32nd Congress of the International Council of the Aeronautical Sciences
September 6-10, 2021
Pudong Shangri-La, Shanghai, China

GAZE-BASED INTERFACE FOR FLIGHT CONTROL: RESULTS OF INITIAL EXPERIMENTS WITH THE SYSTEM

Manuel A. Rodriguez D.[1], Emilia Villani[1] & Petter Krus[2]

[1]Aeronautics Institute of Technology, Brazil.
[2]Linköping University, Sweden.

Abstract

Cockpit flight control interfaces have remained largely unchanged since the beginnings of Aviation (sticks, yokes, pedals and thrust levers). New technologies have emerged that have successfully been employed in other fields of technology regarding Human-Machine interaction, one such technologies is eye tracking.
The pilot is a fundamental element even in modern aircraft with high levels of computerized assistance. Emergency situations or environmental factors can lead to a situation were a pilot is overwhelmed by simultaneous tasks and this can lead to the conditions for an accident.
With these points into consideration, a modernized flight control interface is proposed were Eye tracking is used as a complementary input for yokes, side sticks and pedals. Review of literature and previous research suggests that this application has considerable merit for its potential benefits, this paper is in regard of early experimentation using the system in a highly realistic flight simulator. The development uses equipment available at the Instituto Tecnologico de Aeronautica (ITA) to create a prototype that enables the characterization of the proposed a system.
Results from this set of experiments contributed to the further refinement of the system, error correction, obtaining of operational experience and collection of data and feedback from participants that suggest that the proposed system could have a positive impact in the controlability of an aircraft by an inexperienced pilot during a high-workload situation

Keywords: Flight control interface, Flight control system, Man-Machine Interface, Eye Tracking.

1. General Introduction

Aviation has seen immense progress along more than a century since the first heavier-than-air flight, aircraft became highly complex and reliable machines, and the aerospace industry is now a worldwide reference in state-of-the-art technology. Man-Machine interfaces using body gestures as input methods have become highly reliable in the last two decades giving rise to considerable research in robotics where the command method of machines relies on natural gestures performed by the human operator. Even with the progress of technology and Aviation itself, the interface method between pilot and aircraft is still largely based on sticks and pedals, with research on innovative flight control interfaces being largely limited to unmanned aerial vehicles [5], this work focuses on the initial simulated flight testing experiments of an interface for flight controls that uses on the pilot's visual focus direction (known as "gaze"), this interface allows a pilot to give command inputs to a manned aircraft in a quick and simple procedure that could reduce the proficiency threshold required for a pilot to operate an aircraft in a highly realistic flight simulation environment.
The system has been developed following the Design Science Research methodology [4] where the artifact in question is the interface system itself, initial functionality assessment of the system has proven that it is feasible to control an aircraft using the proposed interface [7], and this initial experience allows for more detailed experiments using more refined control laws to be performed, further expanding into the possibilities of these new flight control interfaces. This paper is in regard to one of such experiments performed during the development process of the proposed system.

2. System Architecture

2.1 Cockpit gaze system

The system uses two types of signal inputs from the pilot, these are referred to as gaze signal and spatial signal. The gaze signal conveys visual focus direction in relation to the head of the pilot while the spatial signal conveys position and orientation of the head of the pilot in relation to a fixed reference system. The gaze signal is obtained from the Tobii Pro Glasses 2 while the spatial signal is obtained from the HTC Vive Trackers.

The eye tracking glasses used rely on retina geometry detection with infrared cameras to obtain a vector for each eye's orientation [1] using a process known as Pupil Corneal Center Reflection (PCCR). These resulting vectors from both eyes can be intercepted to obtain a spatial focus point towards which the actual wearer's gaze is focused, the direction from the point of view of the subject towards this focus point is referred to as the gaze vector.

Spatial signals are captured using components from the HTC Vive virtual reality kit know simply as "Tracker". This equipment relies on a system of rotating laser beam base stations (Lighthouses, as these are named) to perform spatial tracking. It also uses accelerometers and gyroscopes to perform dead reckoning between laser beam sweeps [3]. The spatial signal generated by the Tracker consist two components: a Cartesian three-dimensional vector and a quaternion. The Cartesian vector corresponds to the Tracker's location in space in relation to a fixed frame of reference set by the Lighthouse base stations, while the quaternion corresponds to the Tracker's rotation in relation to this same frame of reference.

In the integrated system, the Tobii Pro Glasses 2 are paired with an HTC Tracker attached to its side as shown in figure 1, the glasses provide a gaze vector relative to their own frame while the Tracker gives the exact location and rotation of the glasses inside the cockpit, combining these two elements it is possible to obtain a pilot's gaze vector that originates at the pilot's point of view in relation to the cockpit, and points towards the direction that he or she is looking at. This last vector is referredd to as the *cockpit gaze vector* in order to distinguish it from the original gaze vector generated by the glasses in relation to their own frame of reference.

Figure 1 – The Wearable hardware used in the proposed system, a Tobii Pro Glasses 2 headset with an attached HTC Tracker.

These signals are then integrated and processed in real time, the cockpit gaze vector is used to determine the point on the simulator's spherical screen where the pilot is focusing, or in what cockpit instrument if the pilot is looking inside of the cockpit. If the gaze focus point is in the screen, the pilot can use a button located in the Yoke to engage the gaze-guided system, this will activate one of four possible control modes that will modify the aircraft trajectory based on the location that the pilot was looking at. for the experiment presented in this paper only one of those control modes is tested; heading change mode, where the lateral gaze angle offset from the "nose" of the aircraft is proportional to a change in the heading of the aircraft at the moment that the system is engaged.

2.2 Flight Simulator

The flight simulation in this architecture is done at the *Simulador de Voo com Plataforma Robotica* (SIVOR) of the Aeronautics Institute of Technology (ITA), it is shown in figure 2.

Figure 2 – Composite image of SIVOR at the *Centro de Competência em Manufatura* (CCM), one of ITA's laboratories. image credit: ITA.

SIVOR is a prototype flight simulator developed by ITA and Embraer that uses an industrial robotic arm mounted on rails as the motion platform instead of the more common Stewart platform with six prismatic actuators, SIVOR's approach increases the available work space and in theory allows to perform sustained accelerations for longer periods of time in comparison to common simulators. SIVOR's cockpit is modelled after an Embraer Phenom 300 and contains both sidestick and yoke inputs while its flight dynamics model represents an Embraer Legacy 500.

3. Experiment

This experiment was designed as a Factorial Experiment with two factors; Interface type and induced Workload level. This paper is in regard to flights using one of the Latero-Directional control modes only. Figure 3 shows some of the subjects participating in the experiment.
Also, this was the first set of experiments that was executed in SIVOR using the Yoke as the physical interface, previous experiments were done on a simple desktop computer, the highly realistic cockpit and simulation model of SIVOR presented a deep immersion for pilots participating in the experiment due to the physical sense of being in a cockpit and having a field of view covering 180 degrees.
The factors for this experiment consist of:

- Interface type: with three levels; Yoke input, Gaze input and Multimodal input (combination of both Yoke and Gaze);

Figure 3 – Composite image showing three of the eight participating pilots, the cockpit shown is SIVOR during the experiment.

- Workload: with two levels; Low and High. Low workload had the pilot follow only the main task of the experiment, while the High level introduces a secondary task in order to divert the pilot's mental and temporal resources from the main task;

Figure 4 illustrates the combinations of these factors for the experiment.

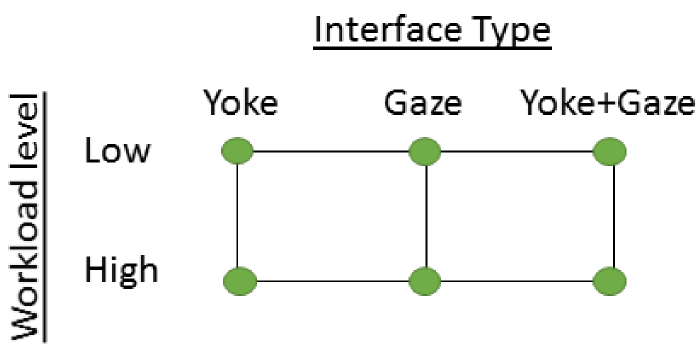

Figure 4 – Combination of factors for the experiment.

3.1 Experiment Setup

SIVOR has full motion capabilities, but hardware limitations with the HTC Lighthouse base stations meant that it must remain stationary during this phase of experiments, specifically this hardware requires sensing of the local gravity vector in order to define its attitude in relation to the physical world, and since at least one of these base stations must be located inside the cockpit, having motion enabled would add an acceleration component to the signal sensed by the Lighthouse's accelerometers. This would disturb the Lighthouse's references forcing it into reset each time the cockpit is moved.
The pilots were seated on the left seat of SIVOR's cockpit while wearing the Tobii Pro Glasses 2 Eye Tracker with an HTC Tracker attached to the right side of the glasses' frame. A button on the Yoke allowed the pilots to issue a gaze command. The sound system in the cockpit gave the pilots basic feedback regarding the status of the gaze guidance via aural alarms. A tablet computer was available

to the pilot at all times to fill a set of questionnaires, the test operator was also present in the back of the cockpit to provide assistance to the pilot should it be required.
Each pilot fills three types of questionnaires in this experiment; a pre-experiment questionnaire regarding their background, a post-experiment questionnaire for giving their feedback on the system, and the NASA-TLX workload assessment. The content and intent of each of these questionnaires is described in section 3.4

3.2 Experiment's Tasks

The simulated aircraft began the experiment in flight stabilized at 10.000 feet Above Sea Level near Sao Jose dos Campos, the pilot's main task is to "catch" a red ball that is floating in a fixed position in the sky, Figure 5 shows a depiction of the pilot approaching this red ball in SIVOR.

Figure 5 – A Test subject in SIVOR guiding the aircraft to the target using gaze.

The pilot had control only of the Latero-Directional axis of the aircraft by means either of the Yoke, the Gaze interface or both depending on the interface type being tested at the moment. The aircraft's altitude was maintained automatically and the airspeed was also maintained by an Auto throttle system when the workload level of the experiment was set to low, and maintained manually by the pilot when the workload level was set to high, the pilot controls airspeed by means of the two thrust levers present on the cockpit.
The red ball target always maintains the same altitude as the aircraft in case that it varied during flight, so the pilot's only concern to reach the ball is to adjust the aircraft's heading to point it towards the ball. Each time that the ball passed by one of the sides of the aircraft (90 degrees to either side from the nose of the aircraft) it was considered that the pilot had reached it, and a new ball was generated at 9 Kilometers from the aircraft, always within the pilot's field of view to allow the pilot to find it by just looking around the cockpit's windows and then having to maneuver the aircraft to reach the next ball.
Reaching the ball is considered the primary task of this experiment, the distance from the ball to the aircraft when the ball passes by either side of the cockpit is registered as the performance metric for this task. The secondary task is considered to be the manual control of airspeed within specific limits and the total amount of time when airspeed is within these limits is considered the performance metric for this task. When the secondary task is present on the pilot's duties it is considered that the induced workload level is set to high, and that it is low when it is absent.

3.3 Procedure

Before beginning the experiment, the pilots are briefed on all of the necessary information and on the objectives of the experiment, they proceed to fill a initial questionnaire regarding their background and then proceed to wear the Eye tracking glasses and enter the cockpit, at this point they are introduced to the controls that they would be operating and proceed to have a training session were they get accustomed to the handling of the aircraft via all of the interfaces. This training session also includes the target red balls in order to have the pilot familiarized with their appearance and to learn to perform their main task. The secondary task of maintaining a specific airspeed is also introduced and practiced.

When the pilot felt satisfied that all of the information has been understood and that they are ready to begin the experiment, they receive the initial NASA-TLX factor pairing procedure, here they have to sort all six TLX factors on a pairwise manner indicating which of each pair is, in their view, the most important contributor to the workload being introduced by the tasks that they have been given.
Each pilot performed a total of six experiment Flights, it is considered that a Flight begins with the pilot ready take control of the aircraft, a total of six balls appeared per flight one after the other was reached. Once the pilot reached six balls the flight was considered to be completed, at which point the pilot released control of the aircraft.
The first three Flights used each of the three interfaces while keeping the auto throttle enabled to maintain the induced workload level at low. Flights numbered four through six repeated the use of each of the three interfaces in the same sequence but this time with manual throttle control by the pilot to raise the induced workload level to high. At this workload level the pilots had to maintain an airspeed of 245, 250 or 255 knots +/- 2, the target airspeed was changed every time that a ball was reached and was communicated verbally to the pilot by the test operator.
At the end of each Flight the pilot fills a NASA-TLX questionnaire regarding the flight that they just performed, in this case they give a value to each of the six TLX factors ranging from zero to one hundred. Furthermore, a final questionnaire is filled by each pilot after all six flights have been completed, this one is for documenting the pilot's feedback on the system. The following list contains a summary of the steps that each pilot undertakes from the beginning to the end of the experiment.

1. Briefing;
2. Pre-Flight Questionnaire;
3. System Calibration and setup;
4. Practice Run;
5. NASA TLX pairwise;
6. Flight number 1: Yoke only, no secondary task, 6 balls;
7. TLX scales;
8. Flight number 2: Gaze only, no secondary task, 6 balls;
9. TLX scales;
10. Flight number 3: Multimodal, no secondary task, 6 balls;
11. TLX scales;
12. Flight number 4: Yoke only, manual throttle, 6 balls;
13. TLX scales;
14. Flight number 5: Gaze only, manual throttle, 6 balls;
15. TLX scales;
16. Flight number 6: Multimodal, manual throttle, 6 balls;
17. TLX scales;
18. Post-Flight Questionnaire;
19. Debriefing.

A multimodal flight is one where both yoke and gaze inputs can be used by the pilot at his or her own discretion, the term comes from the availability of different possible input methods by the pilot [6].
A total of eight pilots were able to participate in this round of experiments.

3.4 Questionnaires

Each participant receives three different questionnaires to fill at various points during the experiment, these are filled using a tablet computer provided by the test operator, Figure 6 shows this moment.

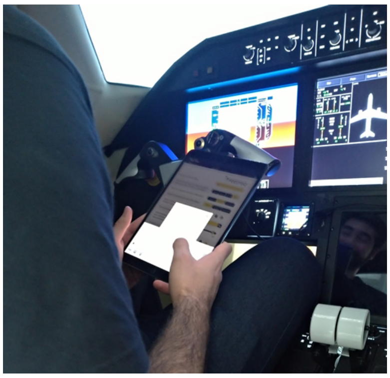

Figure 6 – A Test subject in SIVOR filling one of the questionnaires of the experiment.

3.4.1 NASA Task Load Index (TLX)

The NASA Task Load Index (NASA-TLX) is part of a methodology for evaluating workload as perceived by the subject [2], in this case it is used for comparing the impact that the interfaces and workload levels have on each pilot and quantifies six factors: mental demand, physical demand, temporal demand, performance, effort and frustration.

At the beginning of the testing procedure the pilot is presented with the definition of each of these factors, and then each of these factors is presented in pairs with all the others pair by pair, the subject has to indicate which of the pair being presented is the most important contributor to the workload level being experienced on the task that they have been presented, this is repeated until the subject has ranked all pairs. This part of TLX is performed every time that the experimenter desires to probe the workload being experienced by the subject, in this part the subject assigns a score to each factor ranging from zero to one hundred. Details regarding the implementation and analysis of this method can be found at [2].

3.4.2 Pre-Flight questionnaire

The pre-flight questionnaire is intended to obtain background information from the pilot, its contents are as follow:

- Flight simulation experience: None, little, intermediate or high;
- Experience as pilot: None, little/Student pilot, intermediate/Private pilot or high/Professional pilot;
- Experience in SIVOR: None, little, intermediate high or part of SIVOR team;
- Dominant Eye: left or right;
- Glasses: if corrective glasses are used, state the correction strength of the lens;

3.4.3 Post-Flight questionnaire

The post-flight questionnaire documents feedback from the pilot regarding the experience with the system, its contents are:

- How did workload increased when maintaining airspeed was necessary?: None, little, intermediate or high;
- What was the easiest configuration to use?: Yoke, gaze or multimodal;
- What was the Hardest Configuration to use?: Yoke, gaze or multimodal;

- What is your perceived accuracy of the gaze system?: Low, medium, high or variable;
- What was the percentage of correct gaze command interpretation?: scale from 0 to 100 percent;
- How comfortable is the hardware?: scale from 1 to 5, higher is more comfortable;
- Overall satisfaction using the system: scale from 1 to 5, higher is more satisfactory;
- Additional comments: free textual input to register any other comments.

4. Results

Preliminary results showed large variability by pilot; for some pilots the gaze inputs proved beneficial but in other cases it either had no impact or was detrimental. This could be in part due to some pilot having insufficient training time to adapt themselves to the tasks that were given for the experiment, but another contributing factor could be the lack of complete accuracy of the implementation of the gaze interface used during the experiment as it had certain calibration issues that were solved after the experiment was completed.
Variability of results even on the same pilot could support these theories, although improvements to the design of the experiment itself can indeed be performed to facilitate the acquisition of meaningful and significant data.
On pilots that showed improved performance, the system allowed them to focus more on their secondary task after they gave a flight command using their gaze.

4.1 Flight data

A total of 48 Flights were performed between 8 pilots, this amounts to 288 total target balls across all flights. Each flight recorded 57 simulation parameters that include aircraft state variables like longitude, latitude, altitude, airspeed and heading, but also system-specific variables like guidance mode, guidance target, gaze vector, element in focus, command button status, etc. It is possible to reconstruct the most important aspects of each flight using this recorded data.
Analysis methods for the recorded data has included Analysis of Variance (ANOVA), Information Entropy and Pattern search for correlations in the observed data. preliminary findings suggest that on the available dataset the most influential element is the pilot, this means that comparing the effectiveness of the gaze interface is difficult between pilots as the statistical significance of each interface configuration tested becomes low due to the large variations caused by the pilot's behavior itself and most of the collected data seldom fits any common distribution.
The following figures present data recorded from the simulated flights performed for this experiment, each flight has been divided in six parts to observe the capture of each target ball individually.
Figure 7 contains data regarding both the primary and secondary workloads; the first subplot (on the top) displays the pilot's vertical gaze angle on the vertical axis, on this scale negative angles indicate that the pilot is looking "up" at SIVOR's projection screen were the visual representation of the outside world is being shown. Positive angles indicate that the pilot is looking "down" at the instrument panel on the cockpit. The exact angle toward any particular elements depends on variables that are not necessary to consider for this analysis but that are nonetheless recorded in case the necessity arises, however it is known that the pilot only has two points of interest: the simulation screen (for the primary task of catching the ball) and the airspeed indicator (for the secondary task of maintaining airspeed). Knowing that these two points of interest have clearly different vertical angles from the point of view of the pilot, it becomes possible to assume for simplicity that when the pilot's gaze is focused "downwards" then the airspeed indicator is the focus of visual attention at that moment, and that when the gaze is focused "upwards" the focus of visual attention is the simulation screen. the exact angle towards these elements may vary slightly if for example the pilot moves forward or backwards.
The same subplot also can contain two discrete elements besides the gaze angles; the points in time when the gaze command button is pressed to issue a gaze command and the points when a change of attention occurs. these points are shown as yellow and red circles respectively. in the case shown

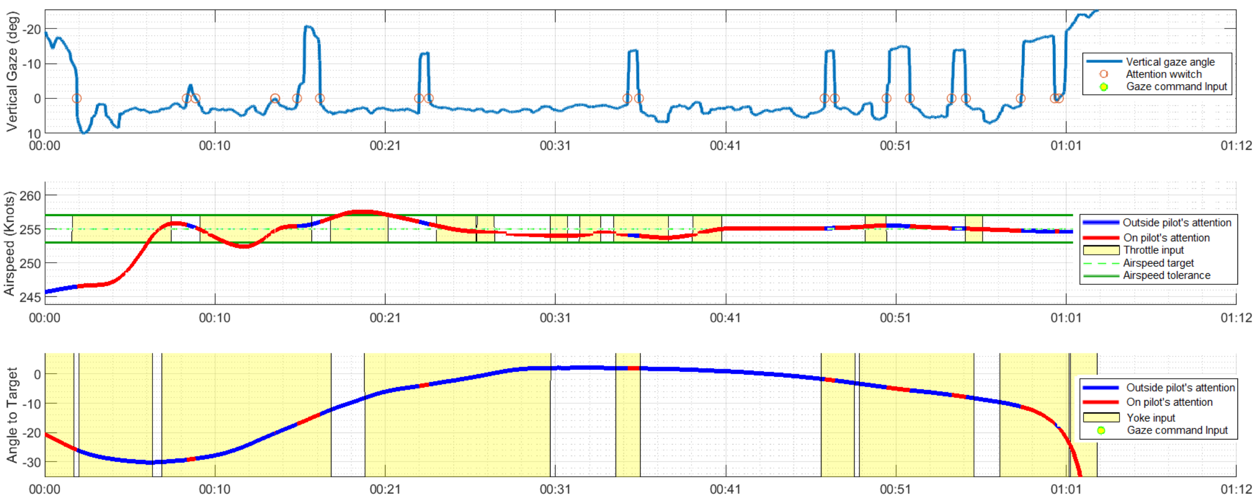

Figure 7 – Data for a Yoke-only flight with high workload (due to the secondary task of maintaining airspeed manually), the time interval is displayed in the horizontal axis and begins when the ball appears 9 kilometers from the aircraft and ends when the pilot captures it.

in figure 7 the flight depicted is Yoke-only, therefore no gaze command points are to be expected, but this description is valid for several other figures that follow the same display methods.

It is considered that an attention change has occurred each time that the pilot's gaze is focused on one of the two elements of interest and proceeds to look at the other element, this inversion is counted by detecting the vertical gaze signal crossing of a specific reference angle, a rising edge and falling edge pairing is considered a single attention change event as the pilot could be focusing on one element and only temporally checked the current status of the other element before reversing to the original element, this only counts as a single attention change event.

The second subplot (middle) of the same Figure 7 shows the Knots Indicated Airspeed (KIAS) in the vertical axis as a representation of the secondary workload were airspeed has to be maintained within certain margins from a specific target that changes every time that a ball is reached. For the segment depicted in Figure 7 this airspeed target is 255 Knots with a tolerance of plus or minus 2 knots as described in section 3.2This target and limits are represented by the dashed and solid green lines respectively.

Airspeed itself is represented in this subplot as data in either blue or red color, this variation in color is associated with the attention of the pilot; the data is red when the pilot is looking at the airspeed indicator and therefore knows that the airspeed is at that value in that moment in time, and it is blue when the pilot is looking elsewhere and therefore does not directly knows what the airspeed is at that moment, but might have a "mental image" of the situation since the last time that she or he paid attention to the airspeed indicator and might remember the last value and/or trend of accelerating or decelerating.

The last piece of information contained in this subplot is the intervals of time when the pilot is operating the thrust levers for adjusting airspeed, this is represented by the yellow-filled regions between the airspeed limits. When these regions are yellow, it represents that the pilot is moving the thrust levers in that interval of time, and that the position of the thrust levers is constant when this regions are white.

Moving the thrust levers is considered as "work" or "effort" being performed by the pilot, this represents that some resources are being allocated to the secondary task of maintaining airspeed. This regions are obtained by differentiating the position of the thrust lever for the number one engine (left engine), if the differentiation at any given point in time is non-zero, it means that the thrust lever has been moved. It is assumed that the left and right thrust levers are moved in parallel, thus differentiation of only one is necessary.

The last subplot (bottom) of figure 7 is presented in a similar style to that of the middle one, but its vertical axis represents the angle between the nose of the aircraft and the target ball at any moment

in time. Positive angles indicate that the target ball is at the left of the center of the aircraft and negative angles indicate that it is to the right.
This angle is also represented in red or blue and these colors represent the same principle as in the case of airspeed; when this line is red it means that the pilot is looking at the projection screen and therefore presumably knows where the ball is located, and it is blue when the pilot is looking elsewhere but the projection screen and therefore does not knows the actual position of the target ball, but might have a "mental image" of the last known position and trend of the ball's relative location.
This part of the figure also contains yellow-colored regions, these represent activity on the roll axis of the Yoke; when these regions are yellow it represent that the Yoke is being moved by the pilot to give roll commands, and it is white when the yoke's roll position is constant. This is obtained from differentiating the yoke's lateral angle, when this differentiation is non-zero it means that the yoke has been moved and therefore that activity present.
These regions represent "work" or "effort" being done by the pilot to accomplish the primary task of reaching the target ball by giving yoke inputs.
This part of the figure can also contain yellow dots representing gaze commands being given by the pilot, these discrete events are represented in the horizontal axis on the moment that they occur, while the vertical position of the dot represents the angle change that was commanded. If the pilot is looking at the ball and the system is correctly determining the gaze vector, then these points would always coincide with the current angle offset of the target ball, if a command is received with any error in the gaze vector, then the yellow dot will be far from the line representing the ball at that moment.
The situation presented in Figure 7 contains a yoke-only flight, therefore no gaze command points are seen at this figure.

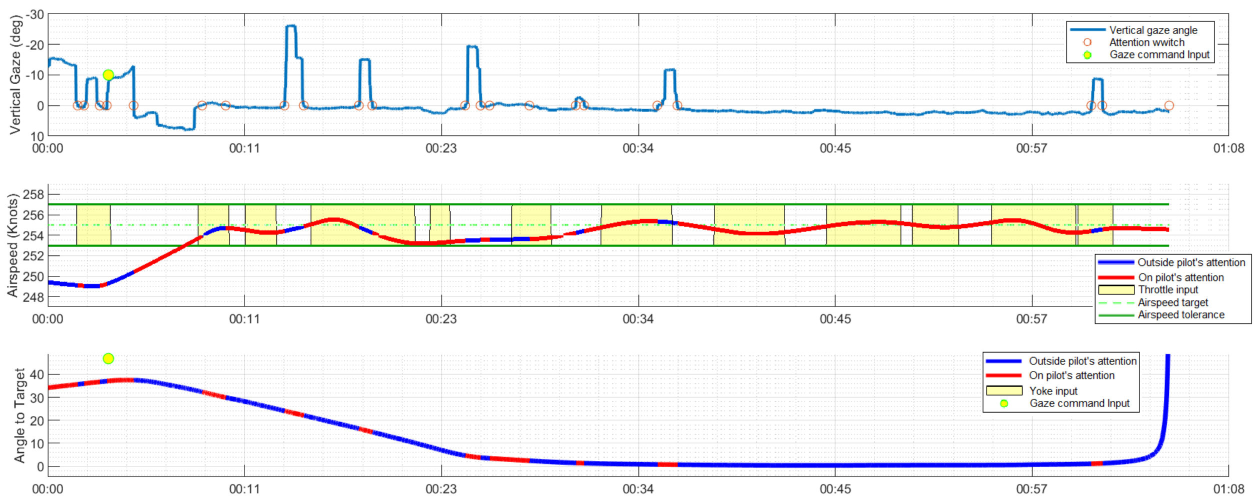

Figure 8 – Gaze-only data for a single target ball.

Figure 8 follows the same information display style as the previous figure, but in this case it displays information for a gaze-only flight by a different pilot, all of these figures start when a new ball appears and end when this ball is reached.
When comparing Figures 7 and 8 several differences can be noted; attention change events are distributed along the flight in the case of the Yoke-only flight whereas in the gaze-only flight these attention changes are grouped towards the beginning of the flight.
Next, it can be seen that only one instance of gaze command was necessary to align the aircraft towards the ball in figure 8, although this is an exception as usually several correction commands are required. Nevertheless, it can be seen that since the aircraft is now pointing at the target ball and that the primary task is seemingly accomplished, the pilot decides to focus more on the secondary task (maintaining airspeed). This is evident by the reduction in attention inversions after a monitoring period has ended when the aircraft points to the ball about 24 seconds into the flight, and due to the increase in throttle inputs in the gaze flight when compared to the yoke flight.
While a complete analysis of this data is still underway, it can be argued that since more resources

are available to the pilot due to the lack of necessity to give any other inputs to the primary task, then the pilot decides to allocate all of this resources to the secondary task to perform it as best as possible since the primary task appears to be completed.

In the case of the yoke-only flight, the airspeed was within limits for 80% of the time in comparison to 87% of the time for the gaze-only flight. Even across different pilots, this is a trend that is observed in the whole experiment: secondary task performance is improved when gaze is used, or at the very least its impact on the workload level of the pilot is reduced.

The fact that the attention spans focused on a single task are also larger on the gaze-only flight can also be counted as a potential positive impact: the pilot does not have to multitask as often.

However, more often than not the use of gaze in unimodal configuration is correlated to a reduction in the performance of the primary task due to the inaccuracies still present in the system.

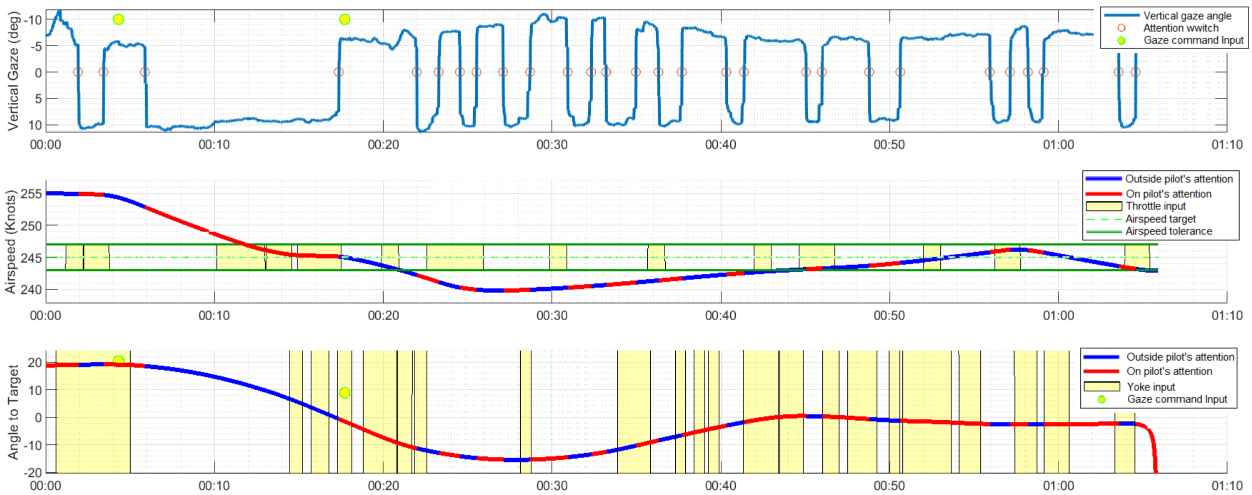

Figure 9 – Multimodal interface data for a single target ball.

The system is designed to be used as a multimodal interface, and while using each element in isolation allows to study the effect and performance with each one by separate, the most promising scenario is still the multimodal configuration. Figure 9 shows this scenario. In this case, a different attention switching pattern can be seen in the vertical gaze signal; the pilot issues a gaze command about five seconds into the flight, monitors that the maneuver is being executed for about two seconds, and then switches attention to the airspeed indicator until it falls within the accepted margins about 18 seconds into the flight. At this point, attention is given again to the first task but this time an erroneous interpretation is taken by the system; the ball is slightly to the right but it takes the command as a turn to the left, the pilot immediately recognizes this and overrides the system by giving yoke inputs until a turn in the correct direction is initiated about 21 seconds into the flight. Then, attention is given again to the airspeed indicator where the pilot realizes now that he did an unnecessary adjustment of the throttle that resulted in airspeed falling bellow required tolerances, and proceeds to perform a correction.

By this point, the pilot either considers that the gaze system will not take an appropriate command or simply prefers to reach the target using the yoke, so no further gaze commands are given but now the pilot has to monitor both the primary and secondary tasks actively, switching attention between both at continuous intervals of about two seconds for each task, this results in a delay to get airspeed back into tolerances about 43 seconds into the flight and also about the same time to position the target ball in front of the aircraft.

These three examples given in Figures 7, 8 and 9 were picked to show some cases were some interesting patterns can be observed, several other instances of similar patters can be observed across the experiment. Statistical analysis of the data has been performed and the only apparent element with high significance across the whole dataset appears to be the pilot, yet if a single individual pilot is observed then more consistent conclusions can be made, but these would be valid only for that pilot, with a few pilots sharing certain behaviours and results across them. In order to draw certain

conclusions more analysis is necessary, be it by employing different methods or by utilizing any other of the recorded variables to properly uncover the effect of each factor. Observations made in this experiment are already influencing the design of the next experiment to account for this variability across pilots.
The previous examples regarded behaviours and performance from an overt point of view of the system, yet some patters can also be seen that might be useful for a passive operation of the system to monitor the pilot's attention, figure 10 shows one of these cases.

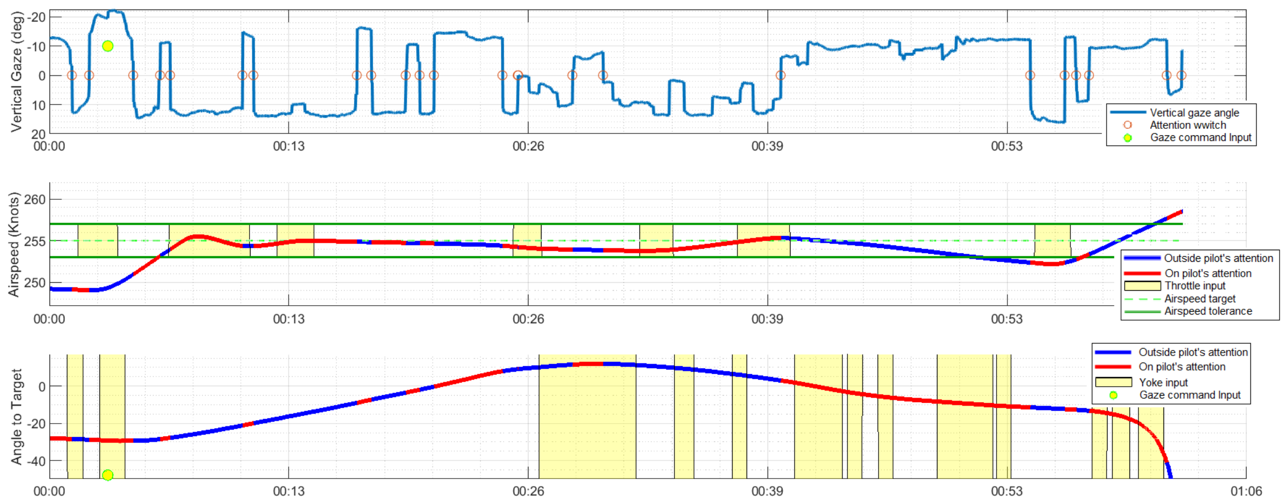

Figure 10 – Multimodal fight data, note the irregular and unstable gaze angle towards the end of the flight and the resulting excursion of airspeed outside both tolerances and continuous deviation of the target ball form the center of the aircraft (the nose of the aircraft is at zero degrees)

In this case, the first half of the flight proceeds as expected with the pilot giving a gaze input and then monitoring the aircraft while most of the attention is dedicated to the secondary workload. In this first half of the flight the airspeed is stabilized near the actual target and the ball is near the center of the aircraft, with the pilot giving yoke inputs to do final adjustments.
About thirty seconds into the flight, the gaze pattern changes; it is no longer stable once an attention change is performed but it erratically varies up and down, this coincides first with the target ball moving away from the center line of the aircraft's nose and later with the airspeed drifting away from its limits, all of this while the pilot still gives yoke inputs but also while the instability in the gaze vector continues.
Eventually, the pilot gives one final throttle correction that overshoots the airspeed to the opposite tolerance limit, at the same time that the ball continues to drift away to one side of the aircraft until it passes 90 degrees to one of the sides, at the moment of which it is considered that the aircraft reached it and the segment ends.
No recorded equipment anomaly occurred during this segment, and this particular pilot performed very well in other segments and flights, but it was the halfway through the last flight of this pilot after over two hours of experimentation at the end of a regular workday for the subject. Therefore it is possible that this instability be associated with a reduction in attention, fatigue or excess of workload, the moment when the final throttle command is given and the airspeed overshoots to the other extreme of the limit could indicate frustration at the situation. This is however a case that also requires further study before drawing any definitive conclusions, few other instances of similar events were recorded in the experiment.
Figure 11 part (a) shows how the performance metric for the primary task varies across all six flights by the first pilot that participated in the experiment. the first three columns correspond tho the first three flights with yoke-only, gaze-only and multimodal configurations without any secondary task, the next three columns represent the same order of interface configuration but this time with the added secondary workload of maintaining airspeed.

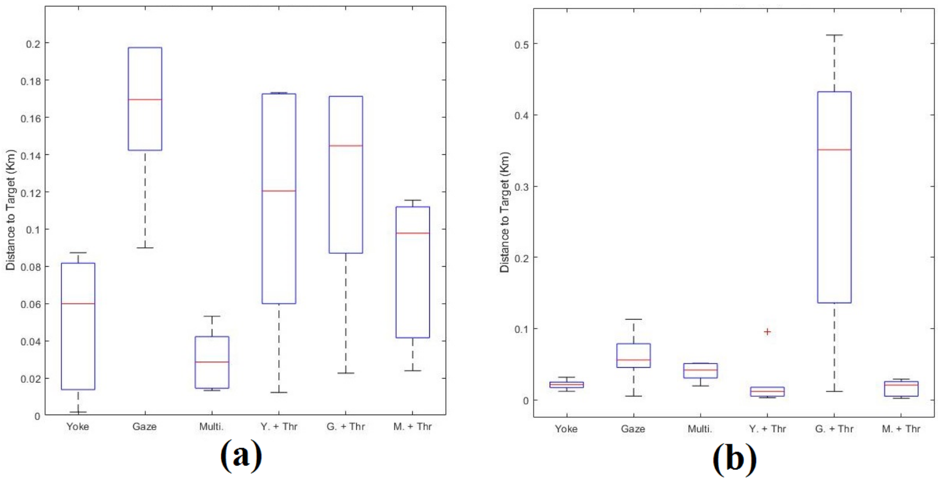

Figure 11 – (a): Primary task performance for Pilot number 1 across all six flights corresponding to the three configurations and two workload levels tested. The performance metric for the primary task is the final distance between the aircraft and the center of the ball once it passes to one of the sides of the aircraft. (b): Primary task performance for pilot number 2.

Figure 11 part (b) is similar to Figure 11 part (a) but it shows the second pilot to perform the experiment, these two figures illustrate how different one pilot's performance can be from the other but most importantly: it shows that each pilots is affected differently by each factor, pilot number 1 suffered a notable drop in performance in the gaze-only flight with no secondary task when compared to the yoke-only flight of the same workload level, but this drop in performance is lees noticeable once a high level of workload is introduced by including the secondary task.
Pilot number 2 on the other hand, has better overall performance even with high workload, excluding the gaze-only high workload scenario. This can be attributed to pilot number 2 having "learned" faster during the experiment compared to pilot number 1, by the fourth flight (corresponding to the fourth column, the yoke-only high workload flight) this pilot had more domain of the situation and was able to improve his performance even though the difficulty was increased in relation to the initial flights.

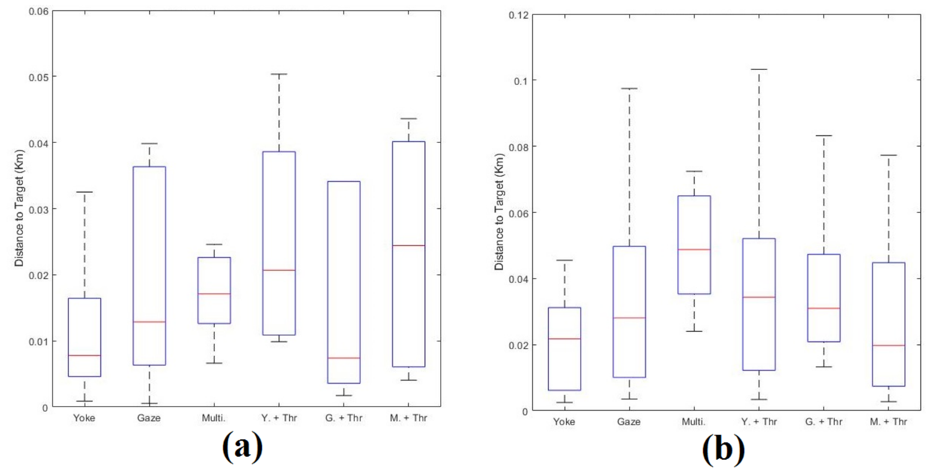

Figure 12 – (a): Primary task performance for pilot number 4. (b): Primary task performance for pilot number 8.

Figure 12 part (a) shows the same type of plot for another Pilot, in this case Pilot number 4. Here a different pattern can be seen in comparison to the two previous pilots, pilot number 4 had the most consistency across his own flights if compared to the consistency of other pilots among themselves, but it is still not clear the effect that each factor have on this pilot.
Pilot number 8 also displays a different pattern as shown in Figure 12 part (b), its performance

decreases on the first three flights but it then improves between the fourth and sixth flights despite the increase in workload due to the introduction of the secondary workload.

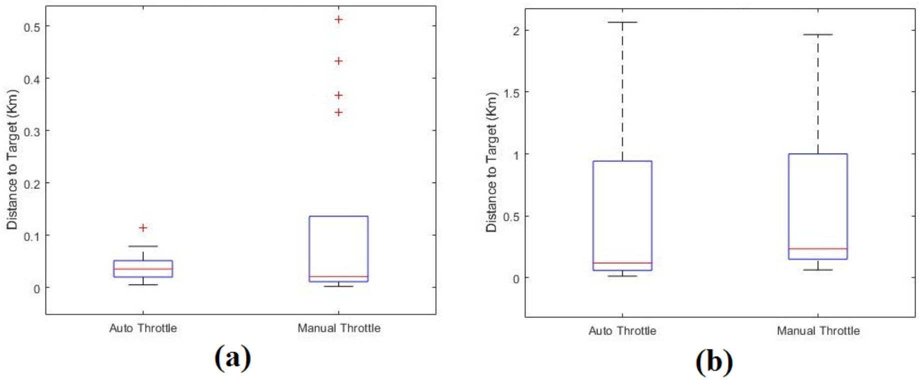

Figure 13 – (a): Impact of Workload level (secondary task) on primary task performance for pilot number 2. (b): Impact of Workload level (secondary task) on primary task performance for pilot number 6.

These example illustrate some of the difficulties encountered analyzing the data from this experiment, the best approach for drawing conclusions that has been encountered so far is to isolate each pilot and make separate conclusions for each one, combining data across pilots results in low significance on all factors even after excluding outliers. The next round of experiments needs be designed with this in mind in order to prevent recurrence of this problem.

4.2 Pre-flight Questionnaire

This section contains the summary of all pilot's answers to the pre-flight questionnaire filled before the experiment began, this contains data regarding their background.

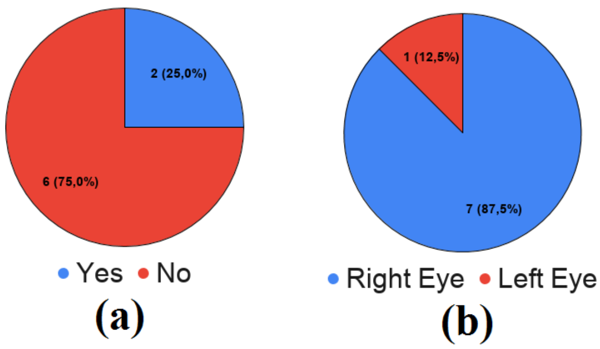

Figure 14 – (a): Do you use corrective glasses. (b): Dominant Eye of the Subject.

Figure 14 part (a) shows that only two of the participant pilots used glasses, their lens correction factor was 5,5 and 1,0 and neither presented similar results between them. Figure 14 part (b) shows that only one of the participant pilots had a left dominant eye, this factor is relevant only in the calibration of the equipment.

Figure 15 part (a) shows that half of participants had some form of piloting experience either in manned aircraft or unmanned aerial vehicle operation, the other half did not have experience at all. the relevance of this experience and its impact on the experiment's results is still being studied. Figure 15 part (b) shows that only three participants had no previous experience on SIVOR, this could be relevant as the familiarity with the simulator could affect the learning time for each participant, this correlation, however, has not yet been found if it exists. Two of the participants were members of the SIVOR development team. Figure 15 part (c) shows that again three participants had no previous experience on in any form of flight simulation, video games are not considered simulators in this

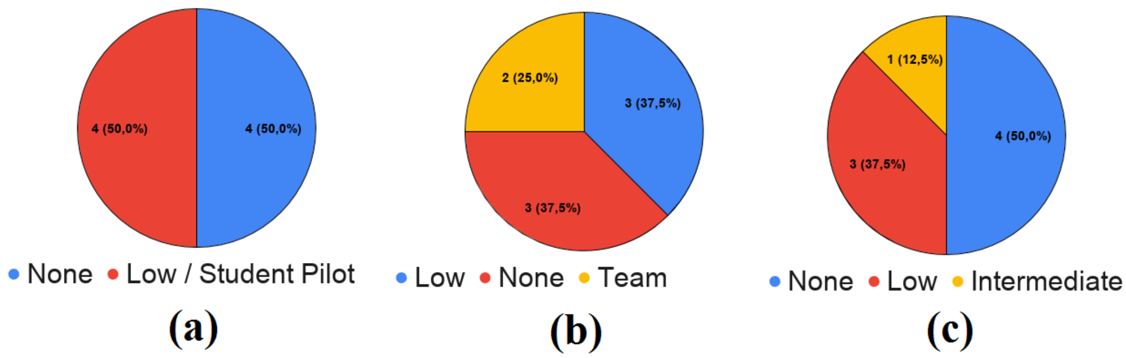

Figure 15 – (a): Experience as Pilot. (b): Familiarity with SIVOR. (c): Flight Simulation Experience.

context. This could be relevant as flight simulation experience could improve the airmanship of a pilot.

4.3 Post-flight Questionnaire

This section contains the pilot's responses to the post-flight questionnaire filled after they concluded the experiment.

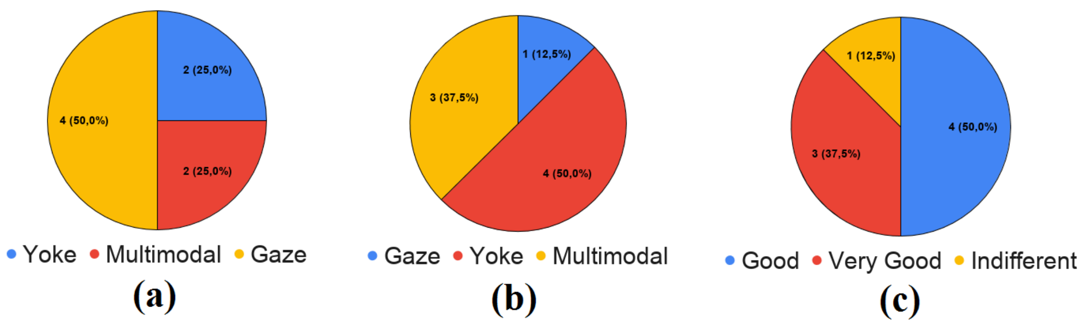

Figure 16 – (a): Easiest Interface to use. (b): Hardest Interface to use. (c): How was the experience using the system?.

Figures 16 part (a) and 16 part (b) show that half of the pilots perceived the gaze-only configuration to be the easiest to use in opposition to the yoke, whom two of those same four participants regarded as the hardest, the other two of that group regarded the multimodal configuration as the hardest to use. In Figure 16 part (c) it can be seen that all except one of the pilots had a pleasant experience performing the experiment.

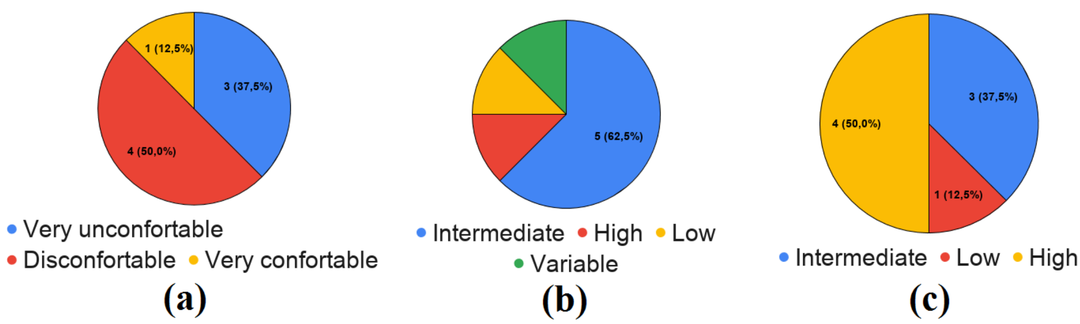

Figure 17 – (a): Level of Comfort of the Hardware. (b): Perceived system Accuracy. (c): Workload increase due to secondary Task.

Figure 17 part (a) shows that all except one of the pilots found the wearable hardware to be uncomfortable due to the unbalanced weight distribution and the excessive pressure that the glasses exert on the pilot's nose when they are tightened to prevent accidental displacements. Figures 17 part (b) and 18 regard the reliability of the system, most pilots finding intermediate accuracy and a mean 71% correct command interpretation across all pilots.

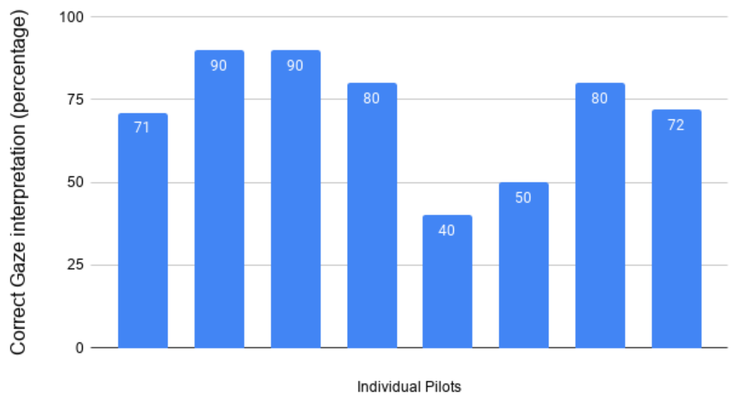

Figure 18 – Perceived correct interpretation of Gaze commands.

Figure 17 part (c) allows the pilots to express from their point of view how was the impact of the workload factor induced by the secondary workload on their ability to carry on both the primary and secondary tasks, half of the pilots found that it had a high impact while the rest except one fount it to be intermediate. The pilot stating that it had a low impact still had changes in performance on the primary task during the experiment.

4.4 Task Load Index

The TLX results are most useful to identify the individual contribution of its six factors to the overall workload, but in this case it is interesting to observe the weighted average of all factors as a measure of workload change across different flights. Then, grouping this data by the experiment's factors of interface type and workload in the form of presence or absence of a secondary task, and subsequently normalizing the data by re-scaling it in relation to the yoke-only flight for use as a baseline, a possible pattern emerges from the data shown in Figures 19 and 20.

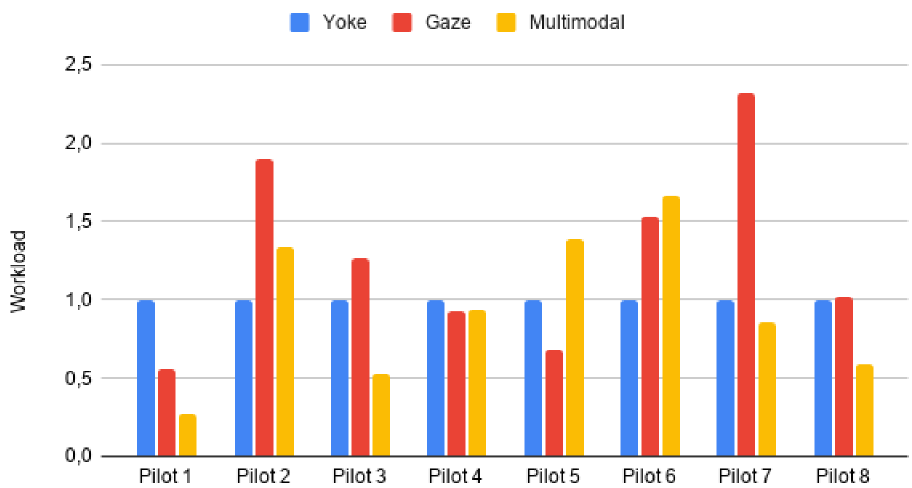

Figure 19 – Relative change in workload by interface used, normalized in relation to yoke-only flights. No secondary task.

During the flights with low induced workload level (no secondary task to perform), five pilots found that their workload increased when they had to perform the gaze-only scenario in relation to the yoke-only

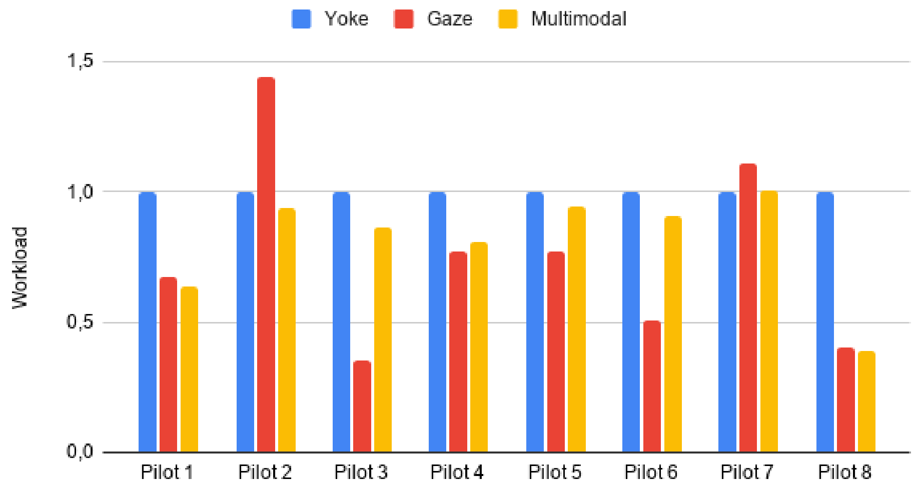

Figure 20 – Relative change in workload by interface used, normalized in relation to yoke-only flights. With secondary task.

scenario, while three pilots perceived that their workload increased with the multimodal configuration also in comparison with the yoke-only scenario.
Then, during the flights with high induced workload (those with the secondary task present), only two pilots perceived that the gaze-only scenario increased their workload (down from five in the previous case) and no pilot thought that the workload increased with the multimodal configuration in this case (down from three in the previous case).
This could represent an important impact of this system on the pilots as it would effectively reduce their workload as perceived by the pilots themselves.

4.5 Feedback from participants

This section contains a summary of feedback from the test subjects in the form of comments during debriefing or registered in the "additional comments" section of the post-flight questionnaire.

- On multimodal flights, it was compelling to use the gaze interface due to the ease of use compared to operating the yoke;
- One pilot noted that after a short while, he stopped noticing the eye tracker and acted naturally as he got used to giving commands to the aircraft using gaze inputs, noting that it was effortless to do so (only look at the target and press a button);
- More feedback on the current state of the system could be useful, the pilot making this comment noted that having a periodic reminder on the state of the gaze system would be more useful than the currently implemented aural signals at activation and deactivation;
- Visual feedback on the direction that the system detected a command input could be useful to increase situational awareness on the state of the system;
- The system's hardware at its current state is very uncomfortable, pilots commented that after about half an hour of continuous use they would appreciate to take it off;
- Three pilots noted that it was easier for them to give gaze commands to the left side of the aircraft in comparison to the right side and felt that these commands were also more accurate. This might be related to the fact that they were seated on the left seat;
- Five pilots noted that it was most effective to give large trajectory adjustment using the gaze input and afterwards give small corrections using the yoke, this allowed to dedicate more time to the secondary workload while the gaze system controlled the aircraft, but its accuracy meant that is was still necessary to use the yoke in the final corrections;

- One subject noted that if accuracy is improved, the effect on workload reduction of the gaze system will be increased by a large margin.

5. Conclusions

This set of experiments provided valuable information regarding the operation and capabilities of the system, the data and experience obtained was used to implement a series of improvements to the overall reliability, accuracy and ergonomics of the system as a whole. Feedback from participants suggest that the system could have a more relevant impact as the cockpit workload increases. Another set of experiments incorporating the improvements to the system would have been already conducted by the time of the publication of this paper.

Gaze guidance overall showed larger margins of error when the target was reached in comparison to the yoke guidance, this can however be traced down to known calibration issues with the system. A set of improvements specific for this issues was implemented as a result of this experiment.

Some participants reported ergonomic issues with the hardware, this was caused by the HTC Tracer being attached directly to the eye tracking glasses, which in turn were tightly secured to the front and back of the head of the pilot, causing a large pressure on the nose. Also, the HTC Tracker added weight to only one side of the head, further adding to the discomfort. This is regarded as a highly important note as ergonomics is fundamental in Human-Machine interfaces. A solution for this problem was implemented for this after the experiment reported in this paper and its impact will be assessed in the next round of experiments.

All participants gave positive feedback on the gaze guidance, suggesting that further development would result in a significant impact in reducing the workload induced by having to command the aircraft during a flight with appropriate visual conditions.

6. Acknowledgements

The authors would like to thank the research foundations and agencies that made this research possible through funding for equipment and scholarships, these are CAPES, FAPESP, CNPQ and FINEP. without their support this paper would have not been possible.

References

[1] Tobii. Pro Glasses 2 Product Description.

[2] NASA. Task Load Index (TLX) Manual.

[3] HTC. VIVE Tracker (2018) Developer Guidelines.

[4] Hevner, Alan R. A three cycle view of design science research. *Scandinavian Journal of Information Systems*, pp 87-92, 2007.

[5] Hansen, John. The Use of Gaze to Control Drones. *ETRA - Symposium on Eye Tracking Research Applications*. 2014.

[6] Kratchunova, Daniela. From Interface to Interspace: Conceptual Framework for Multimodal Flight Deck Controls. *International Conference on Virtual, Augmented and Mixed Reality*. 2016.

[7] Rodriguez Diaz, Manuel Alejandro. Initial Flight Simulation Testing of a Gesture-Based Interface for Flight Control. *Aerospace Technology Congress*. 2019.

Copyright Statement