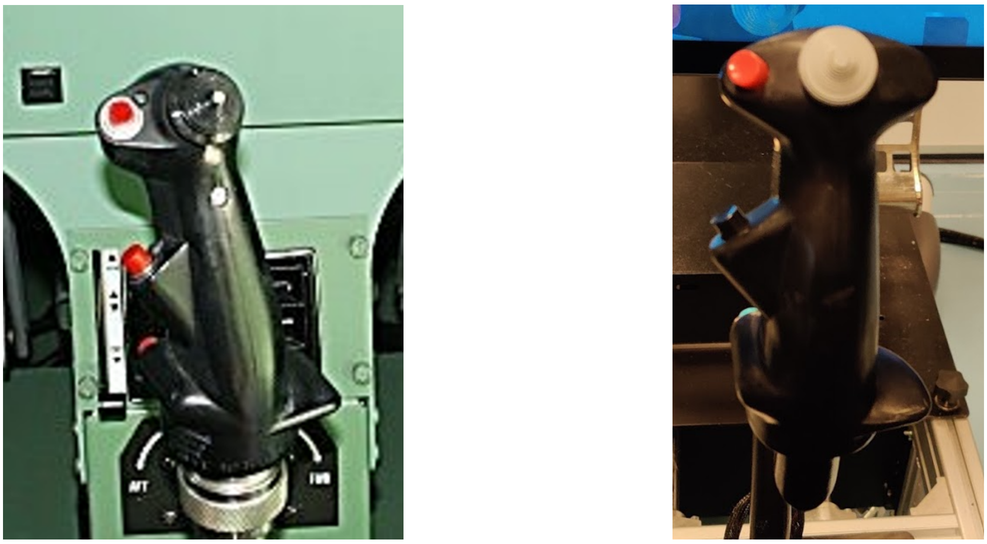
This section will detail the simulation and eye tracking apparatus used. The hardware for both was provided by MultiSim, and part of the eye tracking software has also been implemented within the DSIM simulation framework.

6.2.1. Simulation

The simulation platform, shown in Figure 6.1, consists of a chair with a touchscreen in front of it, a VR headset (the HP G2 Omnicept edition), a flight stick and throttle. The touchscreen can be used to interface with cockpit elements (e.g. gear lever, buttons on the bezel of the electronic flight instruments). The stick is near identical to the one present in the cockpit of the real PC-7 as demonstrated in Figure 6.2. In later iterations, the orientation of the stick will also be matched.

The PC-7 simulation has been developed as a part of this project in collaboration with MultiSIM B.V.. For this reason MultiSim’s simulation framework DSIM was used, together with the accompanying flight model. Several modules were developed, as is laid out in more detail in Appendix A.

Figure 6.1: Simulator setup



(a) PC-7 stick in the real aircraft

(b) PC-7 stick in the simulator

Figure 6.2: PC-7 stick comparison

6.2.2. Eye Tracking

Two VR Headsets with eye tracking capabilities were available for this project. The first of which being the HTC Vive Eye Pro, a VR headset with eye tracking capabilities, was used in early development for testing. Later stages of the experiment and preparation use the HP Reverb G2 Omnicept Edition instead, once it became available commercially.

Compared to the HTC, the HP VR headset has a higher visual resolution, larger field of view, and more accurate eye tracking. The visual resolutions are 1440x1600 and 2160x2160 per eye, respectively. As mentioned in Subsection 2.1.1 the field of view is of importance for scanning behavior, the fields of view are 110 and 114 degrees respectively. These improvements should help increase simulator fidelity and therefore data validity.

The respective manufacturers of the headsets claim an accuracy of 0.5-1.1 degrees and ‘sub degree’ accuracy, respectively. This is deemed to be enough for the purpose of this experiment, it is however important to verify that this claim is in fact true. For this reason each experiment will start and end with a small calibration test to see if the accuracy statement is correct and no drift has occurred during the experiment. This will be done through a DSIM default calibration procedure in which a set number of points will need to be looked at by the participant to complete the calibration.

In the future, a different high-end headset might be used, the Varjo XR-3. The manufacturer of this headset claims to gather more accurate eye tracking data. Also interesting for the Varjo XR-3 is the increase in resolution when the eyes are pointed straight ahead. This increase in resolution can help achieve more realistic scanning patterns during simulation.

6.3. Processing the Eye Tracking Output

Section 3.2 elaborates upon I-VT and I-AOI as identification algorithms for fixations and saccades. Both of these algorithms have been implemented. I-VT is done entirely during post-processing of the data in python, while I-AOI is done in part during the simulation.

Areas of Interest are defined in the 3D-Model of the PC-7 used in the simulator. This 3D model has been built by MultiSim in SketchUp and contains definitions for separate parts. As such, each knob, display or part of the window can be an area of interest if required. For the primary flight display for example, additional invisible collision meshes are placed over separate instruments in the display to detect these more specifically. This means there is an output for a specific instrument (e.g. ADI), but also an output for the background of the instrument panel. During simulation, collision between the gaze-ray and any defined areas of interest is to be tracked and logged. The total data available are thus the gaze direction (ψ, θ), the most recent area of interest hit by the gaze ray and whether an area of interest is currently looked at. The resulting sequence of AOIs looked at was then checked for the minimum duration, being 0.1 second, to end up with the I-AOI fixation identification.

During development, it was found that the eye tracking data produced by the HP G2 Omnicept Edition were too noisy. For this reason a filter was developed. This filter is a low pass filter, shown in Equation 6.1. With state k, filter parameter α and input u. To preserve saccade data the filter also resets if the error becomes too large. α is currently set to 0.13, and the threshold is set to 1.5. This was found to strike a balance between preserving saccades and responsiveness while reducing jitter. Additionally, a low pass filter reduces the effect of glissades on the data.

$$[H]k = (1 - \alpha) \cdot k + \alpha \cdot u \tag{6.1}$$

The raw, unfiltered output of the sensors is also saved. The raw output can be used to apply a different filter, or to remove a bias from the data if a bias is found to be present in post processing.

6.4. Implementation of Measures

This section details the implementation of the data processing. Data presented in this section are purely for development purposes. The aim is to gain insight in the relevant measures and test their implementation. These data were gathered by performing a right to center sweep, one or more instrument checks, and continuing the sweep from center to left. The actual sequence of events is shown in Figure 6.3. All plots in this section are generated with the data from this test. This pertains the first version of the software and the use of the HTC Vive Eye. Data were collected without flying by someone without any flying experience, the data were therefore not representative but merely for verification and validation purposes.

As explained, this section presents testing data. Further phases of this project will involve more data. Chapter 7 will present data gathered with airline pilots in order to validate the methodology presented in this chapter. Chapter 8 then presents the plans for gathering data in an experiment with the aim of answering the research questions.

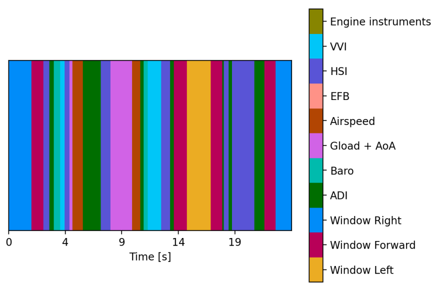
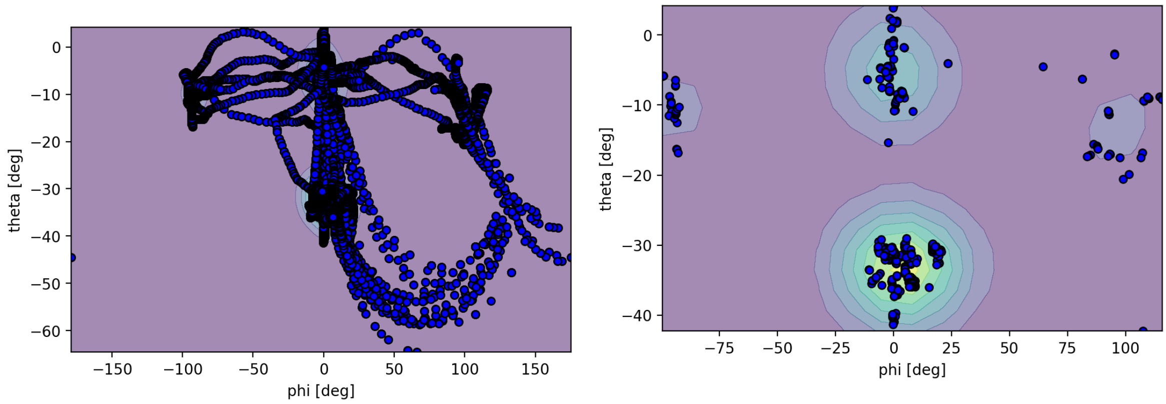


Figure 6.3: Scarf plot for the testing data

D-SIM outputs eye-tracking data from the left, right or combined eye, the latter of these are generated by the SDK of the headset. It has a more accurate depth figure, but it comes with a problem; Blinking makes this part of the API malfunction, resulting in data with artifacts as shown in Figure 6.4a. The artifacts are the arc motions in the bottom right (around 100 phi, -50 theta). These data will be filtered out, since using I-VT these artifacts will have a high point to point velocity. When using I-AOI a sweep such as this one would not reach the minimum fixation duration and thus would also not be valid data. Both I-VT and I-AOI have no issue filtering these artifacts as presented in Figure 6.4. Comparing the two figures demonstrates the use of the I-VT very well, a threshold of 20deg/s, as suggested in literature, results in Figure 6.4b. In the latter the scanning pattern used, being a sweep from the right side to the center, an instrument check, and a sweep from the center to the left side. Besides this being filtered very well by both identification algorithms, the severity of the issue was reduced when switching hardware to the HP Reverb G2.



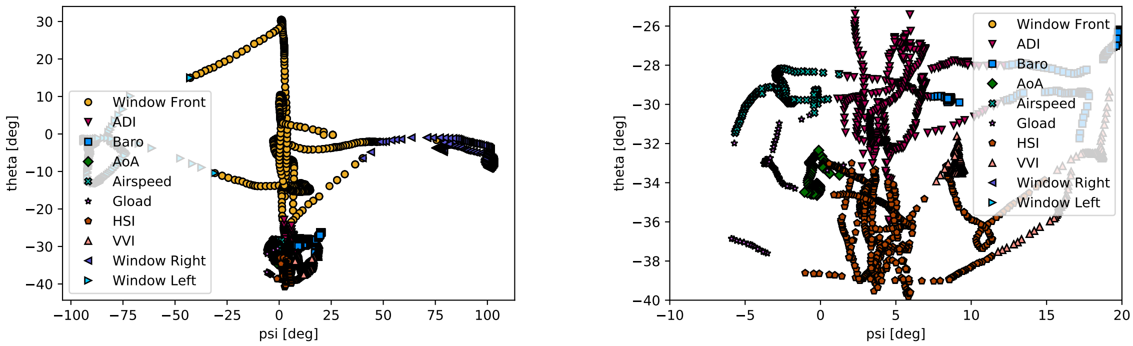
(a) Raw gaze data points and density heatmap

(b) Gaze data points and density heatmap after velocity threshold identification

Figure 6.4: Raw gaze data and fixations resulting from this data, both laid over a heatmap

6.4.1. AOI Fixations, Labeled

This subsection pertains to the analysis of data that have been gathered with AOI-identification enabled. Figure 6.5a shows the full scan path, with each marker showing a single timestep. The PFI is shown more closely in Figure 6.5b. These plots illustrate all data that are collected.

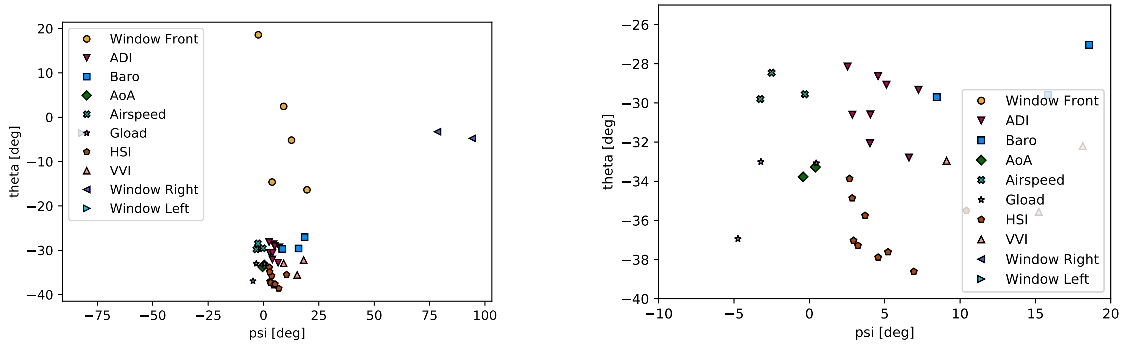


(a) Gaze data

(b) Gaze data of the PFI

Figure 6.5: All data points for the tracked gaze behaviour

Figure 6.6 shows the same two plots, but illustrates only the fixations found using I-AOI (duration threshold = 0.1s) instead. A limitation of the I-AOI algorithm becomes evident in Figure 6.6a since the sweeps over the windows are condensed into more or less a single heading. When looking closely at the plots one might notice that a position on the plot does not exactly correspond with a specific AOI. This is because the plot shows the angles output by the eye-tracker, but does not take into account the head position and distance to the AOI. Important to note is that there is no inaccuracy introduced here, the difference is merely visual. A future version will likely improve upon this, but it is not a priority. This would also allow overlaying the fixations on top of an image of the cockpit for visualization purposes.



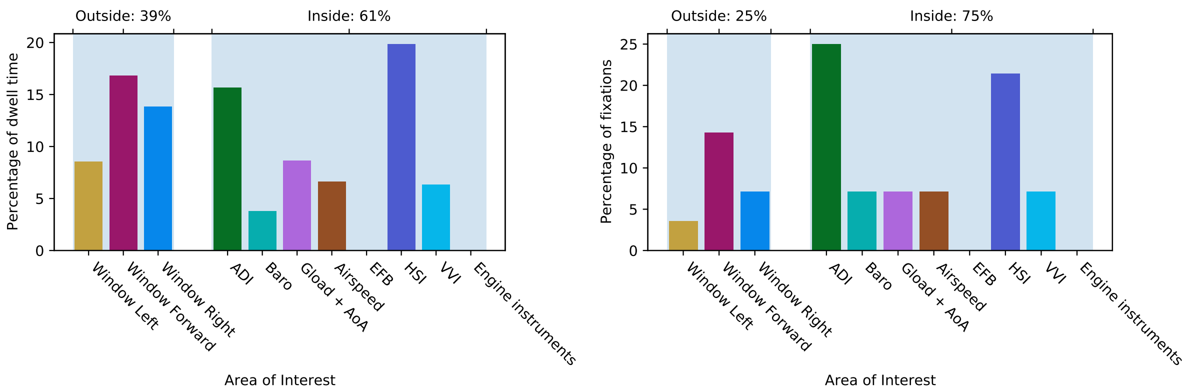
(a) Gaze data, fixations only

(b) Gaze data of the PFI, fixations only

Figure 6.6: AOI-T fixations for the tracked gaze behaviour

6.4.2. Dwell Time and Fixation Percentages

Figure 6.7 shows the percentages of dwell times and fixations per AOI for the test run data, this gives an immediate insight into what instruments are being relied on. Additionally these barplots show the division between time spent looking inside and outside, which provides a more general overview of the way in which scanning time was allocated by the pilot.



(a) Percentage of dwell time per AOI

(b) Percentage of fixations per AOI

Figure 6.7: Division of dwell times and fixations per AOI

6.4.3. GTE

The gaze transition entropy, shown in Equation 3.2, is considered to be of importance to the experiment. A purely deterministic string of AOIs should give a normalized GTE of 0, meaning no entropy, while a purely random string of AOIs should return a GTE of 1. For the purely random case two sets of inputs were tested. One with a constructed string having each possible transition occur once for five AOIs results in a GTE of 1, as is expected for a purely random sequence. The second test input for the pure random case is a randomly generated string of 10000 entries over five AOIs. This consistently leads to an error of $\epsilon < 0.001$. A purely deterministic sequence was also set as input, which lead to a value of zero for the GTE. This means the GTE can be said to be working as expected.

6.4.4. N-gram

The N-gram patterns or frequencies can give some insight into what repeating patterns exist within the data. Figure 6.8 shows the N-gram frequencies for the test data. The 2-gram occurring most frequently (4 times) is the transition from the HSI to the ADI. Evident here is the fact that there are no multiple occurring 5-grams. This makes sense due to the short time span of these test data.

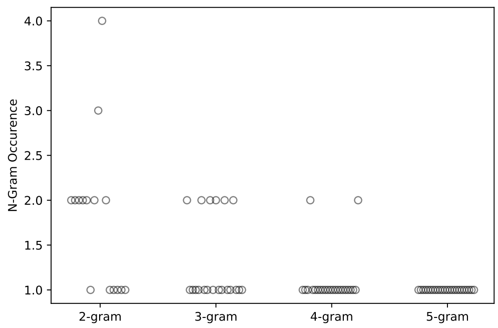
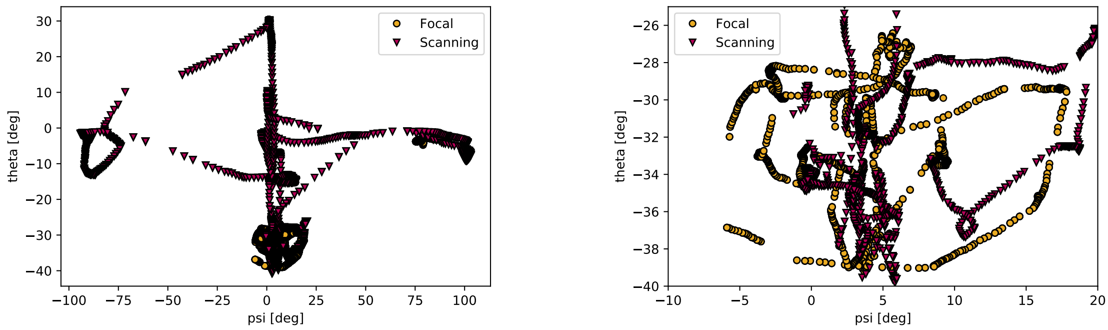


Figure 6.8: N-gram frequencies for the test data

6.4.5. Focal Coefficients

Lastly, the focal coefficients are examined. The plot of these focal coefficients is shown in Figure 6.9.

Evident from this plot is that focal scanning is employed while the instruments are being looked at. Looking out of the cockpit, on the other hand, results in scanning behavior. When looking outside, information is visually more spread out than when looking inside. This results in larger saccade amplitudes. The saccade amplitude is used in the calculation of the focal coefficients, this makes it unclear whether any value is added by looking at these focal coefficients, beyond discerning whether a participant is looking inside or outside.



(a) Focal coefficients for the entire tracking period

(b) Focal coefficients on the PFI

Figure 6.9: Focal coefficients during the test run

6.5. Conclusion

This chapter has shown the implementation of the eye tracking tools to be used in the experiment. As most metrics rely on the use of AOI, the chosen fixation identification algorithm is I-AOI with a minimum fixation duration of 0.1 seconds. The selection of I-AOI also means that glissades are considered to be part of fixations, so long as they remain within an AOI. Additionally, not I-AOI has the advantage over I-VT that more of the less relevant fixations (those not on instruments) are filtered out.

The focal coefficients are likely to be less useful. Their binary value seems to mostly indicate whether or not the gaze is currently positioned inside or outside. This is a side effect of the measured being designed for eye tracking in more static situations.

Among the particularly interesting measures are the dwell time percentages, which provide an overview of how the attention was divided over instruments during the span of the collected data. Combining this with the scarf plot, one can also see in what way this attention was divided to an instrument. It answers the question of in what way the attention is divided, showing either frequent short checks or fixations with an extended duration. The N-gram patterns provide some insight into what patterns are most relied on and are at the basis of the scan pattern. It is however not likely that scan patterns can be discerned simply by looking at the N-gram plot's shape. The GTE on the other hand shows promise in comparing between groups, showing the complexity in the patterns displayed in a sequence.

7 Preliminary Experiment

During the preparation for this thesis an opportunity arose to borrow a couple of participants from a study at TNO. This experiment allowed for verifying and validating the methodology described in Chapter 6 and the tools developed in preparation for the future experiment.

Section 7.1 discusses the participants used in this preliminary experiment. Next, Section 7.2 elaborates briefly on the methodology of this experiment. The results are then presented in Section 7.3 and discussed in Section 7.4. Lastly, a brief conclusion is then shown in Section 7.5.

7.1. Participants

The participating pilots were all civil air transport pilots. None of these pilots had any experience with flying aerobatics, and all had little to no experience flying small aircraft. These pilots had considerable experience with commercial aircraft, as they were pre-selected for the unrelated TNO experiment.

7.2. Methodology

The experiment concerned flying a basic circuit (left handed) twice. A basic circuit starts with take-off, and ends with landing on the same runway. The part of this circuit in which pilots are flying back to the start of the runway, with the aircraft parallel to the runway, is called the downwind. A basic circuit can also be seen in Figure 7.1. Pilots were instructed to fly at 110 kts for the circuit, and were provided with an appropriate power setting for take-off and cruise. The circuit was to be flown at 1,000 ft, and a constant distance from the runway. The pilots were instructed to position the tip of the wing on the runway to assist in keeping the distance to the runway constant.

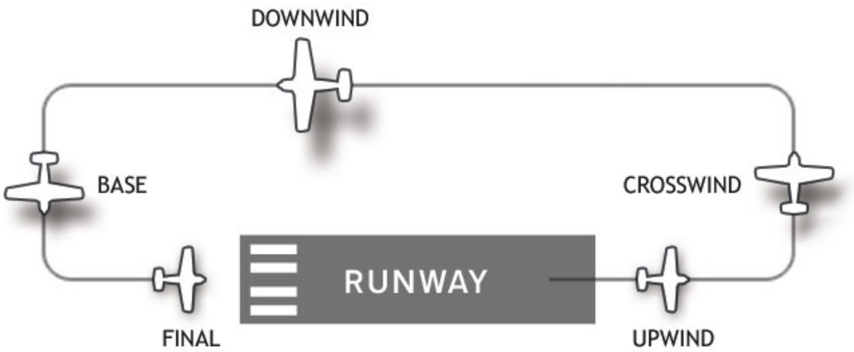


Figure 7.1: Basic circuit with the name of each stage indicated [source:Wikimedia commons]

Additionally the briefing contained some details on the runway and aircraft behavior to help the pilots fly a more controlled circuit without a long training time. Due to time constraints this circuit was flown twice, with the first being for training. A short training period can in part be justified with the experience of each pilot. Appendix B contains further details on the exact briefing and the procedure as followed by the researcher.

Of specific interest to this experiment was the downwind leg of the second circuit, as at this point the participants are in a somewhat steady flight condition, and have been able to get used to the PC-7 simulation's flying behavior in the first circuit. The results of the second circuit are presented in the next section.

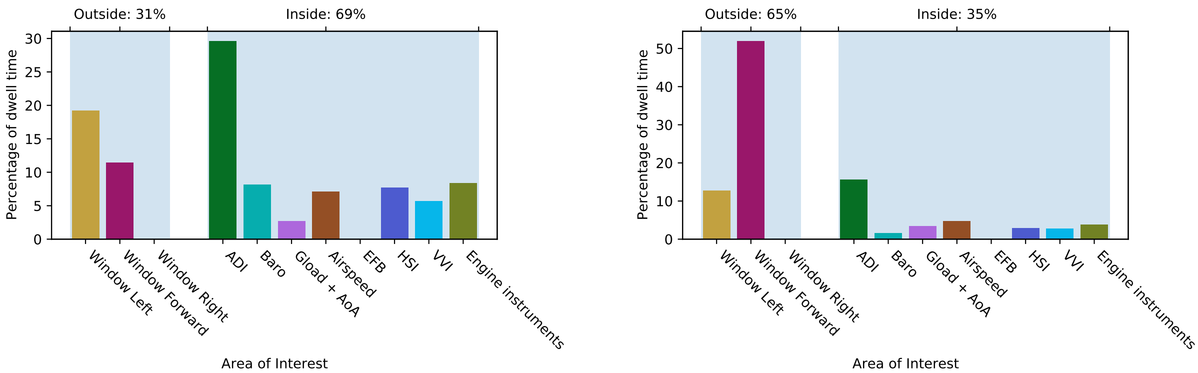
During the first day of experiments the briefing was less clear. More specifically, not all pilots were familiar with a 'standard circuit' leading to some confusion during flying. Data were discarded and the briefing was adjusted. The remaining participants were considered as a separate set of data (n=11) after this and their data are the data that are presented here.

7.3. Results

This section presents the findings from this experiment. The gaze behavior is processed with the tools described Chapter 6. Subsection 7.3.1 assesses the data in terms of dwell time percentages. The overall character of the scan pattern is investigated using the GTE in Subsection 7.3.2 and the N-grams in Subsection 7.3.3.

7.3.1. Dwell times

Figure 7.2 shows the percentage of dwell time for each of the tracked areas of interest during the experiment for two different flight phases, the downwind and the final parts of the circuit. Immediately, it becomes evident that the pilots look inside for the majority of the time (69%) during the downwind leg. This was consistent with the expectations, considering the participants were civil air transport pilots, who flew mostly instrument flight rules (IFR). Interestingly, during approach this is the opposite. Instead, the pilots look at the threshold and occasionally cross check the Attitude Direction Indicator.



(a) Dwell time percentages during the downwind

(b) Dwell time percentages during the final leg

Figure 7.2: Percentage of time spent looking at a specific instrument during a flight phase in the second circuit flown

7.3.2. GTE

For this group of participants, on the downwind, the average gaze transition entropy was found to be 0.478, with $\sigma = 0.0816$. This gives some insight in how the GTE responds to a relatively consistent group of participants over a calm flight phase. Sadly a comparison with literature is difficult since this concerns the normalized gaze transition entropy, while other research seems to use the non-normalized version (e.g. Lounis et al. (2020); Diaz-Piedra et al. (2019)). Furthermore, the GTE is likely sensitive to AOI definition, due to its reliance on the transitions between AOIs. AOI definition can also wildly vary between experiments, with Diaz-Piedra et al. (2019) having divided the field of view in bins, each being 1 degree of visual angle high and wide.

7.3.3. N-gram Patterns and Frequencies

The N-gram patterns can be used to see what common scan patterns exist. Figure 7.3 shows the occurrence of the N-grams in the downwind flight phase.

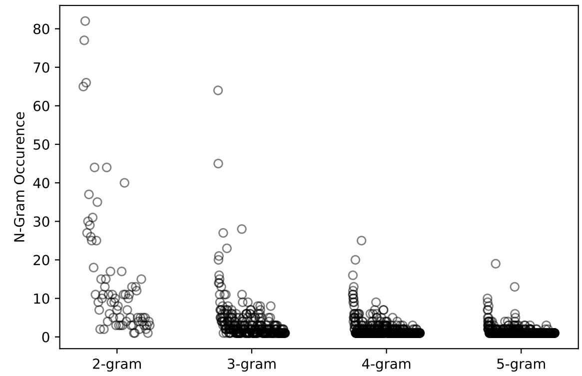


Figure 7.3: Distribution of N-gram frequencies for the downwind

To avoid any bias due to varying downwind length, the results were normalized per participant and shown in Figure 7.4. No significant changes to the order occur when normalizing. For future experiments it would be beneficial to consider a fixed length to avoid the need for normalization.

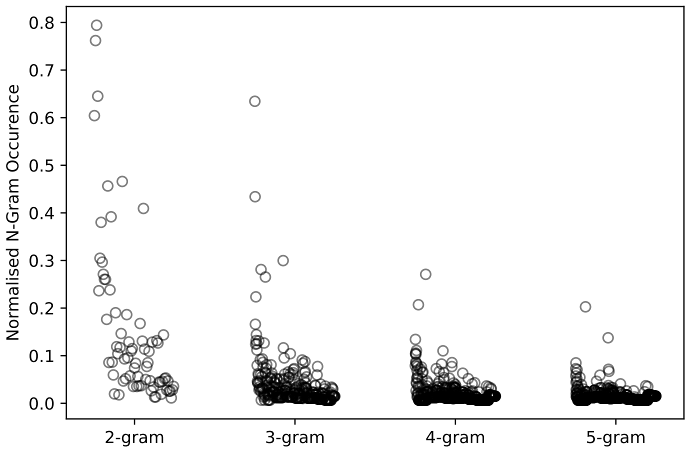


Figure 7.4: Normalized distribution of N-gram frequencies for the downwind

The four most common (normalized) N-gram patterns for the downwind include those shown in Table 7.1 (highest occurrence first), ASI being the Air Speed Indicator. Looking at this table, the reliance on the ADI is clear once more through the apparent scanning pattern of checking one instrument and returning to the ADI.

Table 7.1: Four most common N-gram patterns for 2-5-grams

2-gram	3-gram	4-gram	5-gram
ADI, ASI	ADI, ASI, ADI	ADI, ASI, ADI, ASI	ADI, ASI, ADI, ASI, ADI
ASI, ADI	ADI, Baro, ADI	ADI, ASI, ADI, ASI	ASI, ADI, ASI, ADI, ASI
Baro, ADI	Window (Forward, Left, Forward)	Baro, ADI, ASI, ADI	ADI, ASI, ADI, Baro, ADI
ADI, Baro	ASI, ADI, ASI	ADI, Baro, ADI, Baro	ADI, Baro, ADI, ASI, ADI

7.4. Discussion

When looking at the results or at the replay of the experiment and during debrief, it becomes evident that both VR resolution and eye tracking accuracy deteriorate when deviating from the center of the field of view. Objects and instruments are sharper when at the center of the field of view. Participants were more likely to point their head at the instrument while they might normally just move their eyes. Quantifying such an effect is difficult if not impossible.

One thing that is of note is the value of the G-load + Angle of Attack indicator in Figure 7.2a. The pilots each indicated not having any experience with the instrument and therefore, also not using them. This means that the presence of the G-load+AoA label in this plot is a good indication of any inaccuracies present in the data. These inaccuracies might be due to the decrease in eye-tracking accuracy away from the center of the field of view. If a participant looked down moving just their eyes, the resulting eye tracking output was less accurate than it was if a participant looked down by moving their head.

One way to solve these issues might be to use the Varjo XR-3. This headset's specifications are more advanced than the HP G2 Omnicept that was used for this experiment. Most notable for this are an increased field of view, more accurate eye tracking, increased resolution and the inclusion of a 'focus area' that has an even higher resolution than the peripheral. The more accurate eye tracking being the most important feature here.

Additionally, a solution is being worked on that corrects the data after the fact using DSIM's replay functionality. Such a system could involve buffer AOIs positioned outside of the instrument panel to identify inaccuracies such as biases.

7.5. Conclusion

In conclusion, this preliminary experiment shows that the tools can work in their current state, while still requiring some improvements.

The issues include the limited horizontal field of view of the headset, which is cannot be solved with the currently available technology. The limited field of view leads to more head movements relative to real flight. For example, if the participant wants to look at a point that is too far too the right, they will have to turn their head. Such a movement is much slower than an eye movement, leading to a different overall scanning pattern.

Other issues that can be attributed to the VR headset are the accuracy of the eye tracking. The reduced accuracy has the consequence of possibly needing to reduce the number of AOIs used in the experiment. Fixations on small AOIs are more likely to be interrupted by the noise and bias on these instruments. This is of most concern to the measures relating to the AOIs, as indicated per measure in Table 3.1. As the majority of measures shown here, it is likely necessary to consider larger AOIs to alleviate this problem. One such set of larger AOIs could be as follows: three directions outside (forward, left, right) and the instruments. The eye tracking data quality can perhaps also be improved by better post-processing capabilities and different filter settings.

Another conclusion drawn from this experiment is that the task definition is very important in behavior. This can also be seen when comparing the data from the set with the less optimal briefing.

8 Thesis Experiment Plan

This chapter details the future plans for this project. Section 8.1 will elaborate upon the hypothesis. Section 8.2 details who will participate in this experiment and Section 8.3 explains the methodology.

8.1. Hypothesis

The research gap established in Section 6.1 concerned the effect of changing gaze behavior on flight performance. Chapter 4 shows different ways in gaze behavior could possible be affected. Such methods have been shown to lead to an increase in task performance. One or more of these methods will be used in the experiment to establish the effect on flight performance, in which the hypotheses are as follows:

Exposing novice pilots to expert scanning behavior leads to improved flight performance.

Flight performance here consisting of consistently flying at a given an altitude, speed and heading.

Additionally, the effect on scanning behavior is considered: *Exposing novice pilots to expert scanning behavior leads to faster detection of other flying objects.*

The latter hypothesis serves to illustrate that exposing novice pilots to expert scanning behavior leads to better awareness of the airspace around them.

8.2. Participants

Since the hypothesis concerns both novices and experts, an experiment would require two sets of participants to be able to compare the two. For the experts a minimum of 4 expert military pilots will be available for this experiment, among which are two PC-7 instructors.

Ideally, the second set of participants would be a group with some experience. This requirement is due to realistic scenarios being too hard for a group that has no prior experience. Vygotsky (1978) also establishes these principles with the zones of proximal development. The zones of proximal development can be used to show that a learner cannot immediately perform complex tasks without further training. A suitable group might be those who have recently completed the PC-7 flight training, it is however not yet clear if such a group can be found for this experiment.

Additionally, the use of TU Delft students with no flight experience is considered. This allows for some insight into the untrained behavior. The lack of experience also make it easier to learn something new to the participants, as there is no old behavior to replace. When using participants with no flight experience it is important to take this into account during the experiment design, due to the learning curve. The experiment to confirm the hypotheses is to be between subjects, with the method of the participants' exposure to scanning behavior being the manipulated variable across the groups.

8.3. Methodology

The hardware and software used will likely be the same as in the preliminary experiment described in Chapter 7. Although, some further iterations on both hardware and software may be done in order to resolve some of the issues discussed in Section 7.4 and Section 7.5. The provided simulation software is able to log all variables present in the simulation (e.g. aircraft state, pilot head orientation, gaze direction) for the full experiment.

As this experiment involves trying to change participants' gaze behavior, it is important to consider how this will be done. Chapter 4 shows that creating a spotlight of the gaze behavior is a suitable method for demonstrating what an expert would do in a given situation. The spotlight method is also planned to be used in the experiment.

For this experiment one has to consider the expertise of the participants. Choosing a task that goes too far beyond a novice pilot's skill set (outer layer of the zones of proximal development) would lead to unusable results. As such it is important to consider a task that fits the participant groups well. A simple, yet representative task has to found for this experiment. An example of a task that can be considered too difficult is requiring a stable approach and landing. An alternative is using a stretch of straight and level flight, such as the downwind leg also used in Section 7.2.

The hypothesis will be evaluated by looking at whether the measures mentioned have any statistically significant differences before and after being exposed to expert behavior. Additionally, this relationship will be investigated for different approaches to showing this expert behavior to the participants. The comparison will also be made with participants that have no such exposure.

9 Conclusion

This report has investigated the added value of eye tracking in flight training. It was found that now is a good time for introducing eye tracking within VR simulations due to the increased accessibility of this technology. In addition, many military training organizations, including the the Dutch Air Force's initial flight training (EMVO) are introducing VR into their curricula. These conditions make this a good time to also investigate eye tracking in VR.

Literature has shown that eye tracking and the investigation of scanning patterns have a lot of potential to add value to training programs. Furthermore, a lot of measures that show potential in characterizing scanning behavior have been evaluated, and a few of those have been selected for use in an experiment. One such measure is the Gaze Transition Entropy (GTE) which quantifies the randomness of a scan pattern. Additionally, some important considerations came forward from the literature. It is very important to be aware of the skill level of the participants, since this will affect the tasks they are able to perform, and by extent the complexity of the scanning pattern. For a training curriculum, such as the EMVO's, this would mean that it is not desirable to immediately teach a complex, expert pattern to new students.

As a part of this first project phase, a PC-7 simulation was developed. Additionally eye tracking tools were developed. These eye tracking tools use raw data from the HP G2 Omnicept edition's eye tracker to establish at what instrument one is looking. This data is logged and processed to generate several measures quantifying the scanning behavior. The most straightforward of which is the percentage of dwell time per instrument. These measures can then be used to characterize gaze behavior, which is the first step in using it to enhance training. Upon initial investigation, these measures have shown to work quite well in giving insight into the pilot's behavior.

The eye tracking tools were also verified and used in a preliminary experiment. This preliminary experiment revealed what additional development still needs to be done to ensure usable data. Most important is the need to be able to re-identify what the participants were looking at with a bias or a different filter applied. This should help improve the quality of the data gathered. The experiment also revealed the ease with which these tools can be used. Another experiment will be done with novices to try to affect their behavior using expert gaze behavior. The hypothesis states that this will lead to both improved awareness of the airspace around them, and improved flight performance.

References

Allerton, D. J. (2010). The impact of flight simulation in aerospace [Journal Article]. *Aeronautical Journal*, *114*(1162), 747-756. doi: 10.1017/S0001924000004231

Allsop, J., & Gray, R. (2014). Flying under pressure: Effects of anxiety on attention and gaze behavior in aviation [Journal Article]. *Journal of Applied Research in Memory and Cognition*, *3*(2), 63-71. Retrieved from `https://www.sciencedirect.com/science/article/pii/S2211368114000333` doi: https://doi.org/10.1016/j.jarmac.2014.04.010

Brams, S., Hooge, I. T. C., Ziv, G., Dauwe, S., Evens, K., De Wolf, T., … Helsen, W. F. (2018). Does effective gaze behavior lead to enhanced performance in a complex error-detection cockpit task?

[Journal Article]. *PLOS ONE*, *13*(11), e0207439. Retrieved from https://doi.org/10.1371/journal.pone.0207439 doi: 10.1371/journal.pone.0207439

Carbonell, J. R. (1966). A queueing model of many-instrument visual sampling [Journal Article]. *IEEE Transactions on Human Factors in Electronics*, *HFE-7*(4), 157-164. doi: 10.1109/THFE.1966.232984

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., & Thompson, J. D. (2003). Multiple sequence alignment with the clustal series of programs [Journal Article]. *Nucleic acids research*, *31*(13), 3497-3500.

Chetwood, A. S. A., Kwok, K.-W., Sun, L.-W., Mylonas, G. P., Clark, J., Darzi, A., & Yang, G.-Z. (2012). Collaborative eye tracking: a potential training tool in laparoscopic surgery [Journal Article]. *Surgical Endoscopy*, *26*(7), 2003-2009. Retrieved from https://doi.org/10.1007/s00464-011-2143-x doi: 10.1007/s00464-011-2143-x

Dehais, F., Juaneda, S., & Peysakhovich, V. (2020). Monitoring eye movements in real flight conditions for flight training purpose [Conference Paper]. In V. Peysakhovich, D. Rudi, M. Causse, P. Kiefer, F. Dehais, & M. Raubal (Eds.), *1st international workshop on eye-tracking in aviation (etavi 2020)* (p. 60). ISAE-SUPAERO, Université de Toulouse; Institute of Cartography and Geoinformation (IKG), ETH Zurich. Retrieved from http://hdl.handle.net/20.500.11850/407652 doi: 10.3929/ethz-b-000407652

Diaz-Piedra, C., Rieiro, H., Cherino, A., Fuentes, L. J., Catena, A., & Di Stasi, L. L. (2019). The effects of flight complexity on gaze entropy: An experimental study with fighter pilots [Journal Article]. *Applied Ergonomics*, *77*, 92-99. Retrieved from https://www.sciencedirect.com/science/article/pii/S0003687019300274 doi: https://doi.org/10.1016/j.apergo.2019.01.012

Ditchburn, R. W. (1980). The function of small saccades [Journal Article]. *Vision Research*, *20*(3), 271-272. doi: 10.1016/0042-6989(80)90112-1

Drela, M. (2004). *Avl overview.* Retrieved from https://web.mit.edu/drela/Public/web/avl/ (Accessed on: 22/02/2021)

Dubois, E., Blättler, C., Camachon, C., & Hurter, C. (2015). Eye movements data processing for ab initio military pilot training [Conference Proceedings]. In *Kes international conference on intelligent decision technologies* (Vol. 39). doi: 10.1007/978-3-319-19857-6_12

Ekroot, L., & Cover, T. M. (1993). The entropy of markov trajectories [Journal Article]. *IEEE Transactions on Information Theory*, *39*(4), 1418-1421. doi: 10.1109/18.243461

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems [Journal Article]. *Human Factors*, *37*(1), 32-64. Retrieved from https://doi.org/10.1518/001872095779049543 doi: 10.1518/001872095779049543

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing [Journal Article]. *Vision Research*, *42*(5), 621-636. Retrieved from https://www.sciencedirect.com/science/article/pii/S0042698901003017 doi: https://doi.org/10.1016/S0042-6989(01)00301-7

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory [Journal Article]. *Psychological Review*, *102*(2), 211-245. doi: 10.1037/0033-295X.102.2.211

Faulhaber, A. K., & Friedrich, M. (2019). Eye-tracking metrics as an indicator of workload in commercial single-pilot operations [Book Section]. In (p. 213-225). Springer International Publishing. Retrieved from https://dx.doi.org/10.1007/978-3-030-32423-0_14 doi: 10.1007/978-3-030-32423-0_14

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches [Journal Article]. *Aeronautical Engineering Review*, *9*, 24-29.

Fox, J., Merwin, D., Marsh, R., McConkie, G., & Kramer, A. (1996). Information extraction during instrument flight: An evaluation of the validity of the eye-mind hypothesis [Journal Article]. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *40*(2), 77-81. Retrieved from https://journals.sagepub.com/doi/abs/10.1177/154193129604000215 doi: 10.1177/154193129604000215

Friedrich, M., Lee, S. Y., Bates, P., Martin, W., & Faulhaber, A. K. (2021). The influence of training level on manual flight in connection to performance, scan pattern, and task load [Journal Article]. *Cognition, Technology and Work*. Retrieved from https://dx.doi.org/10.1007/s10111-020-00663-8 doi: 10.1007/s10111-020-00663-8

Fuchs, A. F. (1971). The saccade system [Journal Article]. *The control of eye movements*, 343-362. Retrieved from https://ci.nii.ac.jp/naid/10008713140/en/

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains [Journal Article]. *Educational psychology review*, *23*(4), 523-552.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs [Journal Article]. *International Journal of Industrial Ergonomics*, *24*(6), 631-645. Retrieved from https://www.sciencedirect.com/science/article/pii/S0169814198000687 doi: https://doi.org/10.1016/S0169-8141(98)00068-7

Haider, H., & Frensch, P. A. (1996). The role of information reduction in skill acquisition [Journal Article]. *Cognitive Psychology*, *30*(3), 304-337. Retrieved from https://www.sciencedirect.com/science/article/pii/S0010028596900092 doi: https://doi.org/10.1006/cogp.1996.0009

Hecht, S. (1937). Rods, cones, and the chemical basis of vision [Journal Article]. *Physiological Reviews*, *17*(2), 239-290.

Heiligers, M. M., Van Holten, T., & Mulder, M. (2009). Predicting pilot task demand load during final approach [Journal Article]. *The International Journal of Aviation Psychology*, *19*(4), 391-416. Retrieved from https://doi.org/10.1080/10508410902983987 doi: 10.1080/10508410902983987

Ho, H.-F., Su, H.-S., Li, W.-C., Yu, C.-S., & Braithwaite, G. (2016). Pilots' latency of first fixation and dwell among regions of interest on the flight deck [Conference Proceedings]. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (p. 389-396). Springer International Publishing.

Holmqvist, K., & Nyström, M. (2011). *Eye tracking: A comprehensive guide to methods and measures.* [Book]. Oxford: Oxford University Press.

Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies [Journal Article]. *Journal of Experimental Psychology: Applied*, *12*(2), 67.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention [Journal Article]. *Vision Research*, *40*(10), 1489-1506. Retrieved from https://www.sciencedirect.com/science/article/pii/S0042698999001637 doi: https://doi.org/10.1016/S0042-6989(99)00163-7

Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., & Eika, B. (2012). Conveying clinical reasoning based on visual observation via eye-movement modelling examples [Journal Article]. *Instructional Science*, *40*(5), 813-827. Retrieved from https://doi.org/10.1007/s11251-012-9218-5 doi: 10.1007/s11251-012-9218-5

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning [Journal Article]. *Learning and Instruction*, *25*, 62-70. Retrieved from https://www.sciencedirect.com/science/article/pii/S0959475212000990 doi: https://doi.org/10.1016/j.learninstruc.2012.11.004

Javal, E. (1878). Essai sur la physiologie de la lecture [Journal Article]. *Annales d'Ocilistique*, *80*, 97-117.

Johnson, N., Wiegmann, D., & Wickens, C. (2006). Effects of advanced cockpit displays on general aviation pilots' decisions to continue visual flight rules flight into instrument meteorological conditions [Journal Article]. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(1), 30-34. Retrieved from https://doi.org/10.1177/154193120605000107 doi: 10.1177/154193120605000107

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes [Journal Article]. *Cognitive Psychology*, *8*(4), 441-480. Retrieved from https://www.sciencedirect.com/science/article/pii/0010028576900153 doi: https://doi.org/10.1016/0010-0285(76)90015-3

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension [Journal Article]. *Psychological Review*, *87*(4), 329-354. doi: 10.1037/0033-295X.87.4.329

Kang, Z., & Landry, S. J. (2014). Using scanpaths as a learning method for a conflict detection task of multiple target tracking [Journal Article]. *Human Factors*, *56*(6), 1150-1162. Retrieved from https://doi.org/10.1177/0018720814523066 doi: 10.1177/0018720814523066

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during vfr flight [Conference Proceedings]. In *Proceedings of the 11th international symposium on aviation psychology* (Vol. 6, p. 325-335). Citeseer.

Kolb, H. (1995). *Gross anatomy of the eye* [Book]. University of Utah Health Sciences Center, Salt Lake City (UT). Retrieved from http://europepmc.org/abstract/MED/21413392http://europepmc.org/books/NBK11534https://www.ncbi.nlm.nih.gov/books/NBK11534

Kolb, H., Nelson, R. F., Ahnelt, P. K., Ortuño-Lizarán, I., & Cuenca, N. (2020). The architecture of the human fovea [Journal Article]. *Webvision: The Organization of the Retina and Visual System [Internet]*.

Krejtz, K., Duchowski, A., Krejtz, I., Szarkowska, A., & Kopacz, A. (2016). Discerning ambient/focal attention with coefficient k [Journal Article]. *ACM Trans. Appl. Percept.*, *13*(3), Article 11. Retrieved from https://doi.org/10.1145/2896452 doi: 10.1145/2896452

Krejtz, K., Duchowski, A., Szmidt, T., Krejtz, I., Perilli, F. G., Pires, A., … Villalobos, N. (2015). Gaze transition entropy [Journal Article]. *ACM Trans. Appl. Percept.*, *13*(1), Article 4. Retrieved from https://doi.org/10.1145/2834121 doi: 10.1145/2834121

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: Gaze-tracking study [Journal Article]. *Radiology*, *242*(2), 396-402. Retrieved from https://pubs.rsna.org/doi/abs/10.1148/radiol.2422051997 doi: 10.1148/radiol.2422051997

Lamb, T. D. (2016). Why rods and cones? [Journal Article]. *Eye*, *30*(2), 179-185. Retrieved from https://doi.org/10.1038/eye.2015.236 doi: 10.1038/eye.2015.236

Law, B., Atkins, M. S., Kirkpatrick, A. E., & Lomax, A. J. (2004). Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment [Conference Paper]. In *Proceedings of the 2004 symposium on eye tracking research and applications* (p. 41–48). Association for Computing Machinery. Retrieved from https://doi.org/10.1145/968363.968370 doi: 10.1145/968363.968370

Lefrancois, O., Matton, N., Gourinat, Y., Peysakhovich, V., & Causse, M. (2016). The role of pilots' monitoring strategies in flight performance [Conference Paper]. In *European association for aviation psychology conference eaap32* (p. 1-11). Retrieved from https://oatao.univ-toulouse.fr/16173/

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals [Conference Proceedings]. In *Soviet physics doklady* (Vol. 10, p. 707-710). Soviet Union.

Lounis, C., Peysakhovich, V., & Causse, M. (2020). Lempel-ziv complexity of dwell sequences: visual scanning pattern differences between novice and expert aircraft pilots [Conference Proceedings]. In *1st international workshop on eye-tracking in aviation.*

Lounis, C., Peysakhovich, V., & Causse, M. (2021). Visual scanning strategies in the cockpit are modulated by pilots' expertise: A flight simulator study [Journal Article]. *PLOS ONE*, *16*(2), e0247061. Retrieved from https://dx.doi.org/10.1371/journal.pone.0247061 doi: 10.1371/journal.pone.0247061

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins [Journal Article]. *Journal of molecular biology*, *48*(3), 443-453.

Niehorster, D. C., Hildebrandt, M., Smoker, A., Jarodzka, H., & Dahlström, N. (2020). Towards eye tracking as a support tool for pilot training and assessment [Conference Paper]. In V. Peysakhovich, D. Rudi, M. Causse, P. Kiefer, F. Dehais, & M. Raubal (Eds.), *1st international workshop on eye-tracking in aviation (etavi 2020)* (p. 28). ISAE-SUPAERO, Université de Toulouse; Institute of Cartography and Geoinformation (IKG), ETH Zurich. Retrieved from http://hdl.handle.net/20.500.11850/407625 doi: 10.3929/ethz-b-000407625

NTSB. (1994). A review of flightcrew-involved major accidents of u.s. carriers, 1978 through 1990. document ntsb/sss-94-01. [Journal Article]. *National Transportation Saftey Board*.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data [Journal Article]. *Behavior Research Methods*, *42*(1), 188-204. Retrieved from https://doi.org/10.3758/BRM.42.1.188 doi: 10.3758/BRM.42.1.188

Ognjanovic, S., Thüring, M., Murphy, R. O., & Hölscher, C. (2019). Display clutter and its effects on visual attention distribution and financial risk judgment [Journal Article]. *Applied Ergonomics*, *80*, 168-174. Retrieved from https://www.sciencedirect.com/science/article/pii/

S0003687019300924 doi: https://doi.org/10.1016/j.apergo.2019.05.008

Peißl, S., Wickens, C. D., & Baruah, R. (2018). Eye-tracking measures in aviation: A selective literature review [Journal Article]. *The International Journal of Aerospace Psychology*, *28*(3-4), 98-112. Retrieved from https://doi.org/10.1080/24721840.2018.1514978 doi: 10.1080/24721840.2018.1514978

Ramón Alamán, J., Causse, M., & Peysakhovich, V. (2020). Attentional span of aircraft pilots: did you look at the speed? [Conference Paper]. In V. Peysakhovich, D. Rudi, M. Causse, P. Kiefer, F. Dehais, & M. Raubal (Eds.), *1st international workshop on eye-tracking in aviation (etavi 2020)* (p. 113). ISAE-SUPAERO, Université de Toulouse; Institute of Cartography and Geoinformation (IKG), ETH Zurich. Retrieved from http://hdl.handle.net/20.500.11850/407659 doi: 10.3929/ethz-b-000407659

Reani, M., Peek, N., & Jay, C. (2018). An investigation of the effects of n-gram length in scanpath analysis for eye-tracking research [Conference Paper]. In *Proceedings of the 2018 acm symposium on eye tracking research and applications* (p. Article 1). Association for Computing Machinery. Retrieved from https://doi.org/10.1145/3204493.3204527 doi: 10.1145/3204493.3204527

Reingold, E. M., Charness, N., Pomplun, M., & Stampe, D. M. (2001). Visual span in expert chess players: Evidence from eye movements [Journal Article]. *Psychological Science*, *12*(1), 48-55. Retrieved from https://doi.org/10.1111/1467-9280.00309 doi: 10.1111/1467-9280.00309

Roskam, J. (1985). *Airplane design part v: Component weight estimation* (3rd ed.) [Book]. Lawrence, Kansas, USA: DARcorporation.

Sadraey, M. H. (2012). *Aircraft design: A systems engineering approach* [Book]. John Wiley & Sons.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols [Conference Paper]. In *Proceedings of the 2000 symposium on eye tracking research and applications* (p. 71–78). Association for Computing Machinery. Retrieved from https://doi.org/10.1145/355017.355028 doi: 10.1145/355017.355028

Sen, T., & Megaw, T. (1984). The effects of task variables and prolonged performance on saccadic eye movement parameters [Journal Article]. *Advances in Psychology*, *22*, 103-111.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems [Journal Article]. *IEEE Transactions on Human Factors in Electronics*, *HFE-5*(1), 2-5. doi: 10.1109/THFE.1964.231647

Shannon, C. E. (1948). A mathematical theory of communication [Journal Article]. *The Bell System Technical Journal*, *27*(3), 379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review [Journal Article]. *Journal of Vision*, *11*(5), 13-13. Retrieved from https://doi.org/10.1167/11.5.13 doi: 10.1167/11.5.13

To, M. P. S., Regan, B. C., Wood, D., & Mollon, J. D. (2011). Vision out of the corner of the eye [Journal Article]. *Vision Research*, *51*(1), 203-214. Retrieved from https://www.sciencedirect.com/science/article/pii/S0042698910005614 doi: https://doi.org/10.1016/j.visres.2010.11.008

Van Dijk, H., Van de Merwe, K., & Zon, R. (2011). A coherent impression of the pilots' situation awareness: Studying relevant human factors tools [Journal Article]. *The International Journal of Aviation Psychology*, *21*(4), 343-356. Retrieved from https://doi.org/10.1080/10508414.2011.606747 doi: 10.1080/10508414.2011.606747

van Gog, T., & Jarodzka, H. (2013). Eye tracking as a tool to study and enhance cognitive and metacognitive processes in computer-based learning environments [Book Section]. In *International handbook of metacognition and learning technologies* (p. 143-156). New York, NY: Springer New York. Retrieved from https://doi.org/10.1007/978-1-4419-5546-3_10 doi: 10.1007/978-1-4419-5546-3_10

Van Gog, T., Paas, F., Van Merriënboer, J. J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting [Journal Article]. *Journal of Experimental Psychology: Applied*, *11*(4), 237.

Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving [Journal Article]. *Pragmatics & Cognition*, *3*(2), 199-223.

Velichkovsky, B. M., Helmert, J. R., Joos, M., & Pannasch, S. (2005). Two visual systems and their eye movements: Evidence from static and dynamic scene perception [Conference Proceedings]. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 27).

Vine, S. J., Masters, R. S. W., McGrath, J. S., Bright, E., & Wilson, M. R. (2012). Cheating experience: Guiding novices to adopt the gaze strategies of experts expedites the learning of technical laparoscopic skills [Journal Article]. *Surgery*, *152*(1), 32-40. Retrieved from `https://www.sciencedirect.com/science/article/pii/S003960601200044X` doi: https://doi.org/10.1016/j.surg.2012.02.002

Vlasblom, J., & Rooij, J. v. (2020). Augmented eye: From theory to practice [Conference Paper]. In *AIAA Scitech 2020 forum.* Retrieved from `https://arc.AIAA.org/doi/abs/10.2514/6.2020-0169` doi: 10.2514/6.2020-0169

Vlasblom, J., Zon, R., & van der Pal, J. (2020). The state and promise of eye tracking: from laboratory to operation [Conference Paper]. In *Eth.* doi: 10.3929/ethz-b-000407656

Vygotsky, L. (1978). Mind in society [Journal Article]. *Mind in society the development of higher psychological processes. Cambridge, MA: Harvard University Press*.

Weber, R. B., & Daroff, R. B. (1972). Corrective movements following refixation saccades: Type and control system analysis [Journal Article]. *Vision Research*, *12*(3), 467-475. Retrieved from `https://www.sciencedirect.com/science/article/pii/0042698972900909` doi: https://doi.org/10.1016/0042-6989(72)90090-9

Website, P. A. (2021). *Pilatus aircraft ltd. pc-7 factsheet.* Retrieved from `https://www.pilatus-aircraft.com/data/document/Pilatus-Aircraft-Ltd-PC-7-Factsheet.pdf` (Accessed on: 24-06-2021)

Wickens, C. (2015). Noticing events in the visual workplace: The seev and nseev models [Book Section]. In (p. 749-768). Cambridge University Press. doi: 10.1017/CBO9780511973017.046

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology [Journal Article]. *Human Factors*, *45*(3), 360-380. Retrieved from `<GotoISI>://WOS:000187038200002` doi: 10.1518/hfes.45.3.360.27250

Williams, R. W. (1991). The human retina has a cone-enriched rim [Journal Article]. *Visual Neuroscience*, *6*(4), 403-406. Retrieved from `https://www.cambridge.org/core/article/human-retina-has-a-coneenriched-rim/270D93EB0738DB528FF3CC86E78A9A31` doi: 10.1017/S0952523800006647

Wilson, M., Vine, S., Brewer, J., Bright, E., Masters, R., & McGrath, J. (2011). Gaze training improves technical performances and resistance to distractions in virtual laparoscopic surgery [Conference Proceedings]. In *Sages annual meeting 2011.* Society of American Gastrointestinal and Endoscopic Surgeons (SAGES). The ....

Xiong, W., Wang, Y., Zhou, Q., Liu, Z., & Zhang, X. (2016). The research of eye movement behavior of expert and novice in flight simulation of landing [Conference Proceedings]. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (p. 485-493). Springer International Publishing.

Zaal, P., Lombaerts, T., Mumaw, R., Billman, D., Torron, I., Jamal, S., … Feary, M. (2021). Eye-tracking analysis from a flight-director-use and pilot-monitoring study [Conference Paper]. In *AIAA Aviation 2021 forum.* American Institute of Aeronautics and Astronautics. Retrieved from `https://doi.org/10.2514/6.2021-2995` doi: doi:10.2514/6.2021-299510.2514/6.2021-2995

Ziv, G. (2016). Gaze behavior and visual attention: A review of eye tracking studies in aviation [Journal Article]. *The International Journal of Aviation Psychology*, *26*(3-4), 75-104. Retrieved from `https://doi.org/10.1080/10508414.2017.1313096` doi: 10.1080/10508414.2017.1313096

A
PC-7 Model Description

This Appendix serves to detail the PC-7 model that has been developed in collaboration with MultiSIM B.V. The purpose of this model is to be close to the behaviour of a PC-7, but there is no need for the behaviour to be identical to a real PC-7, this is in part due to the lack of real PC-7 flight data.

A.1. PC-7 Dimensions and Parameters

This section highlights the dimensions and properties of the aircraft. These were derived from Jane's All the World's Aircraft (Janes.com) and the PC-7 turbotrainer manual. Lastly, some geometric parameters (e.g. wing position) were derived from the image shown in Figure A.1.

All derived data are shown in Table A.1

An initial estimate for the aircraft's moment of inertia was derived using the method set up by Roskam (1985). This was later adjusted to be somewhat higher to improve the simulation's flight characteristics. The lift slope for the wings was found using the Athena Vortex Lattice (AVL) software developed by Mark Drela at MIT (Drela, 2004).

Table A.1: PC-7 Parameters

Parameter	Value
Propeller diameter	2.44 $[m]$
Wing Span	10.18 $[m]$
Wing Area	16.28 $[m^2]$
Mean Aerodynamic Chord	1.65 $[m]$
Aspect Ratio	6.3 $[-]$
Dihedral	3 $[deg]$
Sweep at 0.25c	1 $[deg]$
Flap area	1.77 $[m^2]$
Tailplane span	3.665 $[m]$
Horizontal tail area	3.95 $[m^2]$
Elevator area	1.6 $[m^2]$
Vertical tail area	1.98 $[m^2]$
Rudder area	0.9 $[m^2]$
Root Wing Profile	NACA 642A (15%)
Tip Wing Profile	NACA 641A (12%)
MTOW	1900 $[kg]$
stall speed (clean)	77 $[kts]$
wing lift slope	4.41 $[rad^{-1}]$
I_{xx}	5935 $[kgm^2]$
I_{yy}	7479 $[kgm^2]$
I_{zz}	9382 $[kgm^2]$
I_{xz}	600 $[kgm^2]$

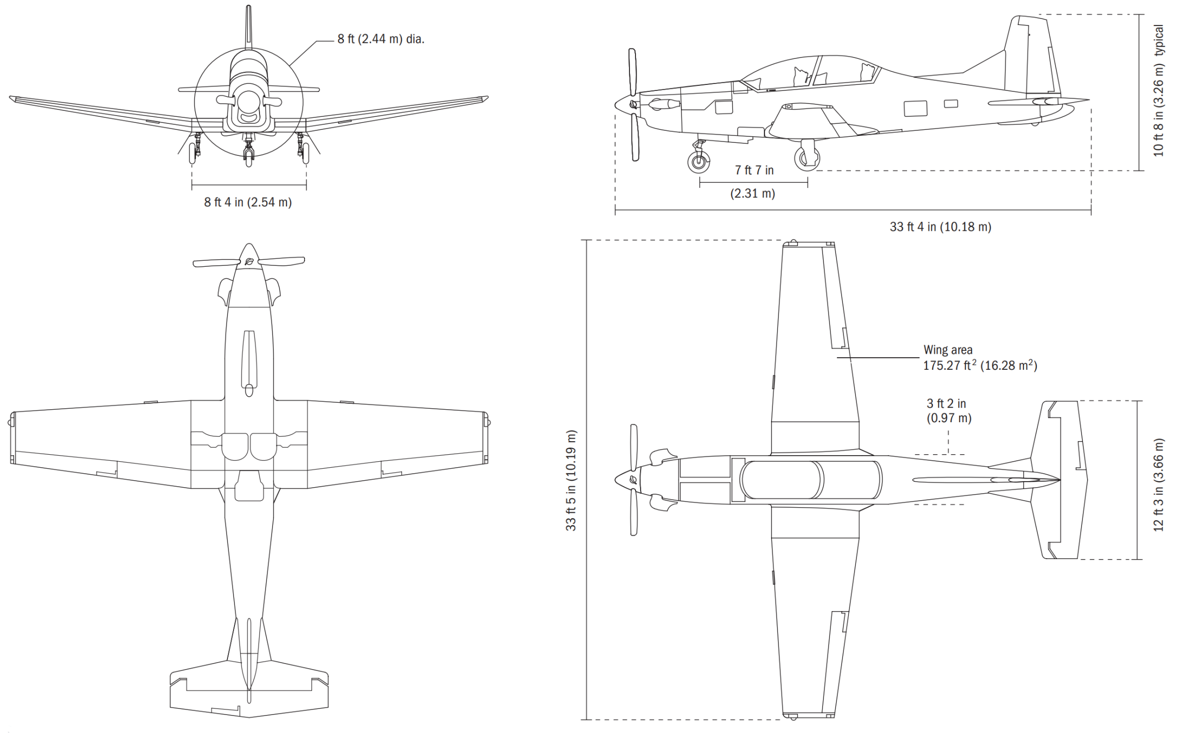


Figure A.1: PC-7 Dimensions Website (2021)

A.2. Flight model Implementation Details

The PC-7's flight model is built upon the generic flight model included with DSIM. This flight model, originally made for helicopter simulation, allows defining separate elements of the aircraft. An example of such an element is a lifting surface as part of the wing or as a control surface. Forces and reaction forces are then generated by all elements included, and the resulting motion is found.

For the implementation itself each wing is split into six parts (for both left and right an inner, outer and center partition). Splitting the wing into parts has several benefits within the DSIM flight model. First of all, this allows the aileron module to be re-purposed as flaps for the inner and center parts, while having the outer part keep functioning as an aileron. Second, the splitting of the wing allows for implementing some variation in parameters over the wingspan. Third, the position of the aileron and flap forces will be more accurate, since the model condenses surfaces into a point-force. Lastly, this allows for the inputs from the propwash module (Section A.3) to be varied over the span of the wing.

The ailerons are modeled based upon the description of forces on an aileron shown by Sadraey (2012). This is also the implementation in the flight model, which means it is fair to assume the forces are represented by the equation shown in Equation A.1. This equation shows the change in section lift coefficient (C_{L_A}) as a function of the lift slope (C_{L_α}), aileron effectiveness (τ_a) and aileron deflection (δ_A). The lift force on the aileron would then be described with Equation A.2

$$C_{L_A} = C_{L_\alpha} \tau_a \cdot \delta_A \tag{A.1}$$

$$L_A = C_{L_A} \cdot \frac{1}{2} \rho V^2 S_a \tag{A.2}$$

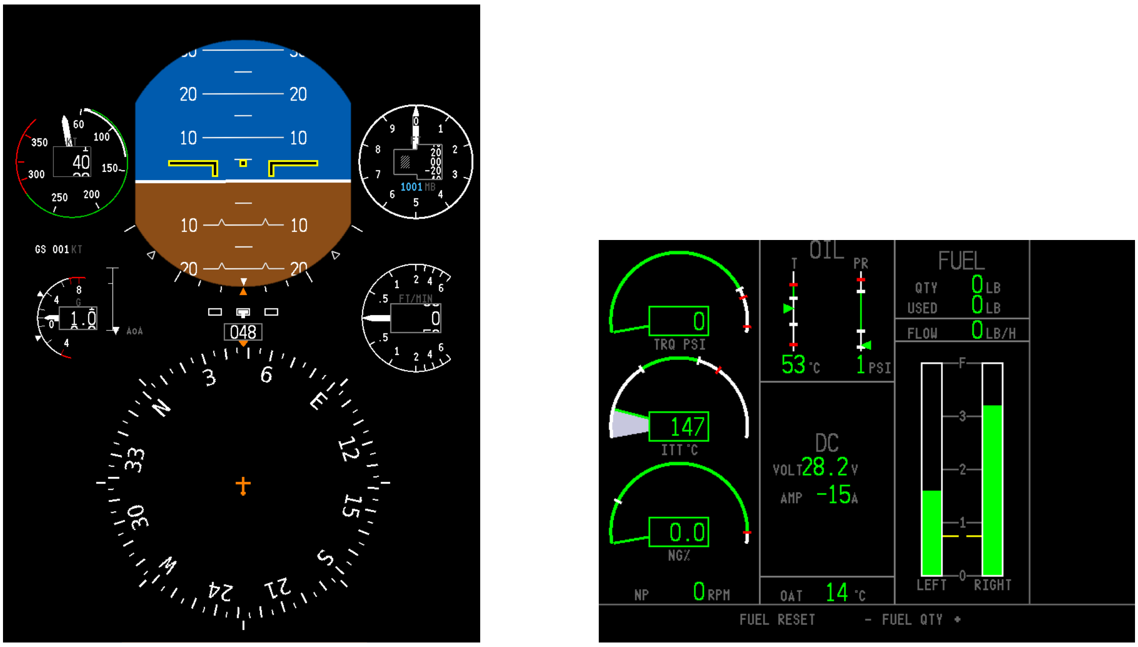
The Brunner control loading hardware allows for setting two parameters and a curve. The curve sets the force-displacement relation for the controls, and can be set using 9 equally spaced points. In another PC-7 simulator this relationship was found to be linear. The other two parameters that can be set are the trim point, and the gain. The former sets the point at which the forces are zero while the latter scales the aforementioned curve. The gain is controlled through a lookup table and the speed, meaning higher speeds lead to higher forces. These gain was tuned according to the forces measured at several speeds in the other simulator and pilot feedback. The speeds used for this are 80kts, 110kts, and 180kts since these are relevant to the PC-7's operations.

A.3. Propwash

The propwash on the vertical and horizontal tail was implemented by taking the vorticity value from the propeller and combining this with the vector of the distance to the surface experiencing propwash to find the local induced velocity vector. The resulting induced velocity was then taken into account in the calculation of the forces present on the lifting surface.

A.4. Electronic Flight Instruments

Two instrument panels were also implemented for the PC-7 model. The primary flight display is shown in Figure A.2a, containing an attitude direction indicator (ADI), vertical Velocity indicator (VVI), G-load indicator, Horizontal Situation Indicator, Angle of Attack (AoA) indicator and barometer. Figure A.2b shows the engine-indicating and crew-alerting system (EICAS) for the PC-7, this was deemed important enough to implement due to the significance of the engine to the way the PC-7 flies.



(a) Electronic Flight Instruments for the PC-7

(b) Engine-indicating and crew-alerting system for the PC-7

Figure A.2: Instruments PC-7

A.5. Engine Model

The PC-7 simulation also included a full engine model. This allows for more accurate flight, as well as having the EICAS be more accurate than it would be in most alternative approaches (e.g. look-up tables). This engine model was developed in Open Modelica by a fellow TU Delft student working with MultiSIM, Mik Swarts.

B Experiment Protocol Used

This appendix details the experiment protocol for the preliminary experiment as described in Chapter 7.

B.1. Pre-briefing

- Clean HP Omnicept's lenses and place in the cleanbox
- Turn on the computers labeled; MC, M7
- Load saved '_prelim' calibrations on the two computers
- Load these into the CTC world
- Check if the eye-tracking filter is set to 0.87 and 1.5
- Calibrate eye-tracking, press escape to return to DWorld
- Validate proper functioning of the eye-tracking
- Set the position of the PC-7 to be on the runway
- Fly to 1000ft and 130kts to verify proper functioning, verify functioning of control loading
- Close throttle and place PC-7 back on the runway with the parking brake
- Put HP Omnicept back into the cleanbox and turn on
- Disinfect hands
- Disinfect equipment.

B.2. Briefing

This section details the notes used in the briefing. As all participants and the briefing were Dutch the following notes will be Dutch.

- Uitleggen waar het experiment om gaat. Experiment gaat om het kijken naar waar piloten naar kijken tijdens het vliegen. (VR eyetracking) Deze data kan dan hopelijk gebruikt worden bij het opleiden van nieuwe piloten voor betere veiligheid. Een goede vergelijking is dat er bij de rijles veel gekeken wordt naar kijkgedrag, terwijl dit bij een kist als de PC-7 waar de instructeur achterin zit moeilijk is.
- Vragen en noteren van ervaring kleine kisten
- Flap settings en de gear worden up/down gedaan wanneer hierom gevraagd wordt. Nog niet bedienbaar via touchscreen.

- Rudder is iets te sterk op de grond, dus mocht het nodig zijn om te sturen is het beter om dit met differential braking (toe brakes) te doen.
- Instrumenten laten zien, vragen of alles herkenbaar is. (pngs hiervan in dezelfde map als dit bestand laten zien)
- Kort het circuitje uitleggen (qua hoogtes etc.), aangeven dat het ook op de EFB staat (deze afbeelding ook laten zien). Nadruk leggen op het langzaam landen. Aviano is lange baan (2.5km), heading 05. Belangrijk om ook goed het 'verkeer' om je heen in de gaten te houden.
- Ook een tweede ronde met touch and go.
- Lezen en tekenen van de consent forms.

B.3. After Briefing

- Bring participant to the simulator hall
- Show controls (Trim, Pedals, Adjustable seat)
- Let participant grab the VR headset from the cleanbox
- After the participant has taken a seat, give the instruction to put on the headset and make sure it can not move
- Start calibration (switch screen on M7 and press the HP Icon)
- Assist in calibration
- Reload DWorld after calibration
- Recenter VR position
- Start recording
- Let participant look at the ADI, the VVI and then the HSI
- Turn off parking brake
- Start experiment

B.4. During Experiment

- Pay attention to flap and gear calls
- Pay attention to the calibration
- When second round is done, enable the parking brake and let the participants check the same instruments

If the calibration turns out to be off, give the instruction to stop the aircraft after the first landing. Turn off the and repeat the calibration procedure.

B.5. EFB

The next page will show the exact image shown to the participants on the EFB, containing the information on the standard circuit that is flown in the experiment.

Take off:

- Take off power **42.5 psi**
- Roteren op **80 kts**
- Gear up
- Klimmen met **110 kts**
- Flaps up
- Circuit hoogte **1000 ft AGL, 1330 ft MSL**

Down-wind:

- Snelheid **110 kts**
- Power **8.5 psi**
- Downwind vliegen met de vleugeltip op de baan
- Gear down, Flaps Landing aan het einde downwind

Base:

- Normale base turn op **45 graden tov threshold**
- (op tijd) daling inzetten

Final:

- Approach speed **90kts**
- Landing **80kts**